Foundations of Nonlinear Algebra

John Perry
University of Southern Mississippi
john.perry@usm.edu
http://www.math.usm.edu/perry/

Table of Contents

Reference sheet for notation

$[r]$	the element $r+n\mathbb{Z}$ of $\mathbb{Z}_n$
$\langle g\rangle$	the group (or ideal) generated by g
A_3	the alternating group on three elements
$A\triangleleft G$	for G a group, A is a normal subgroup of G
$A\triangleleft R$	for R a ring, A is an ideal of R
$[G,G]$	commutator subgroup of a group G
$[x,y]$	for x and y in a group G, the commutator of x and y
$\text{Conj}_a(H)$	the group of conjugations of H by a
$\text{conj}_g(x)$	the automorphism of conjugation by g
D_3	the symmetries of a triangle
$d \mid n$	d divides n
$\deg f$	the degree of the polynomial f
D_n	the dihedral group of symmetries of a regular polygon with n sides
$D_n(\mathbb{R})$	the set of all diagonal matrices whose values along the diagonal is constant
$d\mathbb{Z}$	the set of integer multiples of d
$d\mathbb{Z}$	the set of integer multiples of d
$f(G)$	for f a homomorphism and G a group (or ring), the image of G
$\mathbb{F}(\alpha)$	field extension of $\mathbb{F}$ by $alpha$
$\text{Frac}(R)$	the set of fractions of a commutative ring R
F_S	the set of all functions mapping S to itself
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
$G\cong H$	G is isomorphic to H
$\text{GL}_m(\mathbb{R})$	the general linear group of invertible matrices
$\prod_{i=1}^n G_i$	the ordered n-tuples of $G_1, G_2, \ldots, G_n$
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
$\ker f$	the kernel of the homomorphism f
$\text{lcm}(t,u)$	the least common multiple of the monomials t and u
$\text{lm}(p)$	the leading monomial of the polynomial p
$\text{lv}(p)$	the leading variable of a linear polynomial p
$\mathbb{M}$	the set of monomials in one variable
$\mathbb{M}_n$	the set of monomials in n variables
$N_G(H)$	the normalizer of a subgroup H of G
$\mathbb{N}$	the natural numbers $\{0,1,2,\ldots\}$
$\mathbb{N}^+$	positive integers
Ω_n	the nth roots of unity; that is, all roots of the polynomial x^n-1
$\text{ord}(x)$	the order of x
$P(S)$	the power set of S
Q_8	the group of quaternions
R/A	for R a ring and A an ideal subring of R, R/A is the quotient ring of R with respect to A

$\langle r_1, r_2, \ldots, r_m \rangle$	the ideal generated by $r_1, r_2, \ldots, r_m$
$\mathbb{R}^{>0}$	the set of positive real numbers, a group under multiplication
$R[x_1, x_2, \ldots, x_n]$	the ring of polynomials whose coefficients are in the ground ring R
S_n	the group of all permutations of a list of n elements
$S \times T$	the Cartesian product of the sets S and T
$\mathcal{T}_f$	the support of a polynomial f
T_f	the support of the polynomial f
$\mathrm{tts}(p)$	the trailing terms of p
ω	typically, a primitive root of unity
$Z(G)$	centralizer of a group G
$\mathbb{Z}_n^*$	the set of elements of $\mathbb{Z}_n$ that are *not* zero divisors
$\mathbb{Z}/n\mathbb{Z}$	quotient group (resp. ring) of $\mathbb{Z}$ modulo the subgroup (resp. ideal) $n\mathbb{Z}$
$\mathbb{Z}$	integers
$\mathbb{Z}\left[\sqrt{-5}\right]$	the ring of integers, adjoin $\sqrt{-5}$
$\mathbb{Z}_n$	the quotient group $\mathbb{Z}/n\mathbb{Z}$

A dubious attempt at a narrative hook

We'd like to motivate this study of algebra with some problems that algebraists find interesting. Although we eventually solve them in this text with the aid of algebraic ideas, it might surprise you that, in this class, we're interested not in *how* to solve these problems, but in *why* the solutions work. I could in fact content myself with telling you how to to solve them, and we'd be done soon enough; on to vacation! But then you wouldn't have learned what makes algebra so beautiful *and important.* Think of the difference between spending a day touring a museum, and spending a summer in one of its programs; signs and guides can summarize each exhibit, but you can't really appreciate the importance of most of its items after a quick stroll. This course aims to help you grow beyond being a mere visitor to the museum of mathematics: ideally, your explorations would one day turn up artifacts that would also appear in the museum!

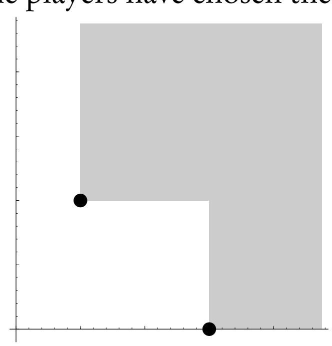
Ideal Nim

The following game generalizes the ancient game of Nim. Our playing board is the first quadrant of the x-y axis. Players first agree on a finite set of points which defines a Forbidden Frontier; that is, any point "southwest" of them is unplayable. (We say that (c,d) is "southwest" of (a,b) if $c < a$ and $d < b$.) They then take turns doing the following:

1. Choose a point (a,b) such that a and b are both integers, and that does not yet lie in a shaded region.
2. Shade the region of points "northeast" of (a,b). (We say that (c,d) is "northeast" of (a,b) if $c \geq a$ and $d \geq b$. Note the equality.)

The winner is the player who makes the last move.

In the example shown below, the players have chosen the points $(1,2)$ and $(3,0)$.



Questions:

- Must the game end? or is it possible to have a game that continues indefinitely? Is this true even if we use an n-dimensional playing board, where $n > 2$? And if so, why?
- Is there a way to count the number of moves remaining, even when there are infinitely many moves?
- Suppose that for each nonnegative integer d, you are forbidden from picking a certain number of points (a,b) such that $a + b = h(d)$, where h is some function of nonnegative integers. It doesn't matter what the points are, only that you may choose a certain number, and no more. Is there a strategy to win?

We answer some of these questions at the end of Chapter 1.

A card trick

Take twelve cards. Ask a friend to choose one, to look at it without showing it to you, then to shuffle them thoroughly. Arrange the cards on a table face up, in rows of three. Ask your friend what column the card is in; call that number α.

Now collect the cards, making sure they remain in the same order as they were when you dealt them. Arrange them on a table face up again, in rows of four. It is essential that you maintain the same order; the first card you placed on the table in rows of three must be the first card you place on the table in rows of four; likewise the last card must remain last. The only difference is the column layout. Ask your friend again what column the card is in; call that number β.

In your head, compute $4\alpha - 3\beta$. If this number does not lie between 1 and 12 inclusive, add or subtract 12 until it is. Call the result x. Starting with the first card, and following the order in which you laid the cards on the table, count to card $\#x$. This will be the card your friend chose.

Mastering this trick takes only a little practice. *Understanding* it requires quite a lot of background! We get to it in Chapter 6.

Questions:

- Would the same trick work with two layouts of 2 columns and 6 columns?
- Could you do something similar with 10 cards? with 15?

Internet commerce

Let's go shopping!!! — No, wait. That's too inconvenient: gas is expensive; the Mississippi sun is hot; and the parking lot is crowded. So… *Let's go shopping… online!!!* Before the merchant sends your product, they'll want payment. This requires you to submit your credit card number. That's not he sort of thing you'd like anyone to be able to read, but once you press the "submit" button, it will bounce happily around a few computers on its way to the company's server. Some of those computers could belong to unsavory characters. Some could even belong to unsavory characters halfway across the world, *even if you and the merchant both reside in the same country!*

How can you communicate the information *securely?* The solution is *public-key cryptography*. The retailer's computer supplies your computer with a special number called an *encryption key*. The retailer broadcasts this in the clear over the internet, and anyone in the world can see it. What's more, anyone in the world can look up the method used to decrypt the message.

You might wonder, *How on earth is this secure?!?* Public-key cryptography works because the *decryption key* remains with the company, hopefully secret. Whew! …or not really. A snooper could reverse-engineer one of the most common techniques using a "simple" grade school procedure: factoring an integer into primes. Say, $35 = 5 \cdot 7$.

*How on earth is **this** secure?!?* Although the procedure is "simple", the size of the integers in current use is about 80 digits. Believe it or not, even a computer will take far too long to factor an 80 digit integer! So your internet commerce is completely safe… unless someone has figured out how to factor 80 digit numbers quickly.[1]

Factorization

How can we factor polynomials like $p(x) = x^6 + 7x^5 + 19x^4 + 27x^3 + 26x^2 + 20x + 8$? There are a number of ways to do it, but the most efficient ways involve *modular arithmetic*. We discuss the theory of modular arithmetic later in the course, but the general principle is like this: pretend

[1]In theory, such a method is known, using *quantum computing*. However, no one knows how to build a quantum computer that can factor 40 digit numbers. *Yet.*

that the only numbers we can use are those on a clock that runs from 1 to 51. As with the twelve-hour clock, when we hit the integer 52, we reset to 1; in general, for any number that does not lie between 1 and 51, we divide by 51 and take the remainder. For example,

$$20 \cdot 3 + 8 = 68 \rightsquigarrow 17.$$

How does this help us factor? When looking for factors of the polynomial p, we can simplify multiplication by working in this modular arithmetic. This makes it easy for us to reject many possible factorizations before we start. In addition, the set $\{1, 2, \ldots, 51\}$ has many interesting properties that we can exploit further.

Conclusion

Non-linear algebra deals with absorbing problems, many of which are quite practical, while others seem merely recreational. Both, however, retain a deep, theoretical character. You may be tempted on many occasions to ask yourself the point of all this abstraction and theory: *Who needs this stuff?* Keep the examples above in mind; they show that algebra is not only useful, but interesting and necessary. Its applications have been profound and broad. Eventually, you will see how algebra addresses the problems above; for now, you can only start to imagine.

Chapter 0:
Foundations

This chapter re-presents ideas you have seen before, but may not have acquired comfort with them. We will emphasize precise definitions and rely heavily on deductive precision, and we will move very quickly and tersely, because this is, after all, supposed to be review. This may make the material seem drier than you remember it the first time around.

Intuition is *very* important in the problem solving process, and you *will* have to develop some intuition to succeed with this material. We *will* emphasize intuitive notions as we introduce new terms. However, ordinary language is often quite vague, and words contain different meanings for different people. You should already have an intuitive familiarity with most of the ideas presented in this section, yet most of you will find it difficult to *deduce* a solution, let alone articulate it in *precise* words.

Gauss, no slouch in either mathematics or science, stated that mathematics is not merely a science, but the queen of the sciences. Good science depends on clarity and reproducibility. This can be hard going for a while, but if you accept it and engage it, you will find it very rewarding.

0.1: Sets and relations

We start with the most fundamental tools in our box.

Sets

A **set** is a collection of objects, called its **members** or **elements**. All sets have the **empty set** as a subset; some people write the empty set as $\{\}$, but we will use $\emptyset$, which is also common.

The sets most fundamental to practical mathematics are

- the **positive integers**, $\mathbb{N}^+ = \{1,2,3,\ldots\}$, also called the **counting numbers**, and
- the **natural numbers**, $\mathbb{N} = \{0,1,2,\ldots\}$.

Notation 0.1.

- If a is a member of a set A, we write $a \in A$.
- We say that a set A is a **subset** of a set B if every element of A is also an element of B. If A is a subset of B, we write $A \subseteq B$. Writing this backwards means the same thing: $B \supseteq A$. This property of **inclusion** can occur even if $A = B$, which is why we write a bar, just as $\leq$ means "less than or equal," so you can think of $\subseteq$ as meaning "subset or equal." Notice that $x \in A$ implies $x \in B$.
- If $A \subseteq B$ but $A \neq B$, we often say that A is a **proper subset** of B. We can still write $A \subseteq B$ in this case, but we may emphasize the inequality by using $A \subsetneq B$. Again, $x \in A$ implies $x \in B$.
- Keep in mind that $A \subsetneq B$ is not the same as $A \not\subseteq B$, which means that A is *not* a subset of B. In this case, we can find $a \in A$ such that $a \notin B$.

Let's review some basic ideas of sets, especially the integers.

Notation 0.2. Not only is $\mathbb{N}^+ \subseteq \mathbb{N}$ true, so is $\mathbb{N}^+ \subsetneq \mathbb{N}$.

Two sets are **equal** if they contain the same elements; that is,

$$S = T \quad \text{if and only if} \quad S \subseteq T \text{ and } S \supseteq T.$$

In addition, can put sets together in several ways. Given two sets S and T, you should already be familiar with their **union**,

$$S \cup T = \{x : x \in S \text{ or } x \in T\},$$

their **intersection**,

$$S \cap T = \{x : x \in S \text{ and } x \in T\},$$

and their **difference**,

$$S \backslash T = \{x : x \in S \text{ and } x \notin T\}.$$

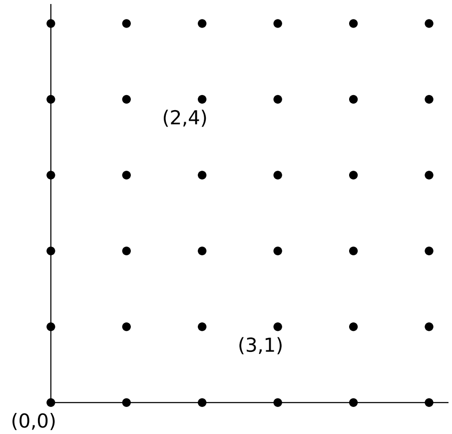
Another way to combine sets is by a set of **ordered pairs**, called a **Cartesian product**,

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

Example 0.3. Suppose $S = \{a, b\}$ and $T = \{x+1, y-1\}$. By definition,

$$S \times T = \{(a, x+1), (a, y-1), (b, x+1), (b, y-1)\}.$$

Example 0.4. Let $S = T = \mathbb{N}$. By substitution, $S \times T = \mathbb{N} \times \mathbb{N}$; the ordered pairs in this Cartesian product consist exclusively of natural numbers. We can visualize this as a **lattice**, where points must have integer co-ordinates:



Let $\mathcal{B} = \{\mathbb{N}^+, \{0\}\}$. The elements of $\mathcal{B}$ are **disjoint** sets, by which we mean that they have nothing in common. In addition, the elements of $\mathcal{B}$ **cover** $\mathbb{Z}$, by which we mean that their union produces the entire set of integers. This phenomenon, where a set can be described the union of smaller, disjoint sets, is important enough to highlight with a definition.

Definition 0.5. Suppose that A is a set and $\mathcal{B}$ is a family of subsets of A, called **classes**. We say that $\mathcal{B}$ is a **partition** of A if
- the classes cover A: that is, $A = \bigcup_{B \in \mathcal{B}} B$; and
- distinct classes are disjoint: that is, if $B_1, B_2 \in \mathcal{B}$ are distinct ($B_1 \neq B_2$), then $B_1 \cap B_2 = \emptyset$.

You will meet some important kinds of partitions in this course.

Relations

One use of a Cartesian product is to define any relationship between the elements of two sets.

Definition 0.6. Any subset of $S \times T$ is **relation on the sets** S **and** T. We call S the **domain** of the relation, and T the **range**. If $S = T$, we may call it a **relation on** S. A **function** is any relation F such that $(a,b) \in F$ implies $(a,c) \notin F$ whenever $c \neq b$.

Even though relations and functions are really just sets, they have a reputation for being more assertive than sets, which seem a rather passive bunch.

Notation 0.7.

- We typically denote relations that are not functions by curious symbols such as $<$ or $\subseteq$. We use $\sim$ as a generic symbol for a relation.
- We write $a \sim b$ instead of $(a,b) \in \sim$. For example, in a moment we will discuss the subset relation $\subseteq$, and we always write $S \subseteq T$ instead of "$(S,T) \in \subseteq$".
- We denote functions by letters, typically f, g, or h, or sometimes the Greek letters, η, φ, ψ, or μ.
- If f is a function, we write $f : S \to T$ instead of $f \subseteq S \times T$, and $f(s) = t$ instead of $(s,t) \in f$. In addition:
  - We call t the **image** of s, and s the **preimage** of t.
  - We define $f(S) = \{t \in T : f(s) = t \ \exists s \in S\}$. This is the **image** of S.

  So "image" has two different meanings: the "image of an element," and the "image of a set."

The definition of a function means that while $a \sim b$ and $a \sim c$ can both be true if $\sim$ is a relation, $f(a) = b$ and $f(a) = c$ cannot both be true if f is a function.

Example 0.8. Let $S = \{0,1,2,\ldots,9\}$ and $T = \{0,1,2,\ldots,19\}$. Define the relations $f, g : S \to T$ by

$$f = \{(0,0),(1,2),(2,4),\ldots,(9,18)\}$$

and

$$g = \{(0,0),(0,1),(1,1),(1,2),(2,2),(2,3),\ldots,(9,9),(9,10)\} .$$

Of the two, only f is a function, as $g(0) = 0$ and $g(0) = 1$.

Here we have defined S as the domain and T as the range of both f and g, but the image of S under f *is not* T. Rather, its image is

$$f(S) = \{0,2,4,\ldots,18\} .$$

The image of 9 under f is 18; the preimage of 18 is therefore 9. Notice that you can think of the "image" as a y-value, and the "preimage" as an x-value.

Operations

Definition 0.9. Let S, T, and U be sets. An **operation from** S **and** T **to** U is a function $f : S \times T \to U$. If $T \subseteq S$, we call f an operation "on" S. An operation f is **closed** if $T \subseteq S = U$.

For each set described above, you have learned a simple arithmetic: addition, subtraction, multiplication, and division. Let's give a concrete meaning to these operations.

We define ***addition of natural numbers*** in the usual way: the sum $a+b$ counts the number of objects in the union of two sets A and B where A has a elements and B has b elements.

Example 0.10. Addition of the natural numbers is a function, $+ : \mathbb{N}\times\mathbb{N}\to\mathbb{N}$. Here $S = T = \mathbb{N}$, so addition is an operation on $\mathbb{N}$. In fact, the sentence, $2+3=5$ can be thought of as $+(2,3)=5$. Hence, addition is an operation on $\mathbb{N}$. Addition is defined for all natural numbers, so it is closed.

We think of ***subtraction of natural numbers*** in terms of removing objects from a set. As you can probably imagine, this presents some difficulties. If you can't imagine why, don't worry: we make this idea precise using several steps, and the introduction of a new set.

- For any $a\in\mathbb{N}^+$, we define its *additive inverse* $-a$ as a new object with the property that $a+(-a)=0$.
  - Since $0+0=0$, we declare that $-0=0$. (It makes sense, anyway.)

  Now define the **integers** as[2]

  $$\mathbb{Z}=\mathbb{N}\cup\{-a : a\in\mathbb{N}\}=\{\ldots,-3,-2,-1,0,1,2,3,\ldots\}\ .$$

  Let's define addition on $\mathbb{Z}$. Let $a,b\in\mathbb{N}^+$ and consider $a+(-b)$:
  - If $a=b$, then substitution implies that $a+(-b)=b+(-b)=0$.
  - Otherwise, let A be any set with a objects.
    - If I can remove a set with b objects from A, then let $c\in\mathbb{N}$ be the number of objects left over. We define $a+(-b)=c$.
    - If I *cannot* remove a set with b objects from A, then let $c\in\mathbb{N}^+$ be the number of objects I would need to add to A so that I could remove b objects in as little time as possible. This satisfies the equation $a+c=b$; we then define $a+(-b)=-c$.
  - Finally, define $(-a)+b=-[a+(-b)]$; that is, $(-a)+b$ is the additive inverse of $a+(-b)$.

  In summary, $a+(-b)$ is either 0 (when a and b are the same), the number of objects left over from taking b elements from a, or the additive inverse of the number of objects we need to add to a set of a elements in order to take away b elements.

  This allows us to define **subtraction of integers** $- : \mathbb{Z}\times\mathbb{Z}\to\mathbb{Z}$ in the following way: for any $a,b\in\mathbb{Z}$, we say that

  $$a-b=a+(-b)\ .$$

  The result is always an integer, so addition and subtraction are both closed on $\mathbb{Z}$.

Example 0.11. We can also consider subtraction of natural numbers as an operation: $- : \mathbb{N}\times\mathbb{N}\to\mathbb{Z}$. Here $S=T=\mathbb{N}$ and $T=\mathbb{Z}$. The trouble with this is that while subtraction is an operation on $\mathbb{N}$, it is not closed on $\mathbb{N}$, because the range ($\mathbb{Z}$) is not the same as the domain ($\mathbb{N}$). Without integers, we cannot speak about subtraction "on" a set.

We define ***multiplication of integers*** as follows.

[2] You may be wondering why we denote the natural numbers by the eminently sensible $\mathbb{N}$, but the integers by $\mathbb{Z}$ instead of $\mathbb{I}$. The answer is that $\mathbb{Z}$ would be very sensible to a German, whose word for "number" is *zahlen*.

- Suppose $a \in \mathbb{N}^+$ and $b \in \mathbb{Z}$.
  - We say that $a \times b$ (also written $a \cdot b$ or ab) is the result of adding a copies of b, or

$$\underbrace{(((b+b)+b)+\cdots b)}_{a}.$$

  - We then define $(-a) \times b = -(a \times b)$. That is, $(-a) \times b$ is the additive inverse of $a \times b$.
  - As a consequence, for any $a \in \mathbb{Z}$, we have

$$a \cdot 0 = \underbrace{(((0+0)+0)+\cdots+0)}_{a}.$$

This, too, is an operation: $\times : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$. We have now defined addition and multiplication for nearly all integers, and they satisfy the following properties:
- $a+b=b+a$ and $ab=ba$ for all $a,b \in \mathbb{N}^+$ (the **commutative property**);
- $a+(b+c)=(a+b)+c$ and $(ab)c=a(bc)$ for all $a,b,c \in \mathbb{N}^+$ (the **associative property**); and
- $a(b+c)=ab+ac$ for all $a,b,c \in \mathbb{Z}$ (the **distributive property**).

We will not prove that they satisfy those properties, but the exercises will ask you to explain them using some examples. How might you do that?

Example 0.12. Let's prove that $5+(-7)=(-7)+5$.
- First we claim that $5+(-7)=-2$.
  - Let A be a set of 5 elements.
  - We cannot take 7 elements from A. (Try it if you doubt me.) So I have to add objects to A.
  - The quickest way to add objects to A in order to remove 7 objects is to add 2 objects. (Try it if you doubt me: adding 3 objects always takes a little longer than adding 2.)
  - So $5+2=7$, and that means $5+(-2)=-7$.
- Now we claim that $(-7)+5=-2$.
  - By definition, $(-7)+5=-[7+(-5)]$, so we can look at that instead.
  - Let A be a set of 7 elements.
  - Taking 5 elements away from A leaves 2, so $7+(-5)=2$.
  - By substitution, $-\left[\underbrace{7+(-5)}\right] = -\left[\underbrace{2}\right] = -2$.
  - Also by substitution, $(-7)+5=-[7+(-5)]=-2$.
- We have shown that $5+(-7)$ and $(-7)+5$ have the same value, -2. Hence $5+(-7)=(-7)+5$.

Unfortunately, this *only* proves the commutative property for 5 and -7.

Notation 0.13. For convenience, we usually write $a-b$ instead of $a+(-b)$.

We have not yet talked about the additive inverses of additive inverses. Suppose $b \in \mathbb{Z} \backslash \mathbb{N}$; by definition, b is an additive inverse of some $a \in \mathbb{N}^+$, $a+b=0$, and $b=-a$. Since we want addition to satisfy the commutative property, we *must* have $b+a=0$, which suggests that we can think of a as the additive inverse of b, as well! That is, $-b=a$. Written another way, $-(-a)=a$.

This also allows us to define the **absolute value** of an integer,

$$|a| = \begin{cases} a, & a \in \mathbb{N}, \\ -a, & a \notin \mathbb{N}. \end{cases}$$

Equivalence relations

Certain relations enjoy a property that we will use quite often.

Definition 0.14. 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 $a \sim c$.

Example 0.15. 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 reflexive.

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. In fact, *it is not,* and what's more, it illustrates an important principle of mathematical investigation. If you have difficulty proving a statement is true, it's worth asking whether you've found an instance *where the statement is false.*

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 $-1 \sim 0$. By the symmetric property, then, $0 \sim -1$. If the transitive property held, we would have $1 \sim -1\ldots$ but the definition of the the relation then requires $1 \times (-1) \in \mathbb{N}$, which is false! The relation is *not* transitive, so $\sim$ *cannot* be an equivalence relation on this set!

Orderings

So far we have avoided referring to any number as "less than" another. For instance, in the discussion of $a + (-b)$, we referred to talked about the number of objects that you could add quickest, and appealed to intuition or experience. But we did not actually say that $2 < 3$, for instance.

We could appeal again to intuition and experience and say that $a < b$ if it takes less time to create a set with a elements than a set with b elements. However, this only takes care of natural numbers; it does not define an ordering on negative numbers.

In fact, were we to extend this approach to negative numbers, we might be tempted to argue that $-2 < -3$. We might then order the integers as

$$\{-1, -2, -3, \ldots, 0, 1, 2, 3, \ldots\} .$$

Believe it or not, this is a perfectly legitimate ordering, which has its advantages! As you know, however, this is not the usual ordering on negative numbers; typically we prefer

$$\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\} .$$

This ordering is actually related to how we defined the sum of a positive and a negative number—in other words, to subtraction.

Definition 0.16. For any two elements $a, b \in \mathbb{Z}$, we say that:
- $a \leq b$ if $b - a \in \mathbb{N}$;
- $a > b$ if $b - a \notin \mathbb{N}$;
- $a < b$ if $b - a \in \mathbb{N}^+$;
- $a \geq b$ if $b - a \notin \mathbb{N}^+$.

So $2 < 3$ because $3 - 2 \in \mathbb{N}^+$ and $3 \geq 2$ because $2 - 3 \notin \mathbb{N}^+$. Notice how the negations work: the negation of $<$ is $\leq$, *not* $>$.

Remark 0.17. Do not yet assume certain "natural" properties of these orderings. For example, it is true that if $a \leq b$, then either $a < b$ or $a = b$. But why? You can reason to it from the definitions given here, so you should do so.

More importantly, you cannot yet assume that if $a \leq b$, then $a + c \leq b + c$. You can reason to this property from the definitions, and you will do so in the exercises.

Some orderings enjoy special properties.

Definition 0.18. Let S be any set. A **linear ordering** on S is a relation $\sim$ where the relation is:
1. *total:* for distinct $a, b \in S$ one of the following holds:
$$a \sim b \quad \text{or} \quad b \sim a.$$
2. *transitive:* for any $a, b, c \in S$
$$[\quad a \sim b \quad \wedge \quad b \sim c \quad] \qquad \Longrightarrow \qquad a \sim c \quad .$$
(Here, $\wedge$ means "and.")
3. *anticyclic:* if $a \sim b$ and $b \sim a$, then $a = b$.

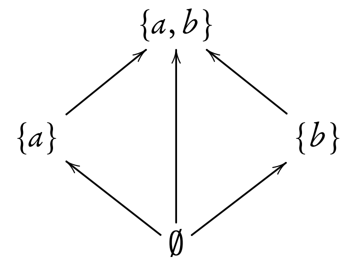
Remark 0.19. A couple of remarks on the terminology used in the definition:
- Remember that "distinct" means $a \neq b$. If $a = b$ then we can have both $a \not\sim b$ and $b \not\sim a$.
- You will more commonly find the name "antisymmetric" for what we call "anticyclic", but "antisymmetric" confuses students, in part because it doesn't seem particularly, well, anti-

symmetric: you still have $a \sim b$ and $b \sim a$, which is in fact a symmetry even if ultimately $a = b$. Besides, the point of the property is to avoid a cycle along the lines of $a < b < a$.

Not all orderings are linear! Suppose we order the subsets of a set $S = \{a, b\}$ by inclusion; that is, $A \sim B$ if and only if $A \subseteq B$. This ordering is *not* linear, because it is not total:

$$\{a\} \not\subseteq \{b\} \quad \text{and} \quad \{b\} \not\subseteq \{a\}.$$

We can visualize this problem if we diagram the ordering: (the arrow $A \to B$ means $A \subseteq B$)



The lack of an arrow from $\{a\}$ to $\{b\}$ or vice versa shows clearly that the two cannot be ordered. By contrast, a diagram makes it clear that the subsets of $T = \{a\}$ can be ordered linearly:

$$\emptyset \longrightarrow \{a\}$$

…although this is a rather boring set.

In any case, the orderings of $\mathbb{Z}$ *are* linear.

Theorem 0.20. The relations $<$, $>$, $\leq$, and $\geq$ are linear orderings of $\mathbb{Z}$.

Our "proof" must rely on some unspoken assumptions: in particular, the arithmetic on $\mathbb{Z}$ that we described before. Try to identify where these assumptions are used, because when you write your own proofs, you have to ask yourself constantly: Where am I using unspoken assumptions? In such places, either the assertion must be something accepted by the audience,[3] or you have to cite a reference your audience accepts, or you have to prove it explicitly. It's beyond the scope of this course to discuss these assumptions in detail, but you should at least try to find them.

Proof. We show that $<$ is linear; the rest are proved similarly. First we show that $<$ is total. □

[An aside. Watch how we reason through the proof; almost everything relies on definitions. We start off looking at the definition of the thing we have to prove: that $<$ is total. Since its definition makes a statement that must hold "for any $a, b \in S$," we must start our proof with two arbitrary elements of S. In this case, $S = \mathbb{Z}$, so we must start with arbitrary $a, b \in \mathbb{Z}$. ("Arbitrary" means no conditions are placed on the objects; here, that means we know nothing about a and b aside from their membership in $\mathbb{Z}$.) We now use the definition of $<$ to show that a and b must indeed satisfy the propert of a linear ordering that $a < b$, $a = b$, or $b < a$. (We get this by replacing $\sim$ in Definition 0.18 by $<$.)]

Proof. (resumed). Let $a, b \in \mathbb{Z}$ be distinct. Subtraction is closed for $\mathbb{Z}$, so $b - a \in \mathbb{Z}$. By definition, $\mathbb{Z} = \mathbb{N}^+ \cup \{0\} \cup \{-1, -2, \ldots\}$. Since $b - a$ must be in one of those three subsets, let's consider each case.

[3] In your case, the *instructor* is the audience.

[Another aside. How did we know to split $\mathbb{Z}$ into three subsets? Here's the thing: *we didn't.* What we did was take a gander at the definition of $<$, which considers $\mathbb{N}^+$ and $\mathbb{N}$. That lets us make an educated guess that we need to consider $\mathbb{N}^+$ as one case and $\{0\}$, which is what's left of $\mathbb{N}$, as another case. That leaves the negatives as a third case. Having settled on that, off we go exploring. Maybe it works, maybe it doesn't, but looking at the definition gave us an educated guess to start with.]

Proof. (resumed anew). ...let's consider each case.

- If $b - a \in \mathbb{N}^+$, then $a < b$.
- If $b - a = 0$, then $b = a$, so $a = b$.
- Otherwise, $b - a \in \{-1, -2, \ldots\}$. By definition, $-(b-a) \in \{1, 2, \ldots\} = \mathbb{N}^+$.

[A final aside. What motivated us to switch from $b - a$ to $-(b-a)$? The definition of $<$ doesn't tell us what to do with $\{-1, -2, \ldots\}$, so we're a bit stuck. We notice, though, that these numbers are the opposite of $\mathbb{N}^+$, regarding which the definition of $<$ tells us a great deal.]

Proof. (conclusion of "distinguish"; proofs of "transitive" and "antisymmetric").

We know that $-(b-a) = a - b$. So in fact $a - b \in \mathbb{N}^+$, and thus $b < a$.

We have shown that $a < b$, $a = b$, or $b < a$. Since a and b were arbitrary in $\mathbb{Z}$, $<$ is total.

We now check that $<$ is transitive. Let $a, b, c \in \mathbb{Z}$ and suppose that $a < b$ and $b < c$. By definition of $<$, we have $b - a, c - b \in \mathbb{Z}$. By algebra,

$$c - a = (c - b) + (b - a) \ ,$$

but the sum of two positive numbers is also positive, so $c - a \in \mathbb{N}^+$. By definition, $a < c$.

We conclude by showing that $<$ is anticyclic. Let $a, b \in \mathbb{Z}$ and suppose that $a < b$ and $b < a$. By definition of $<$, we have $b - a, a - b \in \mathbb{N}^+$. But this is a contradiction: no positive number has a positive "opposite." *In other words, there are no such $a, b \in \mathbb{Z}$!* So the ordering is vacuously anticyclic. □

[A final aside. A statement is "vacuously" true when its domain is the empty set; that is, you can say anything at all about the empty set's elements, and the statement is true. That's what happens in the proof of antisymmetry above. It is not often the case that a statement is vacuously true, but it does sometimes happen. With which of the orderings $\leq$, $\geq$, and $>$ do you think this will happen again? (It happens for at least one, but not all.)]

Linear orderings are already special, but some are *extra* special.

Definition 0.21. Let S be a set and $\prec$ a linear ordering on S. We say that $\prec$ is a **well-ordering** if

Every nonempty subset T of S has a **smallest element** a;
that is, there exists $a \in T$ such that for all $b \in T$, $a \prec b$ or $a = b$.

Example 0.22. The relation $<$ is *not* a well-ordering of $\mathbb{Z}$, because $\mathbb{Z}$ itself has no smallest element under the ordering.

Why not? Proceed by way of contradiction. Assume that $\mathbb{Z}$ has a smallest element, and call it a. Certainly $a-1 \in \mathbb{Z}$ also, but

$$(a-1)-a=-1 \notin \mathbb{N}^{+},$$

so $a \nless a-1$. Likewise $a \neq a-1$. This contradicts the definition of a smallest element, so $\mathbb{Z}$ is not well-ordered by $<$.

[A new aside. How did we know to look at $(a-1)-a$? Again, *we didn't*. We tried a proof by contradiction because we were trying to prove something false, and contradiction is a useful tool for that approach. Beyond that, we know from the definition of $<$ that we should look at the subtraction of two numbers. We need to find a smaller number, and at that point we use our intuition.]

We now assume, *without proof*, the following principle.

The relations $<$ and $\leq$ are well-orderings of $\mathbb{N}$.

That is, any subset of $\mathbb{N}$, ordered by these orderings, has a smallest element. This may sound obvious, but it is very important, and what is remarkable is that *no one can prove it.*[4] It is an assumption about the natural numbers. This is why we state it as a principle (or axiom, if you prefer). In the future, if we talk about the well-ordering of $\mathbb{N}$, we mean the well-ordering $<$.

One consequence of the well-ordering property is the following fact.

Theorem 0.23. Let $a_1 \geq a_2 \geq \cdots$ be a nonincreasing sequence of natural numbers. The sequence eventually stabilizes; that is, at some index i, $a_i = a_{i+1} = \cdots$.

[We won't interrupt this proof with asides. Instead, ask yourself why we took each step. Try to find the motivation *in the definitions of the terms themselves*.]

Proof. Let $A = \{a_1, a_2, \ldots\}$. By definition, $A \subseteq \mathbb{N}$. By the well-ordering principle, A has a least element; call it b. Let $i \in \mathbb{N}^{+}$ such that $a_i = b$. The definition of the sequence tells us that $b = a_i \geq a_{i+1} \geq \cdots$. Thus, $b \geq a_{i+k}$ for all $k \in \mathbb{N}$. Since b is the *smallest* element of A, we know that $a_{i+k} \geq b$ for all $k \in \mathbb{N}$. We have $b \geq a_{i+k} \geq b$, which is possible only if $b = a_{i+k}$. Thus, $a_i = a_{i+1} = \cdots$, as claimed. □

[Did you catch them? We decided to use the well-ordering property because it's a property of $\mathbb{N}$, the natural numbers, and when we're drowning we cast about for any sort of lifesaver. We used the a's to define a set A because the well-ordering property requires us to reason from a set of natural numbers, and hey, A is as handy a name for a set as any other, especially if it contains a's. (I do that a lot in this text.) Since we're using the well-ordering property, we might as well use it on A, and calling the smallest element b helps us distinguish it. After that, we think about what b must be: an element of A has to have the form a_i for some $i \in \mathbb{N}^{+}$. That more or less does the trick.]

[4] You might try to prove the well-ordering of $\mathbb{N}$ using induction. You would in fact succeed, but that requires you to assume induction. Why is induction true? There's a good question! It turns out that you cannot prove that induction is true without assuming the well-ordering of $\mathbb{N}$. In other words, *well-ordering is equivalent to induction:* each implies the other.

Another consequence of the well-ordering property is the principle of:

Theorem 0.24 (Mathematical Induction)**.** Let P be a subset of $\mathbb{N}^+$. If P satisfies (IB) and (IS) where
(IB) $1 \in P$;
(IS) for every $i \in P$, we know that $i+1$ is also in P;
then $P = \mathbb{N}^+$.

There are several versions of mathematical induction that appear: generalized induction, strong induction, weak induction, etc. We present only this one as a theorem, but we use the others without comment.

Proof. Let $S = \mathbb{N}^+ \backslash P$. We prove the contrapositive, so assume that $P \neq \mathbb{N}^+$. Thus $S \neq \emptyset$. Note that $S \subseteq \mathbb{N}^+$. By the well-ordering principle, S has a smallest element; call it n.

- If $n = 1$, then $1 \in S$, so $1 \notin P$. Thus P does not satisfy (IB).
- If $n \neq 1$, then $n - 1 \in \mathbb{N}^+$, so $n > 1$ by the properties of arithmetic. Since n is the smallest element of S and $n - 1 < n$, we deduce that $n - 1 \notin S$. Thus $n - 1 \in P$. Let $i = n - 1$; then $i \in P$ and $i + 1 = n \notin P$. Thus P does not satisfy (IS).

We have shown that if $P \neq \mathbb{N}^+$, then P fails to satisfy at least one of (IB) or (IS). This is the contrapositive of the theorem's statement. If the contrapositive is true, so is the statement itself, and we are done. □

Induction is an enormously useful tool, and we will make use of it from time to time. You may have seen induction stated differently, and that's okay. There are several kinds of induction which are all equivalent. We use the form given here for convenience.

Exercises.

In this first set of exercises, we assume that you are not terribly familiar with creating and writing proofs, so we provide a few outlines, leaving blanks for you to fill in. As we proceed through the material, we expect you to grow more familiar and comfortable with thinking, so we provide fewer outlines, and in the outlines that we do provide, we require you to fill in the blanks with more than one or two words.

Hints to many of the exercises start on p. 419. We'll provide a few more hints in this section than later.

Exercise 0.25.
(a) Let $\sim$ be a relation on $\mathbb{Z}$ such that $a \sim b$ if $a + b \in \mathbb{N}$. Is this an equivalence relation? Why or why not?
(b) Let $\sim$ be a relation on $\mathbb{Z}$ such that $a \sim b$ if $a - b \in \mathbb{N}$. Is this an equivalence relation? Why or why not?

Exercise 0.26. Show that $a \leq |a|$ for all $a \in \mathbb{Z}$.

Exercise 0.27.
(a) Show that $0 < a$ for any $a \in \mathbb{N}^+$.
(b) Show that $0 \leq a$ for any $a \in \mathbb{N}$.

Claim: The relation $\leq$ is total on $\mathbb{Z}$.
Proof:
1. Let $a, b \in$_____. We consider three cases.
2. If $a - b \in \mathbb{N}$, then $b \leq a$ by _____.
3. Otherwise, $a - b \in \mathbb{Z}\backslash\mathbb{N}$. By definition of opposites, $b - a \in$_____.
4. Then $a \leq b$ by _____.
5. Since a and b were arbitrary in _____, we have shown that $\leq$ is total.

Figure 0.1. Material for Exercise 0.28

Exercise 0.28.
(a) Fill in each blank of Figure 0.1 with the justification.
(b) Why is it a good idea to look at the values of $a - b$ and $b - a$?
(c) What is $\mathbb{Z}\backslash\mathbb{N}$?
(d) Show that $\leq$ is also transitive and anticyclic, making it linear.

Exercise 0.29. Let $a \in \mathbb{Z}$. Use the definitions of $<$ and $\leq$ to show that:
(a) $a < a + 1$;
(b) for any $b \in \mathbb{N}$, $a \leq a + b$.

Exercise 0.30. Let $a, b, c \in \mathbb{Z}$.
(a) Prove that if $a \leq b$, then $a = b$ or $a < b$.
(b) Prove that if both $a \leq b$ and $b \leq a$, then $a = b$.
(c) Prove that if $a \leq b$ and $b \leq c$, then $a \leq c$.

Exercise 0.31. Let $a, b \in \mathbb{N}$ and assume that $0 < a < b$. Let $d = b - a$. Show that $d < b$.

Exercise 0.32. Let $a, b, c \in \mathbb{Z}$ and assume that $a \leq b$. Prove that
(a) $a + c \leq b + c$;
(b) if $a \in \mathbb{N}$ and $c \in \mathbb{N}^+$, then $a \leq ac$; and
(c) if $c \in \mathbb{N}^+$, then $ac \leq bc$.

Exercise 0.33. Let $S \subseteq \mathbb{N}$. We know from the well-ordering property that S has a smallest element. Prove that this smallest element is unique.

Exercise 0.34. Let $\lessdot$ be an ordering of $\mathbb{Z}$ such that

$$0 \lessdot -1 \lessdot -2 \lessdot -3 \lessdot \cdots \lessdot 1 \lessdot 2 \lessdot 3 \lessdot \cdots \quad .$$

That is, $\lessdot$ orders 0 smaller than any other integer; any negative smaller than any positive; the positives in the usual way; and the negatives in an unusual way. This is "clearly" a linear ordering: why? Show that it is also a well ordering.

Exercise 0.35. The usual ordering $\prec$ of $\mathbb{Q}$ defines for any $a/b, c/d \in \mathbb{Q}$

$$\frac{a}{b} \prec \frac{c}{d} \iff ad < bc.$$

Let S be a well-ordered set.
Claim: Every strictly decreasing sequence of elements of S is finite.
Proof:

1. Let $a_1, a_2, \ldots \in$_____.
   (a) Assume that the sequence is _____.
   (b) In other words, $a_{i+1} < a_i$ for all $i \in$_____.
2. By way of contradiction, suppose the sequence is _____.
   (a) Let $A = \{a_1, a_2, \ldots\}$.
   (b) By definition of _____, A has a smallest element. Let's call that smallest element b.
   (c) By definition of _____, $b = a_i$ for some $i \in \mathbb{N}^+$.
   (d) By _____, $a_{i+1} < a_i$.
   (e) By definition of _____, $a_{i+1} \in A$.
   (f) This contradicts the choice of b as the _____.
3. The assumption that the sequence is _____ is therefore not consistent with the assumption that the sequence is _____.
4. As claimed, then, _____.

Figure 0.2. Material for Exercise 0.37

Here, $<$ refers to the ordering we defined for $\mathbb{Z}$. (This makes sense because ad and bc are integers.) Also, $\iff$ is a shortcut for writing "if and only if." Definitions are always $\iff$, even if we only say "if."

(a) Show that $\prec$ is a linear ordering.
(b) Show that $\prec$ is compatible with the ordering on $\mathbb{Z}$. Here, "compatible" means if $a, b \in \mathbb{Z}$, then $a, b \in \mathbb{Q}$, and we'd like $a < b$ to be true if and only if $a \prec b$.)

Exercise 0.36. By definition, a function is a relation. Can a function be an equivalence relation?

Exercise 0.37.
(a) Fill in each blank of Figure 0.2 with the justification.
(b) Why should a good proof start the way this one does: with an arbitrary sequence of elements of S?
(c) Why can the proof assume that $a_{i+1} < a_i$?
(c) Why would someone want to look at the smallest element of A?

Definition 0.38. Let $f : S \to U$ be a mapping of sets.

- We say that f is **one-to-one** if for every $a, b \in S$ where $f(a) = f(b)$, we have $a = b$.
- We say that f is **onto** if for every $x \in U$, there exists $a \in S$ such that $f(a) = x$.

Another way of saying that a function $f : S \to U$ is onto is to say that $f(S) = U$; that is, the **image** of S is *all* of U, or that *every* element of U corresponds via f to some element of S. Other ways of saying this is that the range is the image; or, every $u \in U$ has a preimage $s \in S$.

Exercise 0.39. Suppose that $f : S \to U$ is a one-to-one, onto function. Let $g : U \to S$ by

$$g(u) = s \iff f(s) = u.$$

(a) Show that g is also a one-to-one, onto function.
(b) Show that g *undoes* f, in the sense that for any $s \in S$, we have $g(f(s)) = s$.

This justifies the notion of an **inverse function**; if two functions f and g satisfy the relationship of Exercise 0.39, then each is the inverse function of the other, and we write $g = f^{-1}$ and $f = g^{-1}$. Notice how this implies that $f = \left(f^{-1}\right)^{-1}$.

Exercise 0.40. Let f be a function. Show that if f has an inverse function g, then that inverse function is unique.

0.2: Division

The fourth arithmetic "operation" is… not really an operation! See if you can spot why before we tell you.

The Division Theorem

Theorem 0.41 (The Division Theorem for Integers)**.** Let $n, d \in \mathbb{Z}$ with $d \neq 0$. There exist unique $q \in \mathbb{Z}$ and $r \in \mathbb{Z}$ satisfying (D1) and (D2) where
(D1) $n = qd + r$;
(D2) $0 \leq r < |d|$.
We call q the **quotient** and r the **remainder**.

Do you see the problem?

Recall the definition of an operation on a set:

> An **operation from** S **and** T **to** U is a function $f : S \times T \to U$. If $T \subseteq S$, we call f an operation "on" S.

Were division an "operation on the integers," it would take the form

$$\div \quad : \quad \mathbb{Z} \times T \to \mathbb{Z} \qquad ,$$

where $T \subseteq \mathbb{Z}$. This means that $\div$ should *only* result in *one* integer, but division results in *two* integers, rather than *one:* a quotient and a remainder. This means that the form of division is

$$\div \quad : \quad \mathbb{Z} \times T \to \mathbb{Z} \times \mathbb{Z} \qquad .$$

Here $S = \mathbb{Z}$ while $U = \mathbb{Z} \times \mathbb{Z}$. Since $S \neq T$, division of integers does not satisfy our notion of an operation.[5]

[5] This will not be a problem with division of other sets, but it is a problem here!

Proof. 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 q and r exist; second, that r is unique.

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

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

If you have trouble visualizing the sequence, it helps to write out an example.

Example 0.42. Suppose $n = 20$ and $d = 3$. Then $r_0 = 20$, and since $d = 3 \leq 20 = r_0$, we set $r_1 = 17$; since $d = 3 \leq 17 = r_1$, we set $r_2 = 14$, and so on until $r_6 = 2$. We no longer have $d \leq r_6$, so $2 = r_6 = r_7 = r_8 = \ldots$.

In *this* example, we eventually get all the r's to be the same. Is that always the case, regardless of the choice of n and d? That's what the first part of this proof is about.

We claim this sequence is nondecreasing. Why? If $r_{i+1} \neq r_i$, then by (1) we have $d \leq r_i$ and

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

Theorem 0.23 tells us every sequence nondecreasing natural numbers must stabilize; call the stable value r_q. If $d \leq r_q$, then by (1) $r_{q+1} \neq r_q$, contradicting the stability of r_q, so $r_q < d$. In addition, the definition of the sequence requires $r_q \in \mathbb{N}$, so $0 \leq r_q$.In other words, r_q satisfies (D2).

What about (D1)? Consider (1) carefully:

$$\begin{aligned}
r_1 &= n - d & &= n - 1 \cdot d \\
r_2 &= r_1 - d = (n-d) - d & &= n - 2 \cdot d \\
r_3 &= r_2 - d = (n - 2d) - d & &= n - 3 \cdot d \\
&\vdots \\
r_q &= r_{q-1} - d = [n - (q-1)\,d] - d = n - q \cdot d\ .
\end{aligned}$$

That last equation has $n = qd + r$, so q and r satisfy also satisfy (D1). We have shown the existence of q and r to satisfy the theorem!

Uniqueness of q and r: Let $q', r' \in \mathbb{Z}$ such that q' and r' satisfy (D1) and (D2); that is, $n = q'd + r'$ and $0 \leq r' < d$. If we can show that *both* $q = q'$ and $r = r'$, then the arbitrary choice of q' and r' implies that q and r are indeed unique. By substitution,

$$qd + r = n = q'd + r'$$

which we rewrite as

$$r - r' = (q' - q)\,d. \tag{2}$$

Now, this is interesting: the right hand side is an integer multiple of d. So $r - r'$ must also be an integer multiple of d. On the other hand, we have $0 \leq r, r' < d$, so what can $r - r'$ be? The extreme cases here are

- $r = 0$ and $r' = d - 1$, so that $r - r' = -d + 1 > d$; and
- $r = d - 1$ and $r' = 0$, so that $r - r' = d - 1 < d$.

Either way, $-d < r - r' < d$. So the only possible integer multiple of d that $r - r'$ can be is 0. That means $r = r'$. By substitution into (2), we have $0 = \left(q - q'\right) d$. If this happens, then the fact that $d \neq 0$ means $0 = q - q'$. We could explain why here, but it's important enough to separate into a separate fact, which you will see below. For our purposes, the point is that $q = q'$.

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

Definition 0.43. If the remainder from dividing n by d is 0, then $n = qd$. In this case, we say that d **divides** n, or that n is **divisible by** d. If we cannot find such an integer q, then d **does not divide** n.

In the past, you have probably heard of this as "divides evenly;" we omit "evenly".

Notation 0.44. If d divides n, we write $d \mid n$; otherwise, we write $d \nmid n$.

Remark 0.45. Why require $d \neq 0$? That is, why can we not divide by 0? Here are two reasons. First, think about the fact that division is repeated subtraction. In the theorem's proof:

- the theorem requires builds a sequence $n - d, n - 2d, \ldots$;
- if $d = 0$, then $n = n - d = n - 2d = \cdots$; so that
- this technique would never find q, r such that $n = qd + r$ unless $n = r$ already.

"Well," you might think, "that's a problem with your proof, not with the concept." Well, no, not really: the problem lies in the concept itself. If you could find $q, r \in \mathbb{Z}$ such that $n = qd + r$, the very fact that $d = 0$ means $n = q \cdot 0 + r = r$. Now, if $d \leq n = r$ (and often it is!) then we have an insurmountable problem: you can satisfy (D1), but not (D2).

The Zero Product Rule

We promised to explain why $0 = \left(q - q'\right) d$ and $d \neq 0$ implies $0 = q - q'$. We will actually explain something a little more general, called the **Zero Product Rule**.

Lemma 0.46. Let $a, b \in \mathbb{Z}$ and suppose $ab = 0$. Then $a = 0$ or $b = 0$.

How did we use this in the the theorem above? There we had $0 = \left(q - q'\right) d$, so $a = q - q'$ and $b = d$. We knew that $d \neq 0$, so $b \neq 0$. If $a \neq 0$, that would contradict the lemma we are about to prove, so we have to have $q - q' = a = 0$.

Proof. We actually prove the contrapositive. (Remember that? If not, go review it; we discussed it briefly in the previous section.)

Suppose both $a, b \neq 0$. We defined multiplication as repeated addition, so if $a \in \mathbb{N}^+$ then

$$ab = \underbrace{b + b + \cdots + b}_{a \text{ times}},$$

and otherwise

$$ab = -\Big(\underbrace{b + b + \cdots + b}_{|a| \text{ times}}\Big).$$

If $b \in \mathbb{N}^+$, these additions involves the union of sets. Since $b \neq 0$, the set is nonempty. You cannot obtain an empty set from a union of nonempty sets, so in this case $ab \neq 0$. On the other hand, if $b \in \mathbb{Z}\backslash\mathbb{N}$, then

$$ab = b + b + \cdots + b = -[(-b) + (-b) + \cdots + (-b)]$$

or

$$ab = -(b + b + \cdots + b) = -\{-[(-b) + (-b) + \cdots + (-b)]\} .$$

Either way, we are again adding positive numbers, so we end up with a nonempty set. In short, $ab \neq 0$.

We have shown that if both $a, b \neq 0$ then $ab \neq 0$. By the contrapositive, if $ab = 0$ then $a = 0$ or $b = 0$. □

Rational numbers

Until now we've considered the division of discrete, "atomic" objects, such as, a set of toys among three children. If you try to split one toy among three children, you're either going to break the toy or teach the children to share. You can't divide a toy truck into pieces that have equal utility to the children.

Suppose, however, you want to divide a continuous object, such as a pizza. This isn't quite the same problem now, is it? You can easily divide a pizza into pieces that satisfy the three little savages we mentioned a moment ago — in fact, you can do this in multiple ways. We typically model this situation with **rational numbers**, sometimes called the **fractions**, are

$$\mathbb{Q} = \{a/b : a, b \in \mathbb{Z} \text{ and } b \neq 0\} .$$

You can think of a rational number a/b as meaning "divide a bunch of pizzas into b pieces and take a of them." Notice that we denote the rational numbers by a curious choice of letter, $\mathbb{Q}$. The Germans aren't to blame for this one, but rather the Italians, in whose mellifluous tongue $\mathbb{Q}$ stands for *quozienti*, or "quotients".

Thinking again about our hungry children: you can split a pizza into three giant pieces, but that's probably not wise. You could split it instead into six or even nine pieces, and give each child two or three. They'd still be eating the same amoutn of pizza, right? That is,

$$\frac{1}{3} = \frac{2}{6} = \frac{3}{9} .$$

If you don't believe me, you're probably in the wrong math class, but since I've probably made you hungry anyway, try ordering three pizzas, and divide one into three equal slices, another into six equal slices, and the third into nine equal slices. Compare one slice from the three-slice pizza with two slices from the six-slice pizza, then again with three slices from the nine-slice pizza. If they aren't all the same, you've done something terribly wrong, but there is an upside: you can eat the slices, destroy the evidence, then go change your major.

So if you haven't dropped the class yet, notice that

$$\frac{1}{3} = \frac{2}{6} \quad \text{and} \quad 1 \cdot 6 = 2 \cdot 3 .$$

That's not an accident. We can't compare pizzas of different sizes, but if we divide them further so that they do have the same size, we can. So we divide each slice of the three-slice pizza into 6 slices each, giving us

$$\frac{1}{3} = \frac{1 \cdot 6}{\underbrace{3 \cdot 6}_{\text{each into 6!}}} = \frac{6}{18} ,$$

and then divide each slice of the six-slice pizza into 3 slices each, giving us

$$\frac{2}{6} = \frac{2 \cdot 3}{\underbrace{6 \cdot 3}_{\text{each into 3!}}} = \frac{6}{18} .$$

We have now divided our two pizzas into eighteen equally-sized slices, and we can see that the one piece of the first is now six pieces, just as the two pieces of the second is, too. Since the two denominators end up the same, we only have to look at what happens in the numerator: $1 \cdot 6 = 2 \cdot 3$. More generally,

$$\frac{a}{b} = \frac{c}{d} \quad \Longleftrightarrow \quad ad = bc .$$

We can use the same idea to explain addition and subtraciton of $a/b \pm c/d$ precisely:

- if $b = d$, the pieces are all the same size, so it's a straightforward matter to consider $a \pm c$, resulting in $(a \pm c)/b$; but
- if $b \neq d$, you can't work out how many you have if they're different sized pieces, so divide them until they are equally sized pieces, and add the number of pieces; that is,

$$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd} .$$

How about multiplication? Multiplying a rational by an integer is fairly straightforward; it works the same as with integers: just think of the pieces of pizza. Multiplying $2/3$ by $1/2$ is something rather new, though. We can handle it this way: our intuitive notion of "half" of something suggests that half of two slices of a three-slice pizza is one slice of a three-slice pizza. But this is equivalent to two slices of a six-slice pizza; that is,

$$\frac{2}{3} \cdot \frac{1}{2} = \text{“half” of } \frac{2}{3} = \frac{1}{3} = \frac{2}{6} = \frac{2 \cdot 1}{3 \cdot 2} .$$

This reasoning is easily extended, so that in general,

$$\frac{a}{b} \times \frac{1}{d} = \frac{a}{bd} ,$$

and from there we apply multiplication of integers to see that

$$\frac{a}{b}\times\frac{c}{d}=\left(\frac{a}{b}\times\frac{1}{d}\right)\times c=\frac{a}{bd}\times c=\frac{ac}{bd}\,.$$

We come to division: what does the expression $2/3 \div 1/2$ mean? We can think of it this way: *two* slices of a three-slice pizza is half of what? A bit of imagination shows that it's half of *four* slices of a three-slice pizza; that is,

$$\frac{2}{3}\div\frac{1}{2}=\frac{4}{3}\,.$$

Play around with this a bit and you will agree that, in general,

$$\frac{a}{b}\div\frac{1}{d}=\frac{ad}{b}\,.$$

Once you have that, you can get to

$$\frac{a}{b}\div\frac{c}{d}=\frac{ad}{bc}\,.$$

For any $a\in\mathbb{Z}$, we can view a as a rational number, as well. After all, if you have 10 pizzas, then you have 10 one-slice pizzas, so that

$$10=\frac{10}{1}\,.$$

Since every integer is also a rational number, but not every rational number is an integer, we have $\mathbb{Z}\subsetneq\mathbb{Q}$. We have built a chain of sets

$$\mathbb{N}^+\subsetneq\mathbb{N}\subsetneq\mathbb{Z}\subsetneq\mathbb{Q}\,,$$

extending each set with some useful elements. Not bad!

In the exercises, you will generalize the ordering $<$ to the set $\mathbb{Q}$.

Exercises.

Don't forget that there are hints to some problems, starting on page 419. They start to grow more sparse, however. If you're stuck, think about the definitions: manipulating them is often the key to solving these problems.

Exercise 0.47. Identify the quotient and remainder when dividing:
(a) 10 by -5;
(b) -5 by 10;
(c) -10 by 4;
(d) -10 by -4.
Don't forget that *the remainder must be nonnegative!*

Exercise 0.48. Addition, subtraction, multiplication, and division are all operations on $\mathbb{Q}$. For each operation, identify the values of S, T, and U that show that Definition 0.9 is satisfied.

Exercise 0.49. Prove that if $a\in\mathbb{Z}$, $b\in\mathbb{N}^+$, and $a\mid b$, then $a\le b$.

Let $a, b, c \in \mathbb{Z}$.
Claim: If a and b both divide c, then $\operatorname{lcm}(a, b)$ also divides c.
Proof:
1. Let $d = \operatorname{lcm}(a, b)$. By _____, we can choose q, r such that $c = qd + r$ and $0 \leq r < d$.
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 $0 \leq r < d$, and _____. By definition of lcm, $r = 0$.
8. By _____, $c = qd = q\operatorname{lcm}(a, b)$.
9. By definition of _____, $\operatorname{lcm}(a, b)$ divides c.

Figure 0.3. Material for Exercise 0.52

Exercise 0.50. Show that divisibility is transitive for the integers; that is, if $a, b, c \in \mathbb{Z}$, $a \mid b$, and $b \mid c$, then $a \mid c$.

Definition 0.51. 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 precise definition of the least common multiple 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 $b \mid n$).

Exercise 0.52.
(a) Fill in each blank of Figure 0.3 with the justification.
(b) One part of the proof claims that "A similar argument shows that $b \mid r$." State this argument in detail.

Exercise 0.53. 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) If a and b have the same sign, show that that $a \bowtie b$ if they have 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) Show that (b) is not true if a and b have different signs.
(d) 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) Show that every $a \in \mathbb{Q}$ an element of some class.

(e) Show that if $S_a \neq S_b$, then $S_a \cap S_b = \emptyset$.
In other words, $\{S_a\}_{a \in \mathbb{Q}}$ is a partition.

Exercise 0.54. In Lemma on page 19, we proved the Zero Product Rule for elements of $\mathbb{Z}$. Naturally, this is also true for elements of $\mathbb{Q}$. You can show this precisely; that is, don't argue from slices of pizza, but rather take two elements $x, y \in \mathbb{Q}$ and show that if $xy = 0$, then $x = 0$ or $y = 0$. You don't need to use the contrapositive; you can show this directly.

Exercise 0.55. Define an ordering $\prec$ on $\mathbb{Q}$ as follows: for any $a/b, c/d \in \mathbb{Q}$ we say that

$$\frac{a}{b} \prec \frac{c}{d} \quad \Longleftrightarrow \quad ad < bc \quad ,$$

where the comparison of ad and bc is done as ordinary integers.

(a) Use the ordering to sort the numbers $2/3$, $-2/3$, $1/6$, $-1/6$, $1/9$, $-1/9$, $3/4$, $-3/4$. Be sure to show all the comparisons you used.

(b) Show that if a/b and c/d are integers, then the orderings $\prec$ and $<$ agree. In other words, $\prec$ extends $<$, so we might as well just use $<$ as a symbol for both.

(c) Using the pizza analogy, explain why this ordering "makes sense." It might help to review and adapt the explanation we gave of when $a/b = c/d$.

0.3: Linear algebra

Linear algebra is the study of algebraic objects related to linear polynomials. It includes not only matrices and operations on matrices, but vector spaces, bases, and linear transformations. For the most part, we focus on matrices and linear transformations.

Matrices

Definition 0.56. An $m \times n$ **matrix** is a list of m lists (**rows**) of n numbers. If $m = n$, we call the matrix **square**, and say that the **dimension** of the matrix is m.

Notation 0.57. We write the jth element of row i of the matrix A as a_{ij}. If $a_{ij} = 0$ and we are especially lazy, then we often omit writing it in the matrix. If the dimension of A is $m \times n$, then we write $\dim A = m \times n$.

Example 0.58. 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 that its dimension is 3.

Definition 0.59. 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 0.60. We often write A^T for the transpose of A.

Example 0.61. If A is the matrix of the previous example, then

$$A^T = \begin{pmatrix} 1 & & \\ & 1 & 5 \\ 1 & & 1 \end{pmatrix}.$$

While non-square matrices are important, we consider mostly square matrices in this class, with the exception of $m \times 1$ matrices, also called **column vectors**. It is easy to define three operations for matrices:

- We *add* matrices by adding entries in the same row and column. That is, if A and B are $m \times n$ matrices and $C = A + B$, then $c_{ij} = a_{ij} + b_{ij}$ for all $1 \leq i \leq m$ and all $1 \leq j \leq n$. Notice that C is also an $m \times n$ matrix.
- We *subtract* matrices by subtracting entries in the same row and column.
- We *multiply* matrices a little differently. 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

  $$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \, ;$$

  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.

Example 0.62. If A is the matrix of the previous example and

$$B = \begin{pmatrix} 1 & 5 & -1 \\ & 1 & \\ & -5 & 1 \end{pmatrix},$$

then

$$\begin{aligned} AB &= \begin{pmatrix} 1 \cdot 1 + 0 \cdot 0 + 1 \cdot 0 & 1 \cdot 5 + 0 \cdot 1 + 1 \cdot -5 & 1 \cdot -1 + 0 \cdot 0 + 1 \cdot 1 \\ 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 0 & 0 \cdot 5 + 1 \cdot 1 + 0 \cdot -5 & 0 \cdot -1 + 1 \cdot 0 + 0 \cdot 1 \\ 0 \cdot 1 + 5 \cdot 0 + 1 \cdot 0 & 0 \cdot 5 + 5 \cdot 1 + 1 \cdot -5 & 0 \cdot -1 + 5 \cdot 0 + 1 \cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}. \end{aligned}$$

If we write I for this last matrix of three 1's on the diagonal, then something interesting happens:

$$AI = IA = A \quad \text{and} \quad BI = IB = B.$$

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 is an "identity" matrix. In particular, it's the identity of **multiplication**. Is there another identity matrix?

There is certainly another identity matrix for addition: a matrix filled with zeros. (You could probably have guessed that yourself.) But can there be another identity matrix for *multiplication?* In fact, there *cannot*. Rather than show this directly, we wait until Section 1.1; for now, we consolidate our current gains.

Notation 0.63.
- We write $\mathbf{0}$ (that's a **bold**[6] **zero**) for any matrix whose entries are all zero.
- We write I_n for the square matrix of dimension n satisfying
  - $a_{ii} = 1$ for any $i = 1, 2, \ldots, n$; and
  - $a_{ij} = 0$ for any $i \neq j$.

Now, a formal statement of the result.

Theorem 0.64. The zero matrix $\mathbf{0}$ is an identity for matrix addition. The matrix I_n is an identity for matrix multiplication.

Notice that there's a bit of imprecision in this statement. You have to infer from the statement that $n \in \mathbb{N}^+$, $\mathbf{0}$ is an $n \times n$ matrix, and we mean that $\mathbf{0}$ is an identity for addition when we're talking about other matrices of dimension n. We should *not* infer that the statement means that $\mathbf{0}$ is an identity for matrices of dimension $m + 2$; that would be silly, as the addition would be undefined. When reading theorems, you sometimes have to read between the lines.

Proof. Let A be a square matrix of dimension 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 $A + \mathbf{0} = A$. A similar argument shows that $\mathbf{0} + A = A$. Since A is arbitrary, $\mathbf{0}$ 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)

$$a_{ij} \cdot 1 + \sum_{k \neq j} a_{ik} \cdot 0 \quad .$$

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 **inverse** of A is any matrix B such that $A + B = \mathbf{0}$ (if B is an **additive inverse**) and $AB = I_n$ (if B is a **multiplicative inverse**). 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 is **invertible** if it has a multiplicative matrix.

Notation 0.65. We write the additive inverse of a matrix A and $-A$, and the multiplicative inverse of A as A^{-1}.

Example 0.66. The matrices A and B of the previous example are inverses; that is, $A = B^{-1}$ and $B = A^{-1}$.

We want one more property before we move on.

Theorem 0.67. Matrix multiplication is associative.

[6] Only *bold* zeroes can be matrices; cowardly zeroes must be scalars.

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

$$((AB)\,C)_{ij} = \sum_{\ell=1}^{s} \sum_{k=1}^{r} (a_{ik} b_{k\ell})\, c_{\ell j}. \tag{3}$$

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 numbers,

$$(A\,(BC))_{ij} = \sum_{k=1}^{r} \sum_{\ell=1}^{s} (a_{ik} b_{k\ell})\, c_{\ell j}. \tag{4}$$

The only difference between equations (3) and (4) 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! Addition of numbers is commutative, so the order of the terms does not matter; we have

$$((AB)\,C)_{ij} = (A\,(BC))_{ij}\ .$$

We chose arbitrary i and j, so this is true for all entries of the matrices. The matrices are equal, which means $(AB)\,C = A\,(BC)$, □

Linear transformations

Remember how we were able to view functions as both *passive sets* and *dynamic activities?* (We commented on it after Definition 0.6.) We can make the same observation about matrices:

not only are they *passive lists of numbers*, they can act as *dynamic functions* on other matrices. A common example of this is to consider the set D of $n \times 1$ column vectors. If M is an $n \times n$ matrix, we can define a function $f_M : D \to D$ by

$$f_M(\mathbf{x}) = M\mathbf{x}.$$

Read this as, "f_M maps $\mathbf{x}$ to the product of M and $\mathbf{x}$."

Example 0.68. If

$$M = \begin{pmatrix} 1 & & 1 \\ & 1 & \\ & 5 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix},$$

then

$$f_M(\mathbf{x}) = M\mathbf{x} = \begin{pmatrix} -1 \\ 3 \\ 13 \end{pmatrix}.$$

This function is a special example of what we call a *linear transformation*. To define it precisely, we have to use the term **vector space**. If you do not remember that term, or never learned it, first go slap whoever taught you linear algebra, then content yourself with the knowledge that, in this class, it will be enough to know that any set D of all possible column vectors with n rows is a vector space for any $n \in \mathbb{N}^+$. Whatever that is. Then go slap your former linear algebra teacher again.

Definition 0.69. Let V be a vector space over the real numbers $\mathbb{R}$, and f a function on V. We say that f is a **linear transformation** if it preserves

- *scalar multiplication*, that is, $f(av) = af(v)$ for any $a \in \mathbb{R}$ and any $v \in V$, and
- *vector addition*, that is, $f(u+v) = f(u) + f(v)$ for any $u, v \in V$.

Eventually, you will learn about a special kind of function that works very similarly to linear transformations, called a **homomorphism**. For now, let's look at the classic example of a linear transformation, a matrix.

Example 0.70. Recall M and $\mathbf{x}$ from Example 0.68. Let

$$\mathbf{y} = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}.$$

Using the definitions of matrix addition and matrix multiplication, you can verify that

$$M(\mathbf{x} + \mathbf{y}) = \begin{pmatrix} 4 \\ 3 \\ 15 \end{pmatrix},$$

and also

$$Mx + My = \begin{pmatrix} -1 \\ 3 \\ 17 \end{pmatrix} + \begin{pmatrix} 5 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 15 \end{pmatrix}.$$

Now let $a = 4$. Using the definitions of matrix and scalar multiplication, you can verify that

$$M(a\mathbf{x}) = \begin{pmatrix} -4 \\ 12 \\ 68 \end{pmatrix},$$

and also

$$a(M\mathbf{x}) = 4\begin{pmatrix} -1 \\ 3 \\ 17 \end{pmatrix} = \begin{pmatrix} -4 \\ 12 \\ 68 \end{pmatrix}.$$

The example does *not* show that f_M is a linear transformation, because we tested M only with particular vectors $\mathbf{x}$ and $\mathbf{y}$, and with a particular scalar a. To show that f_M is a linear transformation, you'd have to show that f_M preserves scalar multiplication and vector addition on *all* scalars and vectors. Who has time for that? There are infinitely many of them, after all! Better to knock it off with a theorem whose proof relies on symbolic, or "generic", structure.

Theorem 0.71. For any matrix A of dimension n, the function f_A on all $n \times 1$ column vectors defined by $f_A(\mathbf{x}) = A\mathbf{x}$ is a linear transformation.

Proof. Let A be a matrix of dimension n.

First we show that f_A preserves scalar multiplication. Let $c \in \mathbb{R}$ and $\mathbf{x}$ be an $n \times 1$ column vector. By definition of scalar multiplication, the element in row i of $c\mathbf{x}$ is cx_i. By definition of matrix multiplication, the element in row i of $A(c\mathbf{x})$ is

$$\sum_{k=1}^{m} [a_{ik}(cx_k)].$$

Apply the commutative, associative, and distributive properties of the field to rewrite this as

$$c\sum_{k=1}^{m} a_{ik}x_k.$$

On the other hand, the element in row i of $A\mathbf{x}$ is, by definition of matrix multiplication,

$$\sum_{k=1}^{m} a_{ik}x_k.$$

If we multiply it by c, we find that $A(c\mathbf{x}) = cA\mathbf{x}$, as claimed.

We leave it to you to show that f_A preserves vector addition; see Exercise 0.88. □

An important aspect of a linear transformation is the kernel.

Definition 0.72. The **kernel** of a linear transformation f is the set of vectors that are mapped to $\mathbf{0}$. In other words, the kernel is the set

$$\{\mathbf{v} \in V : f(\mathbf{v}) = \mathbf{0}\}.$$

Notation 0.73. We write $\ker f$ for the kernel of f. We also write $\ker M$ when we mean $\ker f_M$.

Example 0.74. Let

$$M = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

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, $\ker M$ has infinitely many elements with an easily described form. We can obtain this description in the following way:

Our task is to find every element of $\ker M$. By definition, any $\mathbf{x} \in \ker M$ satisfies the equation

$$M\mathbf{x} = \mathbf{0},$$

which we can rewrite as

$$\begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

If we simplify the product on the left-hand side, we have

$$\begin{pmatrix} x_1 + 5x_3 \\ x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Each row of the matrices corresponds to an equation:

$$\begin{aligned} x_1 + 5x_3 &= 0 \\ x_2 &= 0 \\ 0 &= 0. \end{aligned}$$

The third equation is a tautology,[7] so in this context it is useless. The second equation tells us that $x_2 = 0$ for any $\mathbf{x} \in \ker M$. The first tells us that $x_1 + 5x_3 = 0$ for any $\mathbf{x} \in \ker M$. This isn't

[7] A "tautology" is a fancy word for a statement that is true in every circumstance. While not useful here, tautologies can be useful sometimes; for instance, when you want to check whether $x = 2$ solves $2x - 8 = -4$, you substitute 2 for x and the left-hand side simplifies to the tautology $-4 = -4$, which means $x = 2$ is indeed a solution. Contrast

enough information to narrow down to *one* vector, because we can choose any value at all for x_3 and obtain a unique value for x_1. For instance, if $x_3 = 12$, then we solve

$$x_1 + 5 \times 12 = 0 \quad \Longrightarrow \quad x = -60$$

and we have found

$$\mathbf{x} = \begin{pmatrix} -60 \\ 0 \\ 12 \end{pmatrix}.$$

It is easy to verify that $\mathbf{x}$ is indeed in $\ker M$, since

$$M\mathbf{x} = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -60 \\ 0 \\ 12 \end{pmatrix} = \begin{pmatrix} 1 \times (-60) + 5 \times 12 \\ 1 \times 0 \\ 0 \times 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad .$$

Likewise, if $x_3 = -3$ then placing $x_1 = 15$ gives $\mathbf{x} \in \ker M$. On the other hand, if $x_3 = 0$ then placing $x_1 = 5$ gives $\mathbf{x} \notin \ker M$, regardless of the value of x_2.

This "ambiguity" is exactly how we know $\ker M$ has infinitely many solutions! Since x_3 can have any value at all, we pick a letter we want — typically t, but I'm feeling in a c'ish mood, so I'll use c — then solve the first equation for x_1. We have

$$x_1 = -5c\ ,$$

so whatever $c \in V$ we pick for x_3, we know that

$$\mathbf{x} = \begin{pmatrix} -5c \\ 0 \\ c \end{pmatrix} \in \ker M \quad .$$

Written a different way,

$$\ker M = \left\{ \mathbf{v} \in V : v = \begin{pmatrix} -5c \\ 0 \\ c \end{pmatrix} \exists c \in \mathbb{F} \right\}.$$

The kernel has important and fascinating properties, many of which we explore later in the course in a more general context.

Determinants

An important property of a square matrix A is its determinant, denoted by $\det A$. We won't explain *why* it's important here, beyond saying that it has the property of being **invariant** when you rewrite the matrix in certain ways (see, for example, Theorem 0.79). We don't even define it terribly precisely; we simply summarize what you ought to know:

- to every matrix, we can associate a unique scalar, called its **determinant**;

---

this to a contradiction, which is true in no circumstances: when you want to check whether $x = 2$ solves $2x - 6 = -4$, you substitute 2 for x and the left-hand side simplifies to the contradiction $-2 = -4$, which means $x = 2$ is not a solution.

- we can compute the determinant using a technique called *expansion by minors* along any row or column; and
- the determinant enjoys a number of useful properties, some of which are listed below.

Example 0.75. Recall the matrix A from Example (0.58). If we expand by minors on the first row, we find that

$$\begin{aligned}\det A &= 1\cdot(-1)^{1+1}\begin{vmatrix}1 & 0\\ 5 & 1\end{vmatrix} + 0\cdot(-1)^{1+2}\begin{vmatrix}0 & 0\\ 0 & 1\end{vmatrix} + 1\cdot(-1)^{1+3}\begin{vmatrix}0 & 1\\ 0 & 5\end{vmatrix}\\ &= 1.\end{aligned}$$

We call a matrix **singular** if its determinant is zero, and **nonsingular** otherwise. The matrix A in the example above is nonsingular.

We now summarize some important properties of the determinant. One caveat: these properties are not necessarily true if the entries of the matrices do not come from $\mathbb{R}$. In many cases, they *are* true when the entries come from other sets, but to go into the details requires more work than we have time for here. One particular property that we state without proof is:

Proposition 0.76. The determinant of a matrix is invariant with respect to the choice of row or column for the expansion by cofactors. That is, it doesn't matter which row or column of a matrix you choose; you always get the same answer for that matrix.

Proving Proposition 0.76 would take a lot of time, and isn't really useful for this course. Any half-decent textbook on linear algebra will have the proof, so you can look it up there, if you like.

For the remaining properties, the proof is either an exercise, or appears in an appendix to this section after the exercises.

Theorem 0.77. If B is the same as the square matrix A, except that row i has been multiplied by a scalar c, then $\det B = c \det A$.

Proof. See page 35. □

Theorem 0.78. For any square matrix A, $\det A = \det A^T$.

Proof. You do it! See Exercise 0.92. □

The next theorem requires some lesser properties, which we relegate to the status of "lemmas", typically used to help prove theorems, but considered interesting in their own right. (Much like Lemma 0.46 on page 19.) First, we state the theorem.

Theorem 0.79. If A is a square matrix and B is a matrix found by adding a multiple of one row of A to another, then $\det A = \det B$.

Now, we state and prove each of the special properties we will need.

Lemma 0.80. If B is the same as the square matrix A, except that row i has been exchanged with row j, then $\det B = -\det A$.

Proof. See page 35. □

Lemma 0.81. If the square matrix A has two identical rows, then $\det A = 0$.

Proof. See page 36. □

Notation 0.82. We write $\mathbf{a}_i$ for the ith row of matrix A.

Lemma 0.83. Let $b_1, \ldots, b_n \in \mathbb{R}$. If

$$A = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} a_{11} + b_1 & a_{12} + b_2 & \cdots & a_{1n} + b_n \\ & \mathbf{a}_2 & & \\ & \vdots & & \\ & \mathbf{a}_n & & \end{pmatrix},$$

then

$$\det B = \det A + \det \begin{pmatrix} b_1 & b_2 & \cdots & b_n \\ & \mathbf{a}_2 & & \\ & \vdots & & \\ & \mathbf{a}_n & & \end{pmatrix}.$$

Proof. See page 36. □

Theorem 0.84. A square matrix A is singular if and only if we can write its first row as a **linear combination** of the others. That is, if we write $\mathbf{a}_i$ for the ith row of A and $\dim A = n$, then we can find $c_2, \ldots, c_n \in \mathbb{R}$ such that

$$\mathbf{a}_1 = c_2 \mathbf{a}_2 + \cdots + c_n \mathbf{a}_n.$$

Proof. You do it! See Exercise 0.84. □

Theorem 0.85. For any two matrices A and B of dimension n, $\det(AB) = \det A \cdot \det B$.

Proof. See page 37. □

Theorem 0.86. An inverse of a matrix A exists if and only if $\det A \neq 0$; that is, if and only if A is nonsingular.

Proof. You do it! See Exercise 0.86. □

Exercises.

Exercise 0.87. Find two 2×2 matrices A and B such that $A \neq 0$ and $B \neq 0$, but $AB = 0$. Explain how this is related to Exercise 3.47.

Exercise 0.88. Show that matrix multiplication distributes over a sum of vectors. In other words, complete the proof of Theorem 0.71.

Exercise 0.89. Let

$$M = \begin{pmatrix} 1 & & 1 \\ & 1 & \\ & 5 & -1 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 2 & 0 & 6 \end{pmatrix}.$$

Show that $\ker M = \{0\}$, and then find $\ker N$.

Exercise 0.90. Use Theorem 0.79 to prove Theorem 0.84. That is, show that a matrix is singular if and only if we can write its first row as a linear combination of the others.

Exercise 0.91. Use Theorems 0.79 and 0.85 to prove Theorem 0.86. That is, show that a matrix has an inverse if and only if its determinant is nonzero.

Exercise 0.92. Prove Theorem 0.78. That is, show that for any matrix A, $\det A = \det A^T$.

Exercise 0.93. Show that $\det A^{-1} = (\det A)^{-1}$.
Note: In the first, we have the inverse of a matrix; in the second, we have the inverse of a number!

Exercise 0.94. Let i be the imaginary number such 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 $\mathbf{jk} = -\mathbf{kj}$.

Exercise 0.95. A matrix A is **orthogonal** if its transpose is also its inverse. Let $n \in \mathbb{N}^+$ and $O(n)$ 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) Suppose A is orthogonal. Show that $\det A = \pm 1$.

Proofs of some properties of determinants.

Notation 0.96. We write $A_{\hat{i}\hat{j}}$ for the submatrix of A formed by removing row i and column j.

Proof of Theorem 0.77. Let A and B satisfy the hypotheses. Write

$$A=\begin{pmatrix}\mathbf{a}_1\\ \vdots\\ \mathbf{a}_{i-1}\\ \mathbf{a}_i\\ \mathbf{a}_{i+1}\\ \vdots\\ \mathbf{a}_n\end{pmatrix}\quad\text{and}\quad B=\begin{pmatrix}\mathbf{a}_1\\ \vdots\\ \mathbf{a}_{i-1}\\ c\mathbf{a}_i\\ \mathbf{a}_{i+1}\\ \vdots\\ \mathbf{a}_n\end{pmatrix}.$$

Expand the determinants of both matrices along row i; then

$$\det A=\sum_{j=1}^{n}a_{ij}\,(-1)^{i+j}\det A_{\hat{i}\hat{j}}\,,$$

while

$$\det B=\sum_{j=1}^{n}\left(ca_{ij}\right)(-1)^{i+j}\det A_{\hat{i}\hat{j}}\,.$$

Apply the distributive property to factor out the common c, and we have

$$\det B=c\sum_{j=1}^{n}a_{ij}\,(-1)^{i+j}\det A_{\hat{i}\hat{j}}=c\det A\,.$$

□

Proof of Lemma 0.80. We prove the lemma for the case $i=1$ and $j=2$; the other cases are similar. We proceed by induction on the dimension n of the matrices.

For the *inductive base*, we consider $n=2$; we have

$$A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\quad\text{and}\quad B=\begin{pmatrix}c & d\\ a & b\end{pmatrix}.$$

Expansion by cofactors gives us $\det A=ad-bc$ and $\det B=bc-ad$. In other words, $\det A=-\det B$.

For the *inductive hypothesis,* we assume that for all matrices of dimension smaller than n, exchanging the first two rows negates the determinant.

For the *inductive step,* expand $\det A$ along column 1. By definition,

$$\det A=\sum_{i=1}^{n}a_{i1}\,(-1)^{i+1}\det A_{\hat{i}\hat{1}}.$$

Rewrite so that the first two elements are not part of the sum:

$$\begin{aligned}\det A &= a_{11}(-1)^{1+1}\det A_{\hat{1}\hat{1}} + a_{21}(-1)^{2+1}\det A_{\hat{2}\hat{1}} + \sum_{i=3}^{n} a_{i1}(-1)^{i+1}\det A_{\hat{i}\hat{1}} \\ &= a_{11}\det A_{\hat{1}\hat{1}} - a_{21}\det A_{\hat{2}\hat{1}} + \sum_{i=3}^{n} a_{i1}(-1)^{i+1}\det A_{\hat{i}\hat{1}}.\end{aligned}$$

In a similar way, we find that

$$\det B = b_{11}\det B_{\hat{1}\hat{1}} - b_{21}\det B_{\hat{2}\hat{1}} + \sum_{i=3}^{n} b_{i1}(-1)^{i+1}\det B_{\hat{i}\hat{1}}.$$

Recall that the difference between A and B is that we exchanged the first two rows of A to obtain B. Thus, $b_{11} = a_{21}$, $b_{21} = a_{11}$, $B_{\hat{1}\hat{1}} = A_{\hat{2}\hat{1}}$, and $B_{\hat{2}\hat{1}} = A_{\hat{1}\hat{1}}$ (it may take a moment to see why the matrices have that relationship, but it's not hard to see, in the end). For $i \geq 3$, however, $b_{i1} = a_{i1}$, while $B_{\hat{i}\hat{1}}$ is *almost* the same as $A_{\hat{i}\hat{1}}$ — the difference except that the first two rows, $\mathbf{a}_1$ and $\mathbf{a}_2$, are exchanged! The dimensions of these matrices are $n-1$, so the inductive hypothesis applies, and $\det B_{\hat{i}\hat{1}} = -\det A_{\hat{i}\hat{1}}$.Making the appropriate substitutions, we find that

$$\begin{aligned}\det B &= a_{21}\det A_{\hat{2}\hat{1}} - a_{11}\det A_{\hat{1}\hat{1}} + \sum_{i=3}^{n} a_{i1}(-1)^{i+1}\left(-\det A_{\hat{i}\hat{1}}\right) \\ &= -\left[a_{11}\det A_{\hat{1}\hat{1}} + a_{21}\det A_{\hat{2}\hat{1}} + \sum_{i=3}^{n} a_{i1}(-1)^{i+1}\left(\det A_{\hat{i}\hat{1}}\right)\right] \\ &= -\det A.\end{aligned}$$

□

Proof of Lemma 0.81. Without loss of generality, we assume that the first two rows of the square matrix A are identical; the other cases are similar. Construct a second matrix B by exchanging the first two rows of A. We can write

$$A = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_1 \\ \mathbf{a}_3 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_1 \\ \mathbf{a}_3 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}.$$

Notice that $A = B$! By substitution, $\det A = \det B$. On the other hand, Lemma 0.80 implies that $\det B = -\det A$. Thus, $\det A = -\det A$, so $2\det A = 0$, so $\det A = 0$. □

Proof of Lemma 0.83. Expand the determinant of B along its first row to see that

$$\det B = \sum_{j=1}^{n}\left(a_{1j} + b_j\right)(-1)^{1+j}\det B_{\hat{1}\hat{j}}.$$

The distributive, associative, and commutative properties allow us to rewrite this equation as

$$\det B = \sum_{j=1}^{n} a_{1j} (-1)^{1+j} \det B_{\hat{1}\hat{j}} + \sum_{j=1}^{n} b_j (-1)^{1+j} \det B_{\hat{1}\hat{j}}.$$

If you look at A and B, you will see that $A_{\hat{1}\hat{j}} = B_{\hat{1}\hat{j}}$ for every $j = 1, \ldots, n$: after all, the only difference between A and B lies in the first row, which is by definition excluded from $A_{\hat{1}\hat{j}}$ and $B_{\hat{1}\hat{j}}$. By substitution, then,

$$\det B = \det A + \det \begin{pmatrix} b_1 & b_2 & \cdots & b_n \\ & & \mathbf{a}_2 & \\ & & \vdots & \\ & & \mathbf{a}_n & \end{pmatrix},$$

as claimed. □

Proof of Theorem 0.79. Without loss of generality, we may assume that we constructed B from A by adding a multiple of the second row to the first. That is,

$$A = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \mathbf{a}_1 + c\mathbf{a}_2 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}.$$

By Lemma 0.83,

$$\det B = \det \begin{pmatrix} \mathbf{a}_1 + c\mathbf{a}_2 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \det \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + \det \begin{pmatrix} c\mathbf{a}_2 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}.$$

Now apply Theorem 0.77 and Lemma 0.81 to see that

$$\det B = \det A + c \det \begin{pmatrix} \mathbf{a}_2 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \det A + c \cdot 0 = \det A.$$

□

Proof of Theorem 0.85. If $\det A = 0$, then Theorem 0.84 tells us that we can find real numbers c_2, $\ldots$, c_n such that $\mathbf{a}_1 = \sum_{k=2}^{n} c_k \mathbf{a}_k$. By properties of matrix multiplication,

$$\mathbf{a}_1 B = \left(\sum_{k=2}^{n} c_k \mathbf{a}_k \right) B = \sum_{k=2}^{n} c_k (\mathbf{a}_k B).$$

Notice that $\mathbf{a}_i B$ is the ith row of AB, so this new equation shows that the first row of AB is a linear combination of the other rows. Theorem 0.84 again implies that $\det(AB) = 0$.

Now suppose $\det A \neq 0$. A fact of linear algebra that we do not repeat here is that we can write

$$A = E_1 E_2 \cdots E_m,$$

where we construct each E_i by applying one of the operations of Theorem 0.77, Lemma 0.80, or Lemma 0.83 to I_n. Thus,

$$\det(AB) = \det(E_1 \cdots E_m B).$$

Let $C = E_2 \cdots E_m B$; we have $\det(AB) = \det(E_1 C)$. We now consider three possible values of E_1.

Case 1: If E_1 is the result of swapping two rows of I_n, then $\det E_1 = -1$. On the other hand, $E_1 C$ is the same as C, except that two rows of C are swapped — the same two rows as in E_1, in fact. So $\det(E_1 C) = -\det C = \det E_1 \cdot \det C$.

Case 2: If E_1 is the result of multiplying a row of I_n by a constant $c \in \mathbb{R}$, then $\det E_1 = c$. On the other hand, $E_1 C$ is the same as C, except that a row of C has been multiplied by a constant $c \in \mathbb{R}$ — the same row as in E_1, in fact. So $\det(E_1 C) = c \det C = \det E_1 \cdot \det C$.

Case 3: If E_1 is the result of adding a multiple of a row of I_n to another row, then $\det E_1 = \det I_n = 1$. On the other hand, $E_1 C$ is the same as C, except that a multiple of a row of C has been added to another row of C — the same two rows as E_1, in fact, and the same multiple. So $\det(E_1 C) = \det C = \det E_1 \det C$.

In each case, we found that $\det(E_1 C) = \det E_1 \det C$. Thus, $\det(AB) = \det E_1 \cdot \det(E_2 \cdots E_m B)$. We now repeat this process for each of the E_i, obtaining

$$\det(AB) = \det E_1 \cdots \det E_m \det B = \det A \det B.$$

□

0.4: "Real", "imaginary", and complex numbers

A significant motivation for the development of algebra was to study roots of polynomials. A polynomial, of course, has the form

$$ax + b, \quad ax^2 + bx + c, \quad ax^3 + bx^2 + cx + d, \quad \ldots$$

and polynomials model many real-world phenomena. A **root** of a polynomial $f(x)$ is any a such that $f(a) = 0$. For example, if $f(x) = x^4 - 1$, then 1 and -1 are both roots of f. However, they are not the *only* roots of f! For the full explanation, you'll need to read about polynomial rings and ideals in Chapters 7 and 8, but we can take some first steps in that direction already.

"Real" numbers

A root of $x^2 - 2$ would be any number α such that $\alpha^2 = 2$. We can refer to α as a square root of 2, or $\sqrt{2}$.

Is $\sqrt{2}$ a "real" number? Is a number worth studying if you need it merely to solve some abstract polynomial? In this case, yes: take a right triangle with two legs, each of length 1; how long is the hypotenuse? By the Pythagorean Theorem,

$$a^2 + b^2 = c^2 \quad \Longrightarrow \quad 1^2 + 1^2 = x^2 \quad \Longrightarrow \quad 2 = x^2 \quad \Longrightarrow \quad x^2 - 2 = 0\,.$$

So the length of the hypotenuse is the root of x^2-2, which we call $\sqrt{2}$.

At this point we encounter a fact that disturbed its discoverers:[8] *This number is not rational!*

Theorem 0.97. $\sqrt{2}\notin\mathbb{Q}$.

Proof. We proceed by contradiction. If $\sqrt{2}$ were rational, then we could write $\sqrt{2}=a/b$.

Without loss of generality, $b>0$. (After all, if $b<0$, then $a/b=-a/-b$, and then Juliet's rose by another name smells just as sweet.)

Also without loss of generality, one of a or b is not divisible by 2. (After all, if both are divisible by 2, then we can find $m,n\in\mathbb{Z}$ with $n>0$ and

$$\frac{a}{b}=\frac{2m}{2n}=\frac{m}{n},$$

and we have $0<n<b$; if both m and n are still even, we can repeat; and by Fact on page 13 we cannot repeat indefinitely.)

Having cleared up some details on a and b, consider the fact that

$$\sqrt{2}=\frac{a}{b} \quad\Longrightarrow\quad 2=\frac{a^2}{b^2} \quad\Longrightarrow\quad 2b^2=a^2 .$$

So a^2 is even. But if a^2 is even, a itself must be even. (After all, if a is not even, then by the Division Theorem $a=2q+1$ for some $q\in\mathbb{Z}$, and then $a^2=2\left(2q^2+2q\right)+1$, so a^2 is not even, a contradiction!) So we can find $m\in\mathbb{Z}$ such that $a=2m$. That means

$$2b^2=a^2=(2m)^2=4m^2 \quad\Longrightarrow\quad b^2=2m^2.$$

So b^2 is even. But if b^2 is even, b itself must be even. (After all, we can replace a in the previous "After all" with b.)

We have found that a and b are both even — but, *wait a minute…* — we observed earlier that we could assume they weren't. That's a contradiction! Something we assumed isn't correct, but:

- it is definitely correct to assume $b>0$, as explained; and
- it is likewise correct to assume a and b cannot both be divisible by 2.

The only assumption that can be incorrect is the one that $\sqrt{2}$ is rational. It must be irrational! □

That means we cannot measure all line segments using rational numbers alone.

[8] The story is that a follower of the Pythagorean cult at Samos discovered this fact. The Pythagoreans had previously believed all numbers are rational numbers. When one of their own showed that the world of mathematics is, in fact, irrational, they rewarded him with a one-way cruise to the bottom of the Mediterranean Sea.
Now, *this is the way I heard the story,* and I've seen it repeated in some other places. (You know, the Internet, and all.) So I can't have invented it whole hog. But some other places (you know, the Internet, and all) give a variant of the story where the Pythagoreans drowned not the discoverer of the secret, but a member who revealed the secret, and in some corners it seems that even the monstrosity of drowning him is completely doubted. So perhaps they were not so irrational as to believe in a rational world.

To remedy the situation, we define $\mathbb{R}$, the **real numbers** as the set that contains a value for every possible measurement of distance along a line with respect to some fixed unit.[9] How does $\mathbb{R}$ relate to the sets of numbers we already know?

- Certainly we have measurements of length 0 units, 1 unit, 2 units, ..., so $\mathbb{N} \subseteq \mathbb{R}$.
- If we think in terms of "direction," we can say that positive lengths are "in front of" a given position, while negative directions are "behind" it, so it does not seem unreasonable to accept that lengths of -1 unit, -2 units, ... are also sensible, so $\mathbb{Z} \subseteq \mathbb{R}$.
- Likewise we can divide a measurement into pieces as easily as a pizza, so we can have lengths of $1/2$ unit, $-5/7$ unit, and so forth. In other words, $\mathbb{Q} \subseteq \mathbb{R}$.
- On the other hand, we know that $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$, so $\mathbb{R} \neq \mathbb{Q}$.

In short,

$$\mathbb{N}^+ \subsetneq \mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}\,.$$

How do we order the real numbers? We leave the details to a class in analysis, but you can treat it as you have in the past.

Do we need anything else? Indeed, we do.

The complex numbers

We mentioned $f(x) = x^4 - 1$ earlier. This factors as $f(x) = (x-1)(x+1)\left(x^2+1\right)$. The roots 1 and -1 show up in the linear factors, and they're the only possible roots of those factors. So, if f has other roots, we would expect them to be roots of x^2+1. However, x^2+1 has no "real" roots, because the square of a real number is nonnegative; adding 1 forces it to be positive.

Let's imagine x^2+1 has a root anyway. If it doesn't make sense, we should find out soon enough and abandon this thought experiment. Let $g(x) = x^2+1$, and say that g has a root, which we'll call i, for "imaginary". Since i is a root of g, we have the equation

$$0 = g(i) = i^2 + 1,$$

or $i^2 = -1$.

We'll create a new set of numbers, $\mathbb{C}$, by adding i to the set $\mathbb{R}$. We'd still like to perform arithmetic on this set, so we have to extend our definitions of addition and multiplication, and maybe add more objects.

Start by assigning $\mathbb{C} = \mathbb{R} \cup \{i\} \cup \ldots$. Is multiplication closed? On the one hand, $2 \in \mathbb{C}$ and $i \in \mathbb{C}$. Since $2i$ is in neither $\mathbb{R}$ nor $\{i\}$, we'll have to add products like it if we want to keep multiplication closed, so we need to add the set $\{bi : b \in \mathbb{R}\}$. Notice that this includes the $i = 1 \cdot i$. On the other hand, $i^2 = -1$, and $-1 \in \mathbb{R}$ already, so certainly $-1 \in \mathbb{C}$, so we're covered there.

Our set has expanded to $\mathbb{C} = \mathbb{R} \cup \{bi : b \in \mathbb{R}\} \cup \ldots$. Can we add in this set? Well, $1 \in \mathbb{C}$ and $2i \in \mathbb{C}$ Since $1+2i$ and is in neither $\mathbb{R}$ nor $\{bi : b \in \mathbb{R}\}$, we'll have to sums like it if we want to keep addition closed, so we need to add the set $\{a+bi : a, b \in \mathbb{R}\}$. Notice that this includes $\mathbb{R}$ since $a = a + 0i$, and it also includes $\{bi : b \in \mathbb{R}\}$ since $bi = 0 + bi$.

That gives us

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\} \cup \ldots.$$

[9]Speaking "precisely", $\mathbb{R}$ is the set of limits of "nice sequences" of rational numbers. By "nice", we mean that the elements of the sequence eventually grow closer together than any rational number. The technical term for this is a **Cauchy sequence**. For more on this, see any textbook on real analysis.

Do we need anything else?

Notice that

$$(a+bi)+(c+di)=(a+c)+(b+d)i ,$$

so addition is now closed. Similarly,

$$(a+bi)(c+di)=ac+adi+bci+bdi^2 ,$$

and $i^2=-1$, so the expression above simplifies to

$$(a+bi)(c+di)=(ac-bd)+(ad+bc)i ,$$

so multiplicaiton is now closed.

What about division? It turns out that we can multiply by the conjugate to obtain a quotient:

$$\frac{1+i}{2+3i} = \frac{1+i}{2+3i}\cdot\frac{2-3i}{2-3i} = \frac{5-i}{13} = \frac{5}{13}-i\cdot\frac{1}{13} .$$

or, in general,

$$\frac{a+bi}{c+di} = \frac{a+bi}{c+di}\cdot\frac{c-di}{c-di} = \frac{(ac+bd)+i(bc-ad)}{c^2+d^2} = \frac{ac+bd}{c^2+d^2}+i\cdot\frac{bc-ad}{c^2+d^2} .$$

In other words, division of nonzero complex numbers is also an operation on nonzero complex numbers.

One question remains: what about roots of complex numbers? This one's a little tougher, so we won't discuss its full generality, but we can establish that $\sqrt[n]{i}\in\mathbb{C}$ in the following way. Recall that $\sqrt[n]{i}$ would be the root of $x^n-i=0$. Let $\alpha=\cos(\pi/2n)+i\sin(\pi/2n)$. Then

$$\alpha^2=\left(\cos\frac{\pi}{2n}+i\sin\frac{\pi}{2n}\right)^2=\left(\cos^2\frac{\pi}{2n}-\sin^2\frac{\pi}{2n}\right)+i\left(2\cos\frac{\pi}{2n}\sin\frac{\pi}{2n}\right) .$$

Recall from trigonometry that $\cos^2\beta-\sin^2\beta=\cos 2\beta$, and that $2\sin\beta\cos\beta=\sin 2\beta$. If we apply this with $\beta=\pi/2n$, we see that the equation above simplifies to

$$\alpha^2=\cos\frac{2\pi}{2n}+i\sin\frac{2\pi}{2n}=\cos\left(2\times\frac{\pi}{2n}\right)+i\sin\left(2\times\frac{\pi}{2n}\right) .$$

(No, we do *not* want to simplify it. You'll see why in a moment.)

Let's raise the roof:

$$\begin{aligned}\alpha^3=\alpha^2\cdot\alpha&=\left(\cos\frac{2\pi}{2n}+i\sin\frac{2\pi}{2n}\right)\left(\cos\frac{\pi}{2n}+i\sin\frac{\pi}{2n}\right)\\&=\left(\cos\frac{2\pi}{2n}\cos\frac{\pi}{2n}-\sin\frac{2\pi}{2n}\sin\frac{\pi}{2n}\right)+i\left(\cos\frac{2\pi}{2n}\sin\frac{\pi}{2n}+\sin\frac{2\pi}{2n}\cos\frac{\pi}{2n}\right) .\end{aligned}$$

Recall from trigonometry that $\cos\beta\cos\gamma-\sin\beta\sin\gamma=\cos(\beta+\gamma)$, and also that $\cos\beta\sin\gamma+\sin\beta\cos\gamma=\sin(\beta+\gamma)$. If we apply this with $\beta=2\pi/2n$ and $\gamma=\pi/2n$, we see that the equation above simplifies to

$$\alpha^3 = \cos\frac{3\pi}{2n} + i\sin\frac{3\pi}{2n} = \cos\left(3\times\frac{\pi}{2n}\right) + i\sin\left(3\times\frac{\pi}{2n}\right).$$

Notice a pattern? This continues indefinitely, so that

$$\alpha^n \;=\; \cos\left(n\times\frac{\pi}{2n}\right) + i\sin\left(n\times\frac{\pi}{2n}\right) \;=\; \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} \;=\; 0 + i\cdot 1 \;=\; i \quad ,$$

or in short,

$$\alpha^n = i\ .$$

So the nth root of i is the complex number $\cos\frac{\pi}{2n} + i\sin\frac{\pi}{2n}$. For $n = 2,3,4,\ldots$ that gives us

$$\sqrt{i} \;=\; \cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\,, \quad \sqrt[3]{i} \;=\; \cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\,, \quad \sqrt[4]{i} \;=\; \cos\frac{\pi}{8} + i\sin\frac{\pi}{8}\,, \quad \ldots$$

In any case, α is both a root of $x^n - i$ and a complex number, so the statement "$\sqrt[n]{i} \in \mathbb{C}$" seems reasonable. The story is of course more sordid than this, and it would take a branch or two of complex analysis to sort out, but the main point is that we can say that $\mathbb{C}$ "crowns" arithmetic and basic algebra.

Definition 0.98. The **complex numbers** are the set

$$\mathbb{C} = \left\{a + bi : a, b \in \mathbb{R},\ i^2 = -1\right\}.$$

The **real part** of $a + bi$ is a, and the **imaginary part** is b.

Remark 0.99. You may wonder whether it is legitimate to just conjure up some imagined number i, and build a new set by adjoining it to $\mathbb{R}$. Isn't that just a little, oh, *imaginary?* No, actually; we can provide two very sound justifications.

First, mathematicians typically model the oscillation of a pendulum by a differential equation of the form $y'' + ay = 0$. As any book in the subject explains, we have good reason to solve this *differential* equations by solving a *polynomial* equation of the form $r^2 + a = 0$. The solutions to this equation are $r = \pm i\sqrt{a}$, so unless the oscillation of a pendulum is "imaginary", i is quite "real".

Second, in Section 8.3 we construct from the real numbers a set that looks an awful lot like these purported complex numbers, using a very sensible approach. We can even show that this set is "identical" to the complex numbers in all the ways that matter.[10] This strongly suggests that the imaginary numbers are quite "real".

The complex plane

We can diagram the real numbers along a line. In fact, that's how we define them intuitively: we said that a real number is any length along a real line. What about the complex numbers: can we diagram them?

[10]It's "identical," but only in quotes, sort of like an algebraic doppelgänger. The correct term is "isomorphic," and by the time we're done explaining what that is, you may well feel that doppelgänger is the right term.

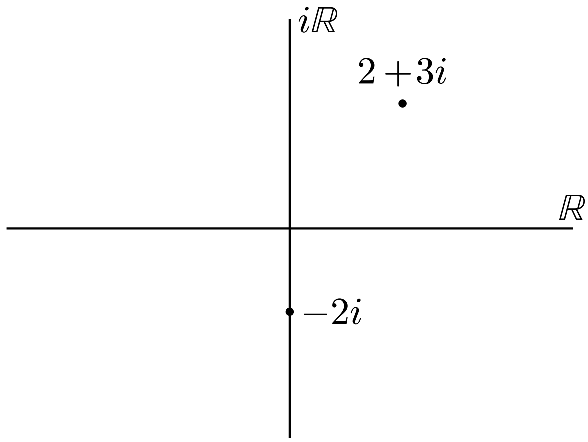


Figure 0.4. Two elements of $\mathbb{C}$, visualized as points on the complex plane

By definition, any complex number is the sum of its real and imaginary parts. We cannot simplify $a+bi$ any further using this representation, so we cannot mix and mingle a and b without changing the complex number. In the same way, when we are given a point (a,b), we cannot mix and mingle a and b without changing that point. The correspondence goes further.

Theorem 0.100. There is a one-to-one, onto function from $\mathbb{C}$ to $\mathbb{R}^2$ that maps 1 to $(1,0)$ and i to $(0,1)$.

Proof. Let $\varphi : \mathbb{C} \to \mathbb{R}^2$ by $\varphi(a+bi) = (a,b)$. That is, we map a complex number to $\mathbb{R}^2$ by sending the real part to the first entry (the x-ordinate) and the imaginary part to the second entry (the y-ordinate). As desired, $\varphi(1) = (1,0)$ and $\varphi(i) = (0,1)$.

Is φ one-to-one and onto? We see that φ is one-to-one by the fact that if $\varphi(a+bi) = \varphi(c+di)$, then $(a,b) = (c,d)$; equality of points in $\mathbb{R}^2$ implies that $a=c$ and $b=d$; equality of complex numbers implies that $a+bi = c+di$. We see that φ is onto by the fact that for any $(a,b) \in \mathbb{R}^2$, $\varphi(a+bi) = (a,b)$. □

Since $\mathbb{R}^2$ has a nice, geometric representation as the x-y plane, we can represent complex numbers in the same way. That motivates our definition of the **complex plane**, which is nothing more than a visualization of $\mathbb{C}$ in $\mathbb{R}^2$.

Take a look at Figure 19. We have labeled the x-axis as $\mathbb{R}$ and the y-axis as $i\mathbb{R}$. We call the former the **real axis** and the latter the **imaginary axis** of the complex plane. This agrees with our mapping above, which sent the real part of a complex number to the x-ordinate, and the imaginary part to the y-ordinate. Thus, the complex number $2+3i$ corresponds to the point $(2,3)$, while the complex number $-2i$ corresponds to the point $(0,-2)$.

Exercises.

Exercise 0.101.

(a) Adapt our proof that $\sqrt{2}$ is irrational to show that $\sqrt{3}$ is also irrational.

(b) Why can we not adapt our proof that $\sqrt{2}$ is irrational to show that $\sqrt{4}$ is irrational? (Aside from the obvious reason that $\sqrt{4} = 2$, I mean. Rather, I want you to identify where the proof fails to generalize in this case.)

Exercise 0.102. Graph the nth roots of i on the complex plane for $i = 2, 3, 4$. Actually, no, let's do this better:

(a) Graph the square root of i on one complex plane, the cube root of i on another complex plane, and the fourth root of i on a third complex plane.
(b) Compute $\left(\sqrt[3]{i}\right)^2$, and graph it on the second complex plane.
(c) Compute $\left(\sqrt[4]{i}\right)^2$ and $\left(\sqrt[4]{i}\right)^3$, and graph both on the third complex plane.
(d) What pattern do you notice in the results?

Exercise 0.103. It turns out that $\mathbb{C}$ is a vector space over $\mathbb{R}$, and the function φ of Theorem 0.100 is a linear transformation. We don't want to get into the details of vector spaces here, so let's accept that $\mathbb{C}$ is indeed a vector space over $\mathbb{R}$, but why don't you verify for yourself that φ is a linear transformation; that is, for any $a, b, c, d \in \mathbb{R}$ we have

$$\varphi((a+bi)+(c+di)) = \varphi(a+bi)+\varphi(c+di)$$

and

$$\varphi(r(a+bi)) = r\varphi(a+bi) \ .$$

Part I

Monoids and groups

Chapter 1: Monoids

Algebra was created to solve problems. Like other branches of mathematics, it started off solving very applied problems of a certain type: problems that "simplify" to polynomial equations. Before now, you focused on techniques that solve the simplest of these: factoring, isolating a variable, and taking roots.

These techniques emerged centuries ago. They work well for all linear and quadratic equations, and it's possible to apply them to any polynomial equation of degree three or four, though it requires some massaging. However, some polynomials of degree five or higher resisted solution by these methods. This prompted algebra to take a radically different turn in the 18th century, leading to the discovery of hidden structures that can be used to analyze a vast array of problems which look superficially different, but are surprisingly similar.

This chapter introduces some important algebraic ideas. We will try to be intuitive, but don't confuse "intuitive" with "vague"; we maintain precision while using concrete examples. One goal is to get you to use these examples when thinking about the more general ideas later on. It will be important for you to *explore* and *play* with the ideas and examples, specializing or generalizing them as needed to attack new problems.

Success in this course will require you to develop both inductive and deductive approaches.

1.1: From integers and monomials to monoids

We move our focus from the "arithmetical" integers to the "algebraic" monomials. In doing so, we notice a similarity in the mathematical structure. That similarity will motivate our first steps into modern algebra, with monoids.

Monomials

Let x represent an unknown quantity. The set of "**univariate monomials** in x" is

$$\mathbb{M} = \{x^a : a \in \mathbb{N}\} \ , \tag{5}$$

where x^a, a "monomial", represents precisely what you'd think: the product of a copies of x. In other words,

$$x^a \quad = \quad \prod_{i=1}^{a} x \quad = \quad \underbrace{x \cdot x \cdot \cdots \cdot x}_{n \text{ times}} .$$

We can extend this notion. Let $x_1, x_2, \ldots, x_n$ represent unknown quantities. The set of "**multivariate monomials** in $x_1, x_2, \ldots, x_n$" is

$$\mathbb{M}_n = \left\{ \prod_{i=1}^{m} \left(x_1^{a_{i1}} x_2^{a_{i2}} \cdots x_n^{a_{in}} \right) : m, a_{ij} \in \mathbb{N} \right\} \ . \tag{6}$$

("Univariate" means "one variable"; "multivariate" means "many variables".) For monomials, we allow neither coefficients nor negative exponents. The definition of $\mathbb{M}_n$ indicates that any of its elements is a "product of products".

Example 1.1. The following are monomials:

$$x^2, \quad 1 = x^0 = x_1^0 x_2^0 \cdots x_n^0, \quad x^2 y^3 x y^4.$$

In the case of $x^2 y^3 x y^4$, we might be able to simplify it to $x^3 y^7$, or we might not — it depends on what they represent. If they represent integers, then yes, we can simplify the product, but if they represent matrices... well, those don't always commute, do they? This motivated the definition of an element of $\mathbb{M}_n$ as a product of products. We could write $x^2 y^3 x y^4$ in those terms as

$$\left(x^2 y^3\right)\left(x y^4\right) = \prod_{i=1}^{m} \left(x_1^{a_{i1}} x_2^{a_{i2}}\right)$$

with $m = 2$, $a_{11} = 2$, $a_{12} = 3$, $a_{21} = 1$, and $a_{22} = 4$.

The following are not monomials:

$$x^{-1} = \frac{1}{x}, \quad \sqrt{x} = x^{\frac{1}{2}}, \quad \sqrt[3]{x^2} = x^{\frac{2}{3}}.$$

Their exponents are not nonnegative numbers: $-1, \frac{1}{2}, \frac{2}{3} \notin \mathbb{Z}$.

Similarities between $\mathbb{M}$ *and* $\mathbb{N}$

We are interested in similarities between $\mathbb{N}$ and $\mathbb{M}$. Why? Suppose that we can identify a structure common to the two sets. If we make the obvious properties of this structure precise, we can determine non-obvious properties that must be true about $\mathbb{N}$, $\mathbb{M}$, and any other set that adheres to the structure. These are the beginnings of *abstract* algebra. People often resist abstraction, but

if we can prove a fact about an abstract structure,
then we don't have to re-prove that fact for any set with that structure!
This ***saves*** *time and* ***increases*** *understanding.*

It is harder at first to think about abstract structures than concrete objects, but time, effort, and determination bring agility.

To begin with, what operation(s) should we normally associate with $\mathbb{M}$? We normally associate addition and multiplication with the natural numbers, but adding monomials does *not* typically give us another monomial: $x^2 + x^4$ is a *polynomial*, not a monomial! If you look back at Definition 0.9, you'll see that this means the monomials *are not* closed under addition.

On the other hand, $x^2 \cdot x^4$ remains a monomial, and in fact $x^a x^b \in \mathbb{M}$ for any choice of $a, b \in \mathbb{N}$. This is true about monomials in any number of variables; that is, the monomials *are* closed under multiplication.

Lemma 1.2. Let $n \in \mathbb{N}^+$. Both $\mathbb{M}$ and $\mathbb{M}_n$ are closed under multiplication.

Proof for $\mathbb{M}$. Let $t, u \in \mathbb{M}$. By definition, there exist $a, b \in \mathbb{N}$ such that $t = x^a$ and $u = x^b$. By definition of monomial multiplication, we see that

$$tu = x^{a+b}.$$

Since addition is closed in $\mathbb{N}$, the expression $a+b$ simplifies to a natural number. Call this number c. By substitution, $tu = x^c$. This has the form of a univariate monomial; compare it with the description of a monomial in equation (5). So, $tu \in \mathbb{M}$. Since we chose t and u to be arbitrary elements of $\mathbb{M}$, and found their product to be an element of $\mathbb{M}$, we conclude that $\mathbb{M}$ is closed under multiplication. □

Easy, right? We won't usually state all those steps explicitly, but we want to do so at least once.

What about $\mathbb{M}_n$? The lemma claims that multiplication is closed there, too, but we haven't proved that yet. I wanted to separate the two, to show how operations you take for granted in the univariate case don't work so well in the multivariate case. The problem here is that the variables might not commute under multiplication. If we knew that they did, we could write something like,

$$tu = x_1^{a_1+b_1} \cdots x_n^{a_n+b_n},$$

so long as the a's and the b's were defined correctly. Unfortunately, if we assume that the variables *are* commutative, then we don't prove the statement for everything that we would like. This requires a little more care in developing the argument. Sometimes, it's just a game of notation, as it will be here.

Proof for $\mathbb{M}_n$. Let $t, u \in \mathbb{M}_n$. By definition, we can write

$$t = \prod_{i=1}^{m_t} \left(x_1^{a_{i1}} \cdots x_n^{a_{in}}\right) \quad \text{and} \quad u = \prod_{i=1}^{m_u} \left(x_1^{b_{i1}} \cdots x_n^{b_{in}}\right).$$

(We give subscripts to m_t and m_u because t and u might have a different number of factors. Using different symbols for m_t and m_u allows for this possibility without requiring it.) By substitution,

$$tu = \left(\prod_{i=1}^{m_t} \left(x_1^{a_{i1}} \cdots x_n^{a_{in}}\right)\right)\left(\prod_{i=1}^{m_u} \left(x_1^{b_{i1}} \cdots x_n^{b_{in}}\right)\right).$$

Intuitively, you want to declare victory; we've written tu as a product of variables, right? All we see are variables, organized into two products.

Unfortunately, we're not quite there yet. To show that $tu \in \mathbb{M}_n$, we must show that we can write it as *one* product of a list of products, rather than two. This turns out to be as easy as making the symbols do what your head is telling you: two lists of products of variables, placed side by side, make one list of products of variables. To show that it's one list, we must identify explicitly how many "small products" are in the "big product". There are m_t in the first, and m_u in the second, which makes $m_t + m_u$ in all. So we know that we should be able to write

$$tu = \prod_{i=1}^{m_t+m_u} \left(x_1^{c_{i1}} \cdots x_n^{c_{in}}\right) \tag{7}$$

for appropriate choices of c_{ij}. The hard part now is identifying the correct values of c_{ij}.

In the list of products, the first few products come from t. How many? There are m_t from t.

The rest are from u. We can specify this precisely using a piecewise function:

$$c_{ij} = \begin{cases} a_{ij}, & 1 \le i \le m_t \\ b_{ij}, & m_t < i. \end{cases}$$

Specifying c_{ij} this way justifies our claim that tu has the form shown in equation (7). That satisfies the requirements of $\mathbb{M}_n$, so we can say that $tu \in \mathbb{M}_n$. Since t and u were chosen arbitrarily from $\mathbb{M}_n$, it is closed under multiplication. □

Life is a little harder when we can't make all the assumptions we would like!

As with the proof for $\mathbb{M}$, we were somewhat pedantic here; don't expect this level of detail all the time. Pedantry has the benefit that you don't have to read between the lines. That means you don't have to think much, only recall previous facts and apply very basic logic. However, pedantry also makes proofs long and boring. While you could shut down much of your brain while reading a pedantic proof, that would be counterproductive. Ideally, you want a reader to *think* while reading a proof, so shutting down the brain is bad. Thus, a good proof does not recount every basic definition or result for the reader, but requires her to make basic recollections and inferences.

Let's look at two more properties.

Lemma 1.3. Let $n \in \mathbb{N}^+$. Multiplication in $\mathbb{M}$ satisfies the commutative property. Multiplication in both $\mathbb{M}$ and $\mathbb{M}_n$ satisfies the associative property.

Proof. We show this to be true for $\mathbb{M}$; the proof for $\mathbb{M}_n$ we will omit (but it can be done as it was above). Let $t, u, v \in \mathbb{M}$. By definition, there exist $a, b, c \in \mathbb{N}$ such that $t = x^a$, $u = x^b$, and $v = x^c$. By definition of monomial multiplication and by the commutative property of addition in $\mathbb{M}$, we see that

$$tu = x^{a+b} = x^{b+a} = ut.$$

As t and u were arbitrary, multiplication of univariate monomials is commutative.

By definition of monomial multiplication and by the associative property of addition in $\mathbb{N}$, we see that

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

□

You might ask yourself, *Do I have to show* every *step?* That depends on what the reader needs to understand the proof. In the equations above, it *is* essential to show that the commutative and associative properties of multiplication in $\mathbb{M}$ depend strictly on the commutative and associative properties of addition in $\mathbb{N}$. Thus, the steps

$$x^{a+b} = x^{b+a} \quad \text{and} \quad x^{a+(b+c)} = x^{(a+b)+c},$$

with the parentheses as indicated, are absolutely crucial, and cannot be omitted from a good proof.[11]

Another property the natural numbers have is that of an identity: both additive and multiplicative. Since we associate only multiplication with the monomials, we should check whether they have a multiplicative identity. I hope this one doesn't surprise you!

Lemma 1.4. Both $\mathbb{M}$ and $\mathbb{M}_n$ have $1 = x^0 = x_1^0 x_2^0 \cdots x_n^0$ as a multiplicative identity.

We won't bother proving this one, but leave it to the exercises.

Monoids

There are quite a few other properties that the integers and the monomials share, but the three properties we have mentioned here are already interesting enough to motivate the following definition.

Definition 1.5. Let M be a set, and $\circ$ an operation on M. We say that the pair $(M, \circ)$ is a **monoid** if it satisfies the following properties:

(closed)	for any $x, y \in M$, we have $x \circ y \in M$;
(associative)	for any $x, y, z \in M$, we have $(x \circ y) \circ z = x \circ (y \circ z)$; and
(identity)	there exists an **identity element** $e \in M$ such that for any $x \in M$, we have $e \circ x = x \circ e = x$.

We may also say that M is a **monoid under** $\circ$.

So far, then, we know the following:

Theorem 1.6. $\mathbb{N}$ is a monoid under both addition and multiplication, while $\mathbb{M}$ and $\mathbb{M}_n$ are monoids under multiplication.

Proof. For $\mathbb{N}$, this is part of its definition. For $\mathbb{M}$ and $\mathbb{M}_n$, see Lemmas 1.2, 1.3, and 1.4. □

We don't generally write the operation in conjunction with the set; we write the set alone, leaving it to the reader to infer the operation. So, for instance, we generally say that "$\mathbb{M}$ as a monoid" and don't mention multiplication.

In some cases, this might lead to ambiguity; after all, both $(\mathbb{N}, +)$ and $(\mathbb{N}, \times)$ are monoids, so which should we prefer? In general, if we refer to "the monoid $\mathbb{N}$," we mean $(\mathbb{N}, +)$. Thus, we can write that $\mathbb{N}$, $\mathbb{M}$, and $\mathbb{M}_n$ are examples of monoids: the first under addition, the others under multiplication.

What other mathematical objects are examples of monoids?

Example 1.7. Let $m, n \in \mathbb{N}^+$. The set of $m \times n$ matrices with integer entries, written $\mathbb{Z}^{m \times n}$, satisfies properties that make it a monoid under addition:

- closure is guaranteed by the definition;
- the associative property is guaranteed by the associative property of its elements; and
- the additive identity is 0, the zero matrix, by Theorem 0.64.

[11]Of course, a professional mathematician would not even prove these things in a scholarly text or article, because they are well-known and easy. On the other hand, a good professional mathematician *would* feel compelled to include in a proof steps that include novel and/or difficult information.

Example 1.8. The set of square matrices with integer entries $\mathbb{Z}^{m\times m}$ satisfies properties that make it a monoid under multiplication:
- closure is guaranteed by the definition;
- the associative property is guaranteed by Theorem 0.67; and
- the multiplicative identity is I_n, by Theorem 0.64.

Here's an example you probably haven't seen before.

Example 1.9. Consider the set $B = \{F, T\}$ with the operation $\wedge$ ("boolean and") where

$$T \wedge T = T \quad \text{and}$$
$$F \wedge x = F \text{ for any } x \in B.$$

We claim that B is a monoid under this operation. To see why, consider that:
- By inspection, the operation is closed: every pair of inputs gives either T or F, both of which are in B.
- By inspection, T acts as an identity:

$$T \wedge F = F \wedge T = F \quad \text{and} \quad T \wedge T = F\ .$$

- By inspection, the operation is associative: for any $x, y \in B$,

$$\begin{aligned}(T \wedge T) \wedge T &= T = T \wedge (T \wedge T)\\ (x \wedge y) \wedge F &= F\ = x \wedge F = x \wedge (y \wedge F)\\ (x \wedge F) \wedge y = F \wedge y &= F\ = x \wedge F = x \wedge (F \wedge y)\\ (F \wedge x) \wedge y = F \wedge y &= F\ = F \wedge z = F \wedge (x \wedge y) \quad (\text{where } z = x \wedge y)\ .\end{aligned}$$

(The phrase "by inspection" means that we are performing an exhaustive check of every possibility. In a set with only two elements, this is easy. As you will discover, it doesn't take long before this becomes tedious: already we have to check $8 = 2^3$ cases for B, and to avoid checking all 8 we tried to be clever by taking arbitrary $x, y \in B$ to reduce the work to 4 cases.)

Example 1.10. Let S be a set, and let F_S be the set of all functions mapping S to itself, with the proviso that for any $f \in F_S$, $f(s)$ is defined for every $s \in S$. For instance,
- $2x + 1 \in F_{\mathbb{Z}}$ because $2z + 1 \in \mathbb{Z}$ for every $z \in \mathbb{Z}$, but
- $1/x \notin F_{\mathbb{Z}}$ because for "most" values of z, $1/z \notin \mathbb{Z}$, and in the case $z = 0$ it isn't even defined!

We can show that F_S is a monoid under composition of functions, since
- for any $f, g \in F_S$, we also have $f \circ g \in F_S$, where $f \circ g$ is the function h such that for any $s \in S$,

$$h(s) = (f \circ g)(s) = f(g(s))$$

  (notice how important it was that $g(s)$ have a defined value regardless of the value of s);
- for any $f, g, h \in F_S$, we have $(f \circ g) \circ h = f \circ (g \circ h)$, since for any $s \in S$,

$$((f \circ g) \circ h)(s) = (f \circ g)(h(s)) = f(g(h(s)))$$

  and

$$(f \circ (g \circ h))(s) = f((g \circ h)(s)) = f(g(h(s)));$$

- if we consider the function $\iota \in F_S$ where $\iota(s) = s$ for all $s \in S$, then for any $f \in F_S$, we have $\iota \circ f = f \circ \iota = f$, since for any $s \in S$,

$$(\iota \circ f)(s) = \iota(f(s)) = f(s)$$

and

$$(f \circ \iota)(s) = f(\iota(s)) = f(s)$$

(we can say that $\iota(f(s)) = f(s)$ because $f(s) \in S$).

Although monoids are useful, they don't capture all the properties that interest us. For example, the Division Theorem requires *two* operations: multiplication (by the quotient) and addition (of the remainder). But monoids think about only one operation! So, there is no "Division Theorem for Monoids"; it simply doesn't make sense. If we want to generalize the Division Theorem to other sets, we will need a more specialized structure. We will actually meet one later! (Section 7.4.)

That doesn't mean we should abandon monoids so quickly, though; other gems may hide in these bushes. For instance, we have talked a lot about identities. A natural question to ask is whether an identity is unique. We actually did ask this about the matrices, back in Section 0.3, but we put the question off at the time. Why? because it's difficult to do with matrices, and elegant to do with monoids.

"Unique" in mathematics means *exactly one*. To prove uniqueness of an object x, you consider a generic object y that shares all the properties of x, then reason to show that $x = y$. This is not a contradiction, because we didn't assume that $x \neq y$ in the first place; we simply wondered about a generic y. We did the same thing in the proof of the Division Theorem (Theorem 0.41 on page 17).

To prove that the identity matrix is unique, you could try multiplying two *arbitrary* matrices, and analyzing the result to show that one of them *has* to be I_n. You would have to work with a matrix of *arbitrary* size, and the proof would start off something like this:

> Let A and B be $n \times n$ matrices, and suppose $AB = A$. By definition of matrix multiplication,
>
> $$AB = \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + \cdots + a_{1n}b_{n2} & \cdots & a_{11}b_{1n} + \cdots + a_{1n}b_{nn} \\ a_{21}b_{11} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + \cdots + a_{2n}b_{n2} & \cdots & a_{21}b_{1n} + \cdots + a_{2n}b_{nn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}b_{11} + \cdots + a_{nn}b_{n1} & a_{n1}b_{12} + \cdots + a_{nn}b_{n2} & \cdots & a_{n1}b_{1n} + \cdots + a_{nn}b_{nn} \end{pmatrix}.$$
>
> Since $AB = A$, we get the system of equations
>
> $$a_{11}b_{11} + \cdots + a_{1n}b_{n1} = a_{11},$$

… and so forth. Ugh.

The abstract structure of a monoid makes this much simpler!

Theorem 1.11. Suppose that M is a monoid, and there exist $e, i \in M$ such that both e and i satisfy the identity property. Then $e = i$, so that the identity of a monoid is unique.

Proof. Suppose that e and i are both identities. Since i is an identity, we know that

$$e = ei.$$

Since e is an identity, we know that

$$ei = i.$$

By substitution,

$$e = i.$$

We chose arbitrary identities of M and showed that they were in fact the same element. This implies there is only one identity. □

This proof is short, sweet, and to the point. It is simpler and more elegant in every way than any investigation we might make of matrices, *and* it works for *any* monoid. In this case, abstraction — working with *generic* elements of a *generic* monoid — proves its worth. The proof is shorter, easier to follow, and easier to replicate.

Exercises.

Exercise 1.12. It turns out that $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are all monoids under both addition and multiplication. This is rather tedious to work out in its entirety, so you can simply assum them after this question. But, *for this question,*

(a) Assuming that $\mathbb{Z}$ is a monoid under multiplication, show that $\mathbb{Q}$ is a monoid under multiplication.
(b) Assuming that $\mathbb{Z}$ is a monoid under addition and multiplication, show that $\mathbb{Q}$ is a monoid under addition.
(c) Assuming that $\mathbb{R}$ is a monoid under addition, show that $\mathbb{C}$ is a monoid under addition.
(d) Assuming that $\mathbb{R}$ is a monoid under multiplication, show that $\mathbb{C}$ is a monoid under multiplication.

Exercise 1.13. Let $t, u, v \in \mathbb{M}_n$.

(a) Show that if $t \mid u$ and $u \mid v$, then $t \mid v$.
(b) Show that if t and u are two monomials such that $t \mid u$ and $u \mid t$, then $t = u$.

Exercise 1.14. Is $\mathbb{N}$ a monoid under:

(a) subtraction?
(b) division?

Be sure to explain your answer.

Exercise 1.15. Is $\mathbb{Z}$ a monoid under:

(a) addition?
(b) subtraction?
(c) multiplication?
(d) division?

Be sure to explain your answer.

Exercise 1.16. 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, explain how it justifies each property.

Exercise 1.17. 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, explain how it justifies each property.

Exercise 1.18. Suppose multiplication of x and y commutes. Show that multiplication in $\mathbb{M}_n$ is both closed and associative.

Exercise 1.19.
(a) Show that $\mathbb{N}[x]$, the ring of polynomials in one variable with integer coefficients, is a monoid under addition.
(b) Show that $\mathbb{N}[x]$ is also a monoid if the operation is multiplication.
(c) Explain why we can replace $\mathbb{N}$ by $\mathbb{Z}$ and the argument would remain valid. (*Hint:* think about the *structure* of these sets.)

Exercise 1.20. Let A be a set of symbols, and L the set of all finite sequences that can be constructed using elements of A. Let $\circ$ represent *concatenation of lists*. For example, $(a, b) \circ (c, d, e, f) = (a, b, c, d, e, f)$. Show that $(L, \circ)$ is a monoid.

Definition 1.21. For any set S, let $P(S)$ denote the set of all subsets of S. We call this the **power set** of S.

Exercise 1.22.
(a) Suppose $S = \{a, b\}$. Compute $P(S)$, and show that it is a monoid under $\cup$ (union).
(b) Let S be *any* set. Show that $P(S)$ is a monoid under $\cup$ (union).

Exercise 1.23.
(a) Suppose $S = \{a, b\}$. Compute $P(S)$, and show that it is a monoid under $\cap$ (intersection).
(b) Let S be *any* set. Show that $P(S)$ is a monoid under $\cap$ (intersection).

Exercise 1.24. Let X be a set that is closed under addition, and under multiplication by real numbers. Let $A \subseteq X$. If $tx + (1-t)y \in A$ for any $t \in [0,1]$ and for any $x, y \in A$, we call A **convex**. Let $\mathcal{C}(X)$ be the set of all convex subsets of X. Show that $\mathcal{C}(X)$ is a monoid under the operation $\cap$.

Definition 1.25. Let X be a set, and $\prec$ a linear ordering of the elements of X. Let $S \subseteq X$. We say that S is **convex with respect to** $\prec$ if, for any $x, y \in S$ and for any $z \in X$, if $x \prec z \prec y$, then $z \in S$ also.

Exercise 1.26. Let $\mathcal{C}_{\prec}(X)$ be the set of all subsets of X that are convex with respect to $\prec$. Show that $\mathcal{C}_{\prec}(X)$ is a monoid under $\cap$.

Exercise 1.27.
(a) Fill in each blank of Figure 1.27 with the justification.
(b) Is $(\mathbb{N}, \text{lcm})$ also a monoid? If so, do we have to change anything about the proof? If not, which property fails?

Exercise 1.28. Recall the usual ordering $<$ on $\mathbb{M}$: $x^a < x^b$ if $a < b$. Show that this is both a linear ordering and a well-ordering.

Remark 1.29. While we can define a well-ordering on $\mathbb{M}_n$, it is a much more complicated proposition, which we take up in Section 11.2.

Exercise 1.30. In Exercise 0.50, you showed that divisibility is transitive in the integers.
(a) Show that divisibility is transitive in *any* monoid; that is, if M is a monoid, $a, b, c \in M$, $a \mid b$, and $b \mid c$, then $a \mid c$. (Of course $a \mid b$ means here that we can find $q \in M$ such that $aq = b$.)
(b) In fact, you don't need all the properties of a monoid for divisibility to be transitive! Which properties *do* you need?

1.2: Isomorphism

We've seen that several important sets share the monoid structure. In particular, $(\mathbb{N}, +)$ and $(\mathbb{M}, \times)$ are very similar. Are they in fact identical *as monoids?* The technical word for this is *isomorphism*; how can we determine whether two monoids are isomorphic? We will look for a way to determine whether their operations behave the same way.

Suppose you need to choose between two offices: how would you decide if they were equally suitable? First, you would need to know what tasks have to be completed, and what materials you need for those tasks. For example, both probably have a desk and a phone, but if your job required a significant amount of data transfer with clients, you would need a high-speed internet port. If both offices had that feature, you'd conclude they were equally suitable, but if one office had only a phone jack, it wouldn't matter how pretty it looked: you'd have to choose the other.

Deciding whether two sets are isomorphic is similar. First we decide what algebraic structure we need. In our case, we've studied only monoids, so for now we care only whether the sets have the same monoid structure. Next, we compare how the sets satisfy those structural properties. If you're looking at finite monoids, an exhaustive comparison of every possible product might

Claim: $\left(\mathbb{N}^+, \text{lcm}\right)$ is a monoid. Note that the operation here looks unusual: instead of something like $x \circ y$, you're looking at $\text{lcm}(x,y)$.

Proof:

1. First we show closure.
   (a) Let $a, b \in$_____, and let $c = \text{lcm}(a,b)$.
   (b) By definition of _____, $c \in \mathbb{N}^+$.
   (c) By definition of _____, $\mathbb{N}^+$ is closed under lcm.
2. Next, we show the associative property. This is one is a bit tedious...
   (a) Let $a, b, c \in$_____.
   (b) Let $m = \text{lcm}(a, \text{lcm}(b,c))$, $n = \text{lcm}(\text{lcm}(a,b), c)$, and $\ell = \text{lcm}(b,c)$. By _____, we know that $\ell, m, n \in \mathbb{N}$.
   (c) We claim that $\text{lcm}(a,b)$ divides m.
      i. By definition of _____, both a and $\text{lcm}(b,c)$ divide m.
      ii. By definition of _____, we can find x such that $m = ax$.
      iii. By definition of _____, both b and c divide m.
      iv. By definition of _____, we can find y such that $m = by$
      v. By definition of _____, both a and b divide m.
      vi. By Exercise _____, $\text{lcm}(a,b)$ divides m.
   (d) Recall that _____ divides m. Both $\text{lcm}(a,b)$ and _____ divide m. (Both blanks expect the same answer.)
   (e) By definition of _____, $n \leq m$.
   (f) A similar argument shows that $m \leq n$; by Exercise _____, $m = n$.
   (g) By _____, $\text{lcm}(a, \text{lcm}(b,c)) = \text{lcm}(\text{lcm}(a,b), c)$.
   (h) Since $a, b, c \in \mathbb{N}$ were arbitrary, we have shown that lcm is associative.
3. Now, we show the identity property.
   (a) Let $a \in$_____.
   (b) Let $\iota =$_____.
   (c) By arithmetic, $\text{lcm}(a, \iota) = a$.
   (d) By definition of _____, ι is the identity of $\mathbb{N}^+$ under lcm.
4. We have shown that $\left(\mathbb{N}^+, \text{lcm}\right)$ satisfies the properties of a monoid.

Figure 1.1. Material for Exercise 1.27

work, but there's a reason we call that approach "exhaustive." Besides, mathematics deals with infinite sets like $\mathbb{N}$ and $\mathbb{M}$ 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 $x \in S$ to precisely one value in T, the output $f(x)$. You have probably heard the geometric interpretation of this: f passes the "vertical line test." Isomorphism *specializes* the notion of a function in a way that tells us important information about a set's structure.

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. (Desk to desk, phone to phone.) What does that mean? Let $x, y, z \in M$ and $a, b, c, \in N$. If M and N have the same structure as monoids, with x corresponding to a, y corresponding to b, and z corresponding to c, then we expect that

$$xy = z \implies ab = c\ .$$

In mathematics, the function notation $f(x) = a$ fills is for "x corresponding to a," so we can rewrite the second equation as

$$f(x)f(y) = f(z)\ .$$

By substitution,

$$f(x)f(y) = f(xy)\ .$$

That's not quite enough. The identity property is pretty important, so the idenfity of M should correspond to the identity of N. So a monoid isomorphism needs to verify that

$$f(e_M) = e_N.$$

In addition, 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, or

f should be onto.

(If you do not remember the definitions of one-to-one and onto, see Definition 0.38 on page 16.)

The following definition summarizes the discussion.[12]

Definition 1.31. Let $(M, \times)$ and $(N, +)$ be monoids. If there exists a function $f : M \longrightarrow N$ such that

- $f(e_M) = e_N$ *(f preserves the identity)*

and

- for all $x, y \in M$,
  $f(xy) = f(x) + f(y)$, *(f preserves the operation)*

then we call f a **homomorphism**. If f is also a bijection, then we say that 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 $(M, \times)$ and $(N, +)$ in the definition partly to suggest our goal of showing that $\mathbb{M}$ and $\mathbb{N}$ are isomorphic, but also because they could stand for *any* monoids. You will see in due course that not all monoids are isomorphic, but first let's see about $\mathbb{M}$ and $\mathbb{N}$.

Example 1.32. Monomials under multiplication $(\mathbb{M}, \times)$ are isomorphic to natural numbers under addition $(\mathbb{N}, +)$.

Before we show the precise details, consider this an an intuitive level. How do we simplify monomial products? by adding exponents. What are the monomial exponents? natural numbers.

[12]The word *homomorphism* comes from the Greek words for *same* and *shape;* the word *isomorphism* comes from the Greek words for *identical* and *shape*. You can consider the *"shapes"* we are comparing to be the monoids' "multiplication table". Isomorphism shows that the "multiplication tables" have the same structure.

So the idea that these monomials are isomorphic to natural numbers should have a certain appeal.

We turn to the details. We have to find a function $f : \mathbb{M} \to \mathbb{N}$ that preserves the identity, preserves the operation, and is a bijection. Let's start with a function that relates monomials and natural numbers… but how?

Our intuitive consideration should help here: monomials have exponents, which are natural numbers. It can't hurt to try mapping each monomial to its exponent:[13]

$$f(x^a) = a.$$

First we show that f is a bijection.

To see that it is one-to-one, let $t, u \in \mathbb{M}$, and assume $f(t) = f(u)$. By definition of $\mathbb{M}$, we can find $a, b \in \mathbb{N}$ such that $t = x^a$ and $u = x^b$. Substituting into $f(t) = f(u)$, we find that $f(x^a) = f\left(x^b\right)$. 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 $t, u \in \mathbb{M}$, 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 $t \in \mathbb{M}$ such that $f(t) = a$. Which t should we choose? We want $f\left(x^?\right) = a$. By definition, $f\left(x^?\right) = ?$. The "natural" choice seems to be $? = a$, or $t = x^a$. That would certainly guarantee $f(t) = a$, but does such an object t exist in $\mathbb{M}$? Any natural exponent is a monomial, so yes, $x^a \in \mathbb{M}$. We showed that for arbitrary $a \in \mathbb{N}$, f maps some element of $\mathbb{M}$ 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 $t, u \in \mathbb{M}$.[14] By definition of $\mathbb{M}$, we can find $a, b \in \mathbb{N}$ such that $t = x^a$ and $u = x^b$. We now manipulate $f(tu)$ using definitions and substitutions to show the operation is preserved:

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

By the transitive property of equality, $f(tu) = f(t)f(u)$, so f preserves the operation.

Does f also preserve the identity? We usually write the identity of $M = \mathbb{M}$ as the symbol 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:

$$f(e_M) = f(1) = f\left(x^0\right) = 0 = e_N.$$

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

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

[13] Okay, maybe it could hurt a little, especially when you're just starting, and your first intuition doesn't pan out, and you have to deal with frustration. After a while you grow inured to the pain and your intutitive muscles grow into lean, mean, proof-building machines. (Not that I know anything about that.)

[14] The definition uses the variables x and y, but those are just letters that stand for arbitrary elements of M. Here $M = \mathbb{M}$ 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 $\mathbb{M}$, because there *is* an element of $\mathbb{M}$ called x. So we choose t and u instead.

Example 1.33. 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

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

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 $(\mathbb{N},+) \not\cong (\mathbb{N},\times)$.

Here's one last example that shows how you can use inspection in small monoids.

Example 1.34. The monoids "Boolean or" and "Boolean and" from Exercise 1.16 and Example 1.9 are isomorphic.

First we have to find our function $f : (B,\vee) \to (B,\wedge)$:

- Since f has to map identity to identity, we have no choice but to send the F, the identity of "or", to T, the identity of "and." So $f(F) = T$.
- What should $f(T)$ be? We have two choices, F and T, but f has to be one-to-one, and we've already assigned $f(F) = T$. So there is no choice but to assign $f(T) = F$.

We have now defined $f : (B,\vee) \to (B,\wedge)$ such that

$$f(x) = \begin{cases} F, & x = T \\ T, & x = F \end{cases} .$$

We build f to preserve the identity and to be one-to-one, so we're in the clear there. Inspection also shows that f is onto: every element of $(B,\wedge)$ has a preimage in $(B,\vee)$.

Does f preserve the operation? Here, unfortunately, we have to perform an exhaustive check, but with $4 = 2^2$ cases it isn't too tedious:

$$f(F \vee F) = f(F) = T = T \wedge T = f(F) \wedge f(F)$$
$$f(F \vee T) = f(T) = F = T \wedge F = f(F) \wedge f(T)$$
$$f(T \vee F) = f(T) = F = F \wedge T = f(T) \wedge f(F)$$
$$f(T \vee T) = f(T) = F = F \wedge F = f(T) \wedge f(T) \quad .$$

Exercises.

Exercise 1.35.

(a) Show that the monoids "Boolean and" and "Boolean or" are isomorphic.

(b) This is *not* the same as Example 1.34 in the sense that you are now going in the opposite direction, so you need a function $f : (B,\wedge) \to (B,\vee)$, and at least one of the properties will be different. But after you do Exercise 1.39 below, that won't be necessary. Which property of an equivalence relation renders this unnecessary?

Exercise 1.36. Show that the monoids "Boolean or" and "Boolean xor" from Exercises 1.16 and 1.17 are *not* isomorphic.

Exercise 1.37. Let $M = \{\{\}, \{a\}\}$.
(a) Show that M is a monoid under the operation $\cup$.
(b) Show that M is isomorphic to the monoid "Boolean or".
(c) Can M be isomorphic to the monoid "Boolean xor"?

Exercise 1.38. 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"?

Exercise 1.39. Let $(M, \times)$, $(N, +)$, and $(P, \sqcap)$ be monoids.
(a) Show that the identity function $\iota(x) = x$ is an isomorphism on M.
(b) Suppose we know $(M, \times) \cong (N, +)$. That means there is an isomorphism $f : M \to N$. One requirement 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.
(c) Suppose we know $(M, \times) \cong (N, +)$ and $(N, +) \cong (P, \sqcap)$. That means 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.

1.3: Direct products

We showed in the last section that the monoid of *univariate* monomials is isomorphic to the monoid of natural numbers; that is, $(\mathbb{M}, \times) \cong (\mathbb{N}, +)$.

What about a *multivariate* monomial? A bivariate monomial corresponds to a pair of natural numbers:
$$x^6 y^3 \text{ looks an awful lot like } (6, 3).$$
The product of two commutative, multivariate monomials corresponds to adding two pair of natural numbers component-wise:
$$(x^6 y^3)(x^2 y) = x^8 y^4 \quad \text{and} \quad (6, 3) + (2, 1) = (8, 4).$$
It's starting to look as if bivariate monomials are isomorphic to the set of pairs of natural numbers — well, that might be true if that were a monoid. Is it?

And what about monomials in more variables? Might they also be isomorphic to sequences of natural numbers? Again, first we have to determine if the set of such sequences are monoids.

So sequences of elements are important. In fact, they're important enough that they already have a name.

Definition 1.40. Let $r>1$ and $S_1, S_2, \ldots, S_r$ be sets. The **Cartesian product** of $S_1, \ldots, S_r$ is the set of all lists of r elements where the ith entry is an element of S_i; that is,

$$S_1 \times \cdots \times S_r = \{(s_1, s_2, \ldots, s_n) : s_i \in S_i\}.$$

If $S_1 = \cdots = S_r$ we may write S^r instead of $S_1 \times \cdots \times S_r$.

We call an element of a Cartesian product a **sequence** or a **tuple**.

Example 1.41. We have already mentioned $\mathbb{N} \times \mathbb{N} = \{(a,b) : a, b \in \mathbb{N}\}$. We can write $\mathbb{N}^2$ instead of $\mathbb{N} \times \mathbb{N}$. Some tuples of $\mathbb{N} \times \mathbb{N}$ are $(0,0)$ and $(5,2)$.

Example 1.42. Another Cartesian product would be $\mathbb{N} \times \mathbb{M}$; tuples of $\mathbb{N} \times \mathbb{M}$ include $(2, x^3)$ and $(0, x^5)$. In general, $\mathbb{N} \times \mathbb{M}$ is the set of all ordered pairs where the first entry is a natural number, and the second is a monomial.

It turns out that if the sets that form a Cartesian product share a common structure, then we can preserve that structure using a *direct product*.

Definition 1.43. Let $r \in \mathbb{N}^+$ and $M_1, M_2, \ldots, M_r$ be monoids. The **direct product** of $M_1, \ldots, M_r$ is the pair

$$(M_1 \times \cdots \times M_r, \otimes)$$

where $M_1 \times \cdots \times M_r$ is the usual Cartesian product, and $\otimes$ is the "natural" operation on $M_1 \times \cdots \times M_r$.

What do we mean by the **natural operation** on $M_1 \times \cdots \times M_r$? Let $x, y \in M_1 \times \cdots \times M_r$; by definition, we can write

$$x = (x_1, \ldots, x_r) \quad \text{and} \quad y = (y_1, \ldots, y_r)$$

where each x_i and each y_i is an element of M_i. Then

$$x \otimes y = (x_1 y_1, x_2 y_2, \ldots, x_r y_r)$$

where each product $x_i y_i$ is performed according to the operation that makes the corresponding M_i a monoid.

Example 1.44. Recall that $\mathbb{N}^2$ is a Cartesian product; when viewed as a direct product, it turns out to be a monoid, just as $\mathbb{N}$ is a monoid. *In this case* we write $\oplus$ for the natural operation, because $\otimes$ looks a little odd when you're adding.[15]

(closure) Let $\mathbf{u}, \mathbf{v} \in \mathbb{N}^2$. By definition of a Cartesian product, $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$. By definition of the natural operation,

$$\mathbf{u} \oplus \mathbf{v} = (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2) \ .$$

[15] We can actually use any symbol we like, and later we will simply write $+$, but for now it's best to highlight that it isn't ordinary addition.

Since $\mathbb{N}$ is a monoid under addition, it is closed under adddition, so $u_1 + v_1, u_2 + v_2 \in \mathbb{N}$ and $(u_1 + v_1, u_2 + v_2) \in \mathbb{N}^2$. By substitution, $\mathbf{u} \oplus \mathbf{v} \in \mathbb{N}^2$. We took arbitrary elements of $\mathbb{N}^2$ and showed their sum is in $\mathbb{N}^2$, so $\mathbb{N}^2$ is closed under addition.

(associative) Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{N}^2$. By definition of a Cartesian product, $\mathbf{u} = (u_1, u_2)$, $\mathbf{v} = (v_1, v_2)$, and $\mathbf{w} = (w_1, w_2)$. By definition of the natural operation,

$$\mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w}) = (u_1, u_2) + (v_1 + w_1, v_2 + w_2) = (u_1 + (v_1 + w_1), u_2 + (v_2 + w_2))$$

and

$$(\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w} = (u_1 + v_1, u_2 + v_2) + (w_1, w_2) = ((u_1 + v_1) + w_1, (u_2 + v_2) + w_2) \ .$$

Since $\mathbb{N}$ is a monoid under addition, the addition is associative, so $u_1 + (v_1 + w_1) = (u_1 + v_1) + w_1$ and $u_2 + (v_2 + w_2) = (u_2 + v_2) + w_2$. By substitution, $\mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w}) = (\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w}$. We took arbitrary elements of $\mathbb{N}^2$ and showed their sum is in $\mathbb{N}^2$, so $\mathbb{N}^2$ is associative under addition.

(An aside. The proof that addition is associative in $\mathbb{N}^2$ depends on the fact that addition is associative in $\mathbb{N}$, so *we had to make that explicit.* This is why we work out the sum two different ways, then explain explicitly that the two sums are the same. We did the same thing when proving the properties of a monoid. *You may not* simply write this:

$$(u_1, u_2) + (v_1 + w_1, v_2 + w_2) = (u_1 + v_1 + w_1, u_2 + v_2 + w_2)$$

because that skips over the fact that v_1 and w_1 are added before u_1. Watch carefully for these habits to which you are accustomed!)

(identity) We claim that $\mathtt{0} = (0,0)$ is the additive identity of $\mathbb{N}^2$. First note that $0 \in \mathbb{N}$ so indeed $\mathtt{0} \in \mathbb{N}^2$. Now let $\mathbf{u} \in \mathbb{N}^2$. By definition of a Cartesian product, $\mathbf{u} = (u_1, u_2)$. By definition of the natural operation,

$$\mathbf{u} + \mathtt{0} = (u_1, u_2) + (0,0) = (u_1 + 0, u_2 + 0) \ .$$

Since 0 is the additive identity of $\mathbb{N}$, this simplifies to $(u_1, u_2) = \mathbf{u}$. By substitution, $\mathbf{u} + \mathtt{0} = \mathbf{u}$. A similar proof shows that $\mathtt{0} + \mathbf{u} = \mathbf{u}$. We took an arbitrary element of $\mathbb{N}^2$ and showed that its sum with $\mathtt{0}$ is itself, so $\mathtt{0}$ is indeed the additive identity of $\mathbb{N}^2$.

(Another aside. Above I wrote, "A similar proof shows that $\mathtt{0} + \mathbf{u} = \mathbf{u}$." You, the student, may be wondering, "Can I do that?" You certainly can — as long as it's true. But how will you know when it's true? In this particular case, we can give two different explanations: (a) the operation is commutative, so it doesn't matter what side $\mathtt{0}$ is on, and (b) it should be clear that, if we reverse the positions of $\mathbf{u}$ and $\mathtt{0}$, nothing really changes in our proof. This *might not* be true if we were dealing with matrices!)

Example 1.45. Recall that $\mathbb{N} \times \mathbb{M}$ is a Cartesian product; we can show that the direct product of $(\mathbb{N}, \times)$ and $(\mathbb{M}, \times)$ is a monoid. (Notice that here we are using *multiplication* in $\mathbb{N}$.)

(closure) Let $t, u \in \mathbb{N} \times \mathbb{M}$. By definition, we can write $t = (a, x^\alpha)$ and $u = \left(b, x^\beta\right)$ for appropriate $a, \alpha, b, \beta \in \mathbb{N}$. Then

$$\begin{aligned} tu &= (a, x^\alpha) \otimes \left(b, x^\beta\right) && \text{(subst.)} \\ &= \left(ab, x^\alpha x^\beta\right) && \text{(def. of } \otimes\text{)} \\ &= \left(ab, x^{\alpha+\beta}\right) \in \mathbb{N} \times \mathbb{M}. && \text{(monomial mult.)} \end{aligned}$$

We took two arbitrary elements of $\mathbb{N} \times \mathbb{M}$, multiplied them according to the new operation, and obtained another element of $\mathbb{N} \times \mathbb{M}$; the operation is therefore closed.

(associative) Let $t, u, v \in \mathbb{N} \times \mathbb{M}$. By definition, we can write $t = (a, x^\alpha)$, $u = \left(b, x^\beta\right)$, and $v = (c, x^\gamma)$ for appropriate $a, \alpha, b, \beta, c, \gamma \in \mathbb{N}$. Then

$$\begin{aligned} t(uv) &= (a, x^\alpha) \otimes \left[\left(b, x^\beta\right) \otimes (c, x^\gamma)\right] \\ &= (a, x^\alpha) \otimes \left(bc, x^\beta x^\gamma\right) \\ &= \left(a(bc), x^\alpha \left(x^\beta x^\gamma\right)\right). \end{aligned}$$

To show that this equals $(tu)v$, we rely on the associative properties of $\mathbb{N}$ and $\mathbb{M}$:

$$\begin{aligned} t(uv) &= \left((ab)c, \left(x^\alpha x^\beta\right) x^\gamma\right) \\ &= \left(ab, x^\alpha x^\beta\right) \otimes (c, x^\gamma) \\ &= \left[(a, x^\alpha) \otimes \left(b, x^\beta\right)\right] \otimes (c, x^\gamma) \\ &= (tu)v. \end{aligned}$$

We took three elements of $\mathbb{N} \times \mathbb{M}$, and showed that the operation was associative for them. Since the elements were arbitrary, the operation is associative on all of $\mathbb{N} \times \mathbb{M}$.

(identity) We claim that the identity of $\mathbb{N} \times \mathbb{M}$ is $(1,1) = \left(1, x^0\right)$. To see why, let $t \in \mathbb{N} \times \mathbb{M}$. By definition, we can write $t = (a, x^\alpha)$ for appropriate $a, \alpha \in \mathbb{N}$. Then

$$\begin{aligned} (1,1) \otimes t &= (1,1) \otimes (a, x^\alpha) \\ &= (1 \times a, 1 \times x^\alpha) \\ &= (a, x^\alpha) = t \end{aligned}$$

and similarly $t \otimes (1,1) = t$. We took an arbitrary element of $\mathbb{N} \times \mathbb{M}$, and showed that $(1,1)$ acted as an identity under the operation $\otimes$ with that element. Since the element was arbitrary, $(1,1)$ must be *the* identity for $\mathbb{N} \times \mathbb{M}$.

Interestingly, if we had used $(\mathbb{N}, +)$ *instead* of $(\mathbb{N}, \times)$ in the previous example, we *still* would have obtained a direct product! Indeed, the direct product of monoids is *always* a monoid!

Theorem 1.46. The direct product of monoids $M_1, \ldots, M_r$ is itself a monoid. Its identity element is $(e_1, e_2, \ldots, e_r)$, where e_i denotes the identity of the monoid M_i.

Proof. You do it! See Exercise 1.51. □

We conclude with this question: are $\mathbb{M}_n$ and $\mathbb{M}^n$ isomorphic?

The two are admittedly not identical: $\mathbb{M}_n$ is the set of *products* of powers of n *distinct* variables, whereas $\mathbb{M}^n$ is a set of *lists* of powers of *one* variable. That is, elements of $\mathbb{M}_n$ have the form $x_1^2 x_2^5 x_4$, while elements of $\mathbb{M}^n$ have the form $\left(x^2, x^5, x\right)$ or perhaps $\left(x^2, x^5, 1, x\right)$. In addition, if the variables of $\mathbb{M}_n$ are *not* commutative, then $\mathbb{M}_n$ and $\mathbb{M}^n$ are not at all similar. Think about $(xy)^4 = xyxyxyxy$; if the variables are commutative, we can combine them into x^4y^4, which looks likes $(4,4)$. If the variables are not commutative, however, it is not at *all* clear how we could get $(xy)^4$ to correspond to an element of $\mathbb{N} \times \mathbb{N}$.

In fact, they cannot: in a non-commutative system we can write

$$(xy)^4 = (xy)(xy)(xy)(xy)$$

…and that's it. But in $\mathbb{M}^2$ we can rewrite

$$(x,x) \otimes (x,x) \otimes (x,x) \otimes (x,x) = \left(x^4, x^4\right).$$

So we need $(xy)^4$ to map to $\left(x^4, x^4\right)$. But we should probably map x^4y^4 to $\left(x^4, x^4\right)$, so the operation is not preserved.

Now, this is only one map, but it's the most sensible map possible. The following proof shows that this is generally true.

Theorem 1.47. $\mathbb{M}_n \cong \mathbb{M}^n$ if and only if the variables of $\mathbb{M}_n$ are commutative.

Proof. Assume the variables of $\mathbb{M}_n$ are commutative. Define $f : \mathbb{M}_n \longrightarrow \mathbb{M}^n$ by

$$f\left(x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}\right) = \left(x^{a_1}, x^{a_2}, \ldots, x^{a_n}\right).$$

The fact that we cannot combine a_i and a_j if $i \neq j$ shows that f is one-to-one, and any element $\left(x^{b_1}, \ldots, x^{b_n}\right)$ of $\mathbb{M}^n$ has a preimage $x_1^{b_1} \cdots x_n^{b_n}$ in $\mathbb{M}_n$; thus, f is a bijection.

Is f also an isomorphism? To see that it is, let $t, u \in \mathbb{M}_n$. By definition, we can write $t = x_1^{a_1} \cdots x_n^{a_n}$ and $u = x_1^{b_1} \cdots x_n^{b_n}$ for appropriate $a_1, b_1 \ldots, a_n, b_n \in \mathbb{N}$. Then

$$\begin{aligned}
f(tu) &= f\left(\left(x_1^{a_1} \cdots x_n^{a_n}\right)\left(x_1^{b_1} \cdots x_n^{b_n}\right)\right) && \text{(substitution)} \\
&= f\left(x_1^{a_1+b_1} \cdots x_n^{a_n+b_n}\right) && \text{(commutative)} \\
&= \left(x^{a_1+b_1}, \ldots, x^{a_n+b_n}\right) && \text{(definition of } f\text{)} \\
&= \left(x^{a_1}, \ldots, x^{a_n}\right) \otimes \left(x^{b_1}, \ldots, x^{b_n}\right) && \text{(def. of } \otimes\text{)} \\
&= f(t) \otimes f(u). && \text{(definition of } f\text{)}
\end{aligned}$$

Hence f is an isomorphism, and we conclude that $\mathbb{M}_n \cong \mathbb{M}^n$.

Conversely, suppose $\mathbb{M}_n \cong \mathbb{M}^n$. By Exercise 1.39, $\mathbb{M}^n \cong \mathbb{M}_n$. By definition, there exists a bijection $f : \mathbb{M}^n \longrightarrow \mathbb{M}_n$ satisfying Definition 1.31. Let $t, u \in \mathbb{M}_n$; by definition, we can find

$a_i, b_j \in \mathbb{N}$ such that $t = x_1^{a_1} \cdots x_n^{a_n}$ and $u = x_1^{b_1} \cdots x_n^{b_n}$. Since f is onto, we can find $v, w \in \mathbb{M}^n$ such that $f(v) = t$ and $f(w) = u$. As a homomorphism, f preserves the operation of $\mathbb{M}^n$ in $\mathbb{M}_n$, so

$$f(v \otimes w) = f(v) f(w), \quad \text{or,} \quad f(v \otimes w) = tu.$$

Exercise 1.52 (with the $S_i = \mathbb{M}$) tells us that v and w, as elements of $\mathbb{M}^n$, commute under its operation, $\otimes$, so $v \otimes w = w \otimes v$. Returning again to the homomorphism property of f, we see that

$$f(w \otimes v) = f(w) \otimes f(v) = ut.$$

Chaining these equalities together gives us

$$tu = ut.$$

We have shown that, if $\mathbb{M}_n \cong \mathbb{M}^n$, then $\mathbb{M}_n$ is commutative. □

Notation 1.48. Although we used $\otimes$ in this section to denote the operation in a direct product, this is not standard; I was trying to emphasize that the product is different for the direct product than for the monoids that created it. In general, the product $x \otimes y$ is written simply as xy. Thus, the last line of the proof above would have $f(t) f(u)$ instead of $f(t) \otimes f(u)$.

Exercises.

Exercise 1.49. Imitate the proof of Example 1.44 to show that $\mathbb{Q}^2$ is a monoid under the natural operation. *Do not use* Theorem 1.46!

Exercise 1.50. Imitate the proof of Example 1.44 to show that $(B, \wedge) \times (B, \vee)$ is a monoid under the natural operation. *Do not use* Theorem 1.46!

Exercise 1.51. Prove Theorem 1.46. Use Example 1.44 as a guide.

Exercise 1.52. Suppose M_1, M_2, ..., and M_n are *commutative* monoids. Show that the direct product $M_1 \times M_2 \times \cdots \times M_n$ is also a *commutative* monoid.

Exercise 1.53. Show that if the variables of $\mathbb{M}_n$ are commutative, then it is isomorphic to $\mathbb{N}^n$.

Exercise 1.54. Show that if the variables of $\mathbb{M}_m$ and $\mathbb{M}_n$ are commutative, then $\mathbb{M}_m \times \mathbb{M}_n$ is isomorphic to $\mathbb{M}_{m+n}$.

Exercise 1.55. Let $\mathbb{T}_S^n$ denote the set of terms in n commutative variables whose coefficients are elements of the set S. For example, $2xy \in \mathbb{T}_{\mathbb{Z}}^2$ and $\pi x^3 \in \mathbb{T}_{\mathbb{R}}^1$.
(a) Show that if S is a monoid, then so is $\mathbb{T}_S^n$.
(b) Show that if S is a monoid, then $\mathbb{T}_S^n \cong S \times \mathbb{M}_n$.

1.4: Absorption and the Ascending Chain Condition

We conclude our study of monoids by introducing a new object, and a fundamental notion.

Absorption

Definition 1.56. Let M be a monoid, and $A \subseteq M$. If $ma \in A$ for every $m \in M$ and $a \in A$, then A **absorbs from** M. We also say that A is an **absorbing subset**, or that A satisfies the **absorption property**.

Notice that if A absorbs from M, then A is closed under multiplication: if $x, y \in A$, then $A \subseteq M$ implies that $x \in M$, so by absorption, $xy \in A$, as well. Unfortunately, that doesn't make A a monoid, as e_M might not be in A.

Example 1.57. Consider $\mathbb{Z}$ as a monoid under multiplication. Write $2\mathbb{Z}$ for the set of even integers; that is,

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

By definition, $2\mathbb{Z}$ is not a monoid, as it lacks the multiplicative identity: $1 \notin 2\mathbb{Z}$. On the other hand,$2\mathbb{Z} \subsetneq \mathbb{Z}$, and if we take any $m \in \mathbb{Z}$ and any $a \in 2\mathbb{Z}$, we can write $a = 2z$ for some $z \in \mathbb{Z}$ and see that

$$ma = m(2z) = 2(mz) \in 2\mathbb{Z}.$$

Since a and m were arbitrary, $2\mathbb{Z}$ absorbs from $\mathbb{Z}$.

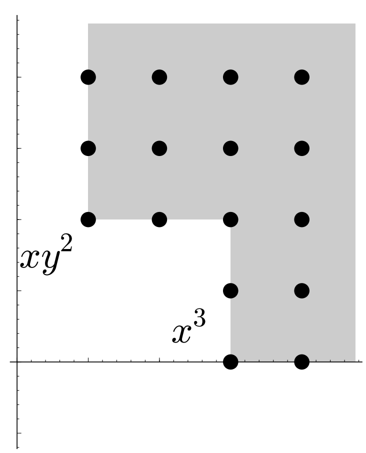
As we refer often to the set of integer multiples of an integer, we have a special notation for it.

Notation 1.58. We write $d\mathbb{Z}$ for the set of integer multiples of d.

That makes $2\mathbb{Z}$ the set of integer multiples of 2, $5\mathbb{Z}$ the set of integer multiples of 5, and so forth. You will show in Exercise 1.68 that $d\mathbb{Z}$ absorbs multiplication from $\mathbb{Z}$, but *not* addition.

The monomials provide another important example of absorption.

Example 1.59. Consider the monoid $\mathbb{M}_2$, and let A be any absorbing subset of it. Suppose that $xy^2, x^3 \in A$, but none of their factors is in A. Since A absorbs from $\mathbb{M}_2$, all the monomial multiples of xy^2 and x^3 are also in A. We can illustrate this with a **monomial diagram**:



Every dot represents a monomial in A; the dot at $(1,2)$ represents the monomial xy^2, and the dots above it represent xy^3, xy^4, ... Notice that multiples of xy^2 and x^3 lie *above and to the right* of these monomials.

The diagram suggests that we can identify special elements of subsets that absorb from the monomials.

Definition 1.60. Suppose A is an absorbing subset of $\mathbb{M}_n$, and $t \in A$. If no other $u \in A$ divides t, then we call t a **generator** of A.

In the diagram above, xy^2 and x^3 are the generators of an absorbing subset corresponding to the monomials covered by the shaded region, extending indefinitely upwards and rightwards. The name "generator" is apt, because every monomial multiple of xy^2 and x^3 is also in A, but nothing "smaller" is in A, in the sense of divisibility.

Lemma 1.61. Every absorbing subset of $\mathbb{M}_n$ has at least one generator.

Proof. Let A be an absorbing subset of $\mathbb{M}_n$, and let $t \in A$. If t is a generator of A, then we are done; otherwise, we can find $u \in A$ such that $t \neq u$ and u divides t. Without loss of generality, we may assume that $\deg_{x_1} u < \deg_{x_1} t$, so choose u such that $\deg_{x_1} u$ is minimal.

If u is a generator of A, then we are done; otherwise, we can find $v \in A$ such that $v \neq u$ and v divides u. Divisibility is transitive, so v also divides t. Since $\deg_{x_1} u$ is minimal among the divisors of t, $\deg_{x_1} u \leq \deg_{x_1} v$, so a different variable must have smaller degree, say $\deg_{x_2} v < \deg_{x_2} u$.

If v is a generator of A, then we are done; otherwise, we can find $w \in A$ such that $w \neq v$ and w divides v and, as before, it divides both t and u, as well. Moreover, if $w = u$ then by substitution $u \mid v$, but this contradicts $\deg_{x_2} v < \deg_{x_2} u$, so $w \neq u$. Since $\deg_{x_1} u$ is minimal among the divisors of t and $\deg_{x_2} v$ is minimal among the divisors of u, $\deg_{x_1} u \leq \deg_{x_1} w$ and $\deg_{x_2} v \leq \deg_{x_2} w$, so a different variable must have smaller degree, say $\deg_{x_3} w < \deg_{x_3} v$.

There are only n variables, and each one is consumed with each new divisor, so if we proceed in this fashion, eventually we run out: the final element of A chosen will be a generator. □

This leads us to a remarkable result.

Monomials and Ideal Nim

For this section, you'll want to review the game Ideal Nim on page 1. The game's rules are inspired from computations some algebraists perform on monomial diagrams. At the time, we pointed out an obvious question:

Must every game end?

If you play a few games in two dimensions, it seems the game must in fact end, and perhaps you can even explain why it does in two dimensions. But what about arbitrary dimension? It's not so easy to visualize that, let alone play a few games and see what happens.

The mathematician Leonard Dickson asked a similar question about a century ago, which we can phrase this way:[16]

Theorem 1.62 (Dickson's Lemma)**.** Every absorbing subset of $\mathbb{M}_n$ has a finite number of generators.

How are Ideal Nim and Dickson's Lemma related? Recall Exercise 1.53: $\mathbb{M}_n \cong \mathbb{N}^n$. A "generator" of an absorbing subset lies furthest southwest in a monomial diagram, and a game of Ideal

[16] This is not the original lemma. Dickson proved a different, related lemma.

Nim ends when no player can choose a point further south or west than those that exist already. That allows us to formulate and prove the following lemma:

Lemma 1.63. Dickson's Lemma is equivalent to termination of every game of Ideal Nim.

The proof is basically "by picture:" we compare activity in the game with its isomorphic image in a monomial diagram.

Proof. First we show that if every game of Ideal Nim must end, Dickson's Lemma is true. Assume that every game of Ideal Nim must end, and let T be an absorbing subset of $\mathbb{M}_n$. By Lemma 1.61, every absorbing subset of $\mathbb{M}_n$ has at least one generator, so we can certainly agree to play a game of Ideal Nim where every point played is required to be a generator of T. As this is a specific game of Ideal Nim, and we have assumed that every game of Ideal Nim must end, this game, too, must end: but it can end only when we run out of generators of T. Hence T can have only finitely many generators.

Now we show that if Dickson's Lemma is true, then every game of Ideal Nim must end. Assume that Dickson's Lemma is true, and consider a game of Ideal Nim. Our argument will have two main steps: the first an examination of the points eliminated by gameplay, the second an examination of the corresponding points of $\mathbb{M}_n$.

- By the rules of the game, if the point Q is played after the point P, we must be able to find some coordinate i such that $q_i < p_i$; otherwise, Q would lie northeast of P, violating the rules. Let $\mathcal{S}$ be the set of points northeast of those points already played; as we just said, for any $P \in \mathcal{S}$, the fact that P is out of play means any point Q lying northeast of P is also out of play.
- With that in mind, let f be an isomorphism from $\mathbb{N}^n$ to $\mathbb{M}_n$: it maps any tuple $(a_1, \ldots, a_n)$ to $x_1^{a_1} \cdots x_n^{a_n}$. Let T be the image of $\mathcal{S}$ in $\mathbb{M}_n$. For any $t \in T$, the previous paragraph implies that for any $u \in \mathbb{M}_n$, if $t \mid u$ then $u \in T$ also. In other words, T is an absorbing set of $\mathbb{M}_n$. We have assumed Dickson's Lemma, which says that T has a finite number of generators. But any generator t of T must correspond to a point actually played in the game; it cannot lie to the "northeast" of another point, else it would be divisible by some other $v \in T$, contradicting its status as a generator. So at some point the players *must actually play* those points that correspond to the generators of T. Once they do, every possible point of T lies northeast of those generators, so the absorbing subset can expand no further. By isomorphism, neither can $\mathcal{S}$ expand further, so the game must have ended.

□

The equivalence proved in Lemma 1.63 means that:

- if we can show every game of Ideal Nim ends, then Dickson's Lemma must be true; and likewise,
- if we can prove Dickson's Lemma, then every game of Ideal Nim ends.

We will take the first route: we now prove that every game of Ideal Nim must end, and by Lemma 1.63 we get Dickson's Lemma for free! Textbooks usually prove Dickson's Lemma directly, however, so if that interests you, see the appendix after this section's exercises.

Theorem 1.64. Every game of Ideal Nim must end.

The proof requires at least three dimensions to "see" every case, so we interleave its main arguments with a running example to illustrate the point of that part.

Proof. On each turn, let $\mathcal{P}$ be the set of points played thus far. For each $k = 1, \ldots, n$ let $d_k = \min_{P \in \mathcal{P}} p_k$; that is, d_k is the smallest entry in the kth coordinate of a played point. The well ordering of $\mathbb{N}$ means that when we recalculate d_k on subsequent turns, each d_k can decrease only finitely many times. Hence there exists a turn where all the d_k have reached their minimum values. At this point, let $D_k = \max_{P \in \mathcal{P}} d_k$; that is, D_k is the largest entry in the kth coordinate of a played point.

Example 1.65. Suppose the points played are $P = (5,2,6)$, $Q = (8,3,7)$, and $R = (6,6,2)$. Suppose further that we have reached the turn where all the d_k have reached their minimum values. In this case, we have $d_1 = 5$, $d_2 = 3$, and $d_3 = 2$, as well as $D_1 = 8$, $D_2 = 6$, and $D_3 = 7$.

Proof (continued). Suppose gameplay continues, which it sometimes can. Let S be a playable point. We have already reached the minimum values of the coordinates, so $d_k \le s_k$ for each k. On the other hand, we also know that $s_k < D_k$ for at least one k; otherwise, S would be northeast of every point played!

Example 1.65 (continued)**.** The point $S = (5,3,5)$ is playable because it is northeast of no point played: after all, $s_3 < p_3$, $s_1 < q_1$, and $s_2 < r_2$. However, the point $(8,6,7)$ is clearly not playable because it is northeast of all three points played.

Proof (continued). Only finitely many points S satisfy the "squeezing" constraint that $d_k \le s_k < D_k$, so a game that does not terminate must eventually select a playable point T such that $t_k \ge D_k$ for some k. Even this can occur only finitely many times; to see why, let $K = \{k_1, \ldots, k_\ell\}$ be the set of "squeezed" indices k_i with $t_{k_i} < D_{k_i}$. Again, $d_{k_i} \le t_{k_i}$ for each of these, so after playing T any playable point U with $t_{k_i} \le u_{k_i}$ is bounded above by the non-squeezed indices, and in fact we must have $u_{k_i} < t_{k_i}$ for some i (otherwise U is northeast of T).

Example 1.65 (continued)**.** Also playable is $T = (10,5,5)$ becuse $t_3 < p_3, q_3$ and $t_2 < r_2$. Notice that $t_1 > D_1$, but as explained in the proof, any future playable point $(u,5,5)$ must have $5 \le u < 10$, and there are only finitely many of those.

Proof (concluded). So there are only finitely many points whose coordinates are all squeezed between d_k and D_k (the S's), and only finitely many *points we can play* that are not squeezed between them (the T's and U's). Once those points are played, no options are left, so the game ends. □

Write S for the set of points eliminated by gameplay in a game of Ideal Nim, and T for the corresponding set of monomials. Both are absorbing sets in their respective monoids. In addition, S and T "expand" with each turn of gameplay, but we just proved that they can expand only finitely many times. This hints at an important concept that we will exploit greatly in Chapter 8.

Suppose $A_1, A_2, \ldots$ absorb from a monoid M, and $A_i \subseteq A_{i+1}$ for each $i \in \mathbb{N}^+$.
Claim: Show that $A = \bigcup_{i=1}^{\infty} A_i$ also absorbs from M.

1. Let $m \in M$ and $a \in A$.
2. By _____, there exists $i \in \mathbb{N}^+$ such that $a \in A_i$.
3. By _____, $ma \in A_i$.
4. By _____, $A_i \subseteq A$.
5. By _____, $ma \in A$.
6. Since _____, this is true for all $m \in M$ and all $a \in A$.
7. By _____, A also absorbs from M.

Figure 1.2. Material for Exercise 1.70

Definition 1.66. Let M be a monoid. Suppose that, for any absorbing subsets $A_1, A_2, \ldots$ of M, we can guarantee that the sets expand only finitely many times — that is, if $A_1 \subseteq A_2 \subseteq \cdots$, then there is some $n \in \mathbb{N}^+$ such that $A_n = A_{n+1} = \cdots$. In this case, we say that M satisfies the **ascending chain condition**, or that M is **Noetherian**.

Dickson's Lemma is a perfect illustration of the Ascending Chain Condition. It also illustrates a relationship between the Ascending Chain Condition and the well-ordering of the integers: we used the well-ordering of the integers repeatedly to prove that $\mathbb{M}_n$ is Noetherian. This property, too, will reappear often.

Exercises.

Exercise 1.67. Show that $2\mathbb{Z}$ is not an absorbing subset of $\mathbb{Z}$ under addition.

Exercise 1.68. Let $d \in \mathbb{Z}$ and $A = d\mathbb{Z}$. Show that A is an absorbing subset of $\mathbb{Z}$ under multiplication.

Exercise 1.69. Let $M = \mathbb{Z}^{2\times 2}$. This is a monoid by 1.8. Let A be the set of 2×2 integer matrices with zero in the first column; that is,

$$\begin{pmatrix} 0 & 3 \\ 0 & -1 \end{pmatrix} \in A, \quad \text{but} \quad \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix} \notin A.$$

Show that A is an absorbing subset of M.

Exercise 1.70. Fill in each blank of Figure 1.70 with its justification.

Definition 1.71. Let S be a set, and $\circ$ an operation. We say that $(S, \circ)$ is a **semigroup** if its operation is closed and associative, although it might not have an identity element. A "semigroup" is "half a group", in that it satisfies half of the properties of a group.

Exercise 1.72. Explain why an absorbing subset of a monoid is a semigroup. Indicate why it might not be a monoid.

Exercise 1.73. Suppose M is a monoid. Let $\mathcal{A} = \{A_1, A_1, \ldots\}$ be a set of absorbing subsets of M, with the additional condition that if $A_i, A_j \in \mathcal{A}$, then $A_i \cup A_j \in \mathcal{A}$, also.

(a) Show that $\mathcal{A}$ is a semigroup under the operation of union ($\cup$).

(b) $\mathcal{A}$ might not be a monoid, but if not it always becomes one by adding only one subset of M to it. Which subset?

(c) Suppose $\mathcal{A}$ is a monoid. What would an absorbing subset of $\mathcal{A}$ look like?

(c) Suppose $\mathcal{A}$ is a monoid, and M is Noetherian. Must $\mathcal{A}$ also be Noetherian?

Exercise 1.74. Our variables here are commutative.

(a) Show that $\mathbb{M}_m \times \mathbb{M}_n$ is Noetherian.

(b) Suppose M and N are Noetherian monoids. Must $M \times N$ be a Noetherian monoid?

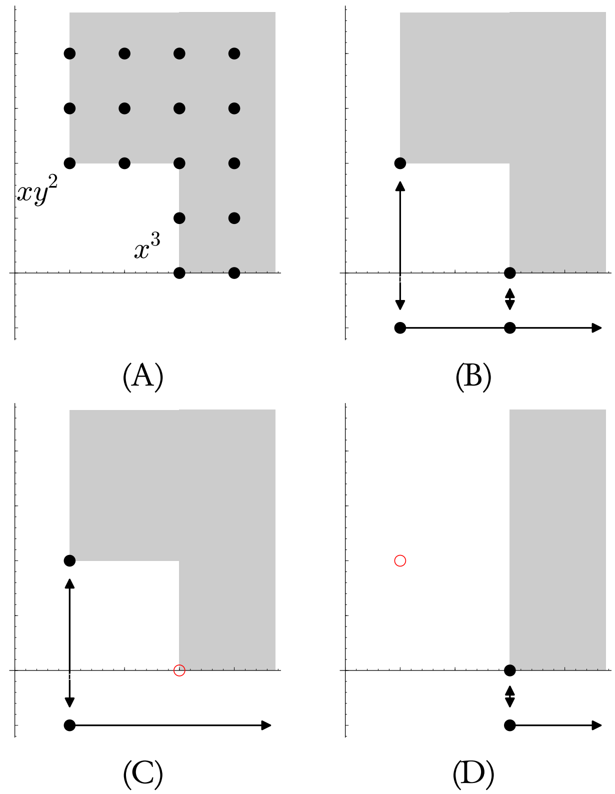


Figure 1.3. Illustration of the proof of Dickson's Lemma.

Traditional proof of Dickson's Lemma

The proof is a little complicated, so we'll illustrate it using some monomial diagrams. In Figure 3(A), we see an absorbing subset A. (The same as you saw before.) Essentially, the argument *projects* A down one dimension, as in Figure 3(B). In this smaller dimension, an argument by induction allows us to choose a finite number of generators, which correspond to elements of A, illustrated in Figure 3(C). These corresponding elements of A are always generators of A, but they might not be *all* the generators of A, shown in Figure 3(C) by the red circle. In that case, we take the remaining generators of A, use them to construct a new absorbing subset, and project again to obtain new generators, as in Figure 3(D). The thing to notice is that, in Figures 3(C) and 3(D), the y-values of the new generators decrease with each projection. This cannot continue indefinitely, as the y-values are natural numbers, and $\mathbb{N}$ is well-ordered. Eventually, the process must stop.

Proof. Let A be an absorbing subset of $\mathbb{M}_n$. We proceed by induction on the dimension, n.

For the *inductive base,* assume $n = 1$. Let S be the set of exponents of monomials in A. Since $S \subseteq \mathbb{N}$, it has a minimal element; call it a. By definition of S, $x^a \in A$. We claim that x^a is, in fact, the one generator of A. To see why, let $u \in A$; by definition of A, $u = x^b$ for some $b \in \mathbb{N}$. Suppose that $u \mid x^a$; by definition of monomial divisibility, $b \leq a$. Since $u \in A$, it follows that $b \in S$. Since a is the *minimal* element of S, $a \leq b$. We already knew that $b \leq a$, so it must be that $a = b$. The claim is proved: no other element of A divides x^a. Thus, x^a is a generator. Any other generators must come from other variables, but $n = 1$, so there are no other variables.

For the *inductive hypothesis,* assume that any absorbing subset of $\mathbb{M}_{n-1}$ has a finite number of generators.

For the *inductive step*, we use A to construct a sequence of absorbing subsets of $\mathbb{M}_{n-1}$ in the following way.

- Let B_1 be the **projection** of A onto $\mathbb{M}_{n-1}$. Technically, this is the set of all monomials in $\mathbb{M}_{n-1}$ such that $t \in B_1$ implies that $t x_n^a \in A$ for some $a \in \mathbb{N}$, but if you simply keep the image of Figure 3(B) in mind, you have the idea.
  We claim that B_1 absorbs from $\mathbb{M}_{n-1}$. To see why, let $t \in B_1$, and let $u \in \mathbb{M}_{n-1}$ be any monomial multiple of t, say $u = tv$. By definition, there exists $a \in \mathbb{N}$ such that $t x_n^a \in A$. Since A absorbs from $\mathbb{M}_n$, and $u \in \mathbb{M}_{n-1} \subsetneq \mathbb{M}_n$, absorption implies that $v\left(t x_n^a\right) \in A$. By substitution, $u x_n^a \in A$, and the definition of B_1 tells us that $ut \in B_1$. We took an arbitrary element t of B_1 and an arbitrary multiple u of t, and found that $u \in B_1$; this means B_1 absorbs from $\mathbb{M}_{n-1}$.
  This result is important! By the inductive hypothesis, B_1 has a finite number of generators; call them $\{t_1, \ldots, t_m\}$. Each of these generators corresponds to an element of A. For instance, t_1 might correspond to $t_1 x_n^{20}$, $t_1 x_n^3$, $t_1 x_n^7$, and so forth. For each i, let U_i be the set of powers of x_n such that $t_1 x_n^i \in A$; in the example of the previous sentence, $U_1 = \{20, 3, 7, \ldots\}$. By definition, each $U_i \subseteq \mathbb{N}$; by the well-ordering property, it has a smallest element; call it a_i. Let $T_1 = \left\{t_1 x_n^{a_1}, \ldots, t_m x_n^{a_m}\right\} \subsetneq A$.
  We now claim that every element of T_1 is a generator of A. Why? By way of contradiction, assume that some $u \in T_1$ is not a generator of A. The definition of a generator means that there exists some other $v \in A$ that divides u, but $u \nmid v$. Choose $a, b \in \mathbb{N}$ such that $u = t x_n^a$ and $v = t' x_n^b$ for some $a, b \in \mathbb{N}$; by projection, $t, t' \in B_1$.
  Things fall apart here! After all, if $v \mid u$, then $t' \mid t$. Both t' and t are in the projection of A to B_1, so T_1 contains generators for each. We know that $t \in T_1$, so it generates itself $(t \cdot 1 = t)$; choose $\widehat{t} \in T_1$ to generate t'. Thus, $\widehat{t} \mid t'$, and since $t' \mid t$, we have $\widehat{t} \mid t$ (Exercise 1.13(a)). However, t is a generator; by definition, one generator cannot divide a *different* generator. Hence, the two generators must be equal: $t = \widehat{t}$. By Exercise 1.13(b), $t = t'$. By substitution, $u = t x_n^a$ and $v = t x_n^b$; since $v \mid u$, $a \leq b$, and since $u \nmid v$, $a \neq b$. Putting them together, $a < b$. This contradicts the definition of T_1, whose elements we chose to have the *smallest* power of x_n.
- If T_1 is a complete list of the generators of A, then we are done. Otherwise, we know that T_1 generates all the elements of B_1, which is the projection of A onto $\mathbb{M}_{n-1}$, and the elements of T_1 had the *smallest* powers of x_n. Let $A^{(1)}$ be subset of A that *is not* generated by T_1. Let B_2 be the absorbing subset generated by the projection of $A^{(1)}$ onto $\mathbb{M}_{n-1}$. As before, B_2 absorbs from $\mathbb{M}_{n-1}$, and the inductive hypothesis implies that it has a finite number of generators, which correspond to a set T_2 of generators of $A^{(1)}$. These elements have larger powers of x_n than those in T_1, but since the elements of T_1 do not generate them, they must have smaller powers in some other variable, as in Figure 3(C).
- As long as T_i is not a complete list of the generators of A, we continue building
  - a subset $A^{(i)}$ of A whose elements are *not* generated by $T_1 \cup \cdots \cup T_i$;
  - an absorbing subset B_{i+1} whose elements are the projections of $A^{(i)}$ onto $\mathbb{M}_{n-1}$, and
  - a set T_{i+1} of generators of A that correspond to generators of B_{i+1}.

Can this process continue indefinitely? No, it cannot. First, if $t \in T_{i+1}$, then write it as $t = t' x_n^a$. On the one hand,

$$t \in A^{(i)} \subsetneq A^{(i-1)} \subsetneq \cdots A^{(1)} \subsetneq A,$$

so t' was an element of every B_j such that $j \leq i$. For each j, t' was divisible by at least one generator u'_j of B_j. However, t was *not* in the absorbing subsets generated by $T_1, \ldots, T_i$. So the $u_j \in T_j$ corresponding to u'_j does *not* divide t. Write $t = x_1^{a_1} \cdots x_n^{a_1}$ and $u_j = x_1^{b_1} \cdots x_n^{b_n}$. Since $u'_j \mid t'$, $b_k \leq a_k$ for each $k = 1, \ldots, n-1$. Since $u \nmid t$, $b_n > a_n$. As i increases, the minimal degree a_n of x_n in each T_i must decrease, as the larger values are chosen earlier.

We have created a strictly decreasing sequence of natural numbers (the minimal degree a_n of x_n). By the well-ordering property, such a sequence cannot continue indefinitely. Thus, we cannot create sets T_i containing new generators of A indefinitely; there are only finitely many such sets. In other words, A has a finite number of generators. □

Chapter 2:
Groups

In Chapter 1, we described monoids. When we look at many monoids, we find that there is no way to undo an operation using the same operation. For instance, the *natural numbers* are a monoid under addition, and there is no way to undo the addition using another addition: while $5+(-5)=0$, the number we used, -5, is not natural. This is not a difficulty with *integers* under addition, as -5 is an integer. Yet when we look at the integers under *multiplication*, we encounter the same issue: we cannot "undo" the multiplication of 2 by 3 using another multiplication: while $6\times\frac{1}{3}=2$, the number $\frac{1}{3}$ is not an integer.

This ability to undo the operation is inherent to a large number of real-world phenomena, so it's a good idea to identify which structures can be used with such problems. This chapter introduces a special kind of monoid, called a *group*. Every element of a group has an inverse element *in the group itself*. We also introduce a number of important groups to which the text returns time and again: the cyclic groups (Section 2.3), the symmetries of a triangle (Section 2.2), and the roots of unity (Section 2.4).

2.1: Groups

This first section looks only at the fundamental properties of groups.

Precise definition, first examples

Definition 2.1. Let G be a set, and $\circ$ a binary operation on G. The pair $(G,\circ)$ is a **group** if it satisfies the three properties of a monoid as well as one more property:

(inverses) each element of the group has an **inverse**; that is, for any $x\in G$ we can find $y\in G$ such that $x\circ y=y\circ x=e$.

We may also say that G is a **group under** $\circ$. We say that $(G,\circ)$ is an **abelian group** if it also satisfies a fifth property:

(commutative) the operation is commutative; that is, $xy=yx$ for all $x,y\in G$.

Notation 2.2. If the operation is addition, we may refer to the group as an **additive group** or a **group under addition**. We also write $-x$ instead of x^{-1}, and $x+(-y)$ or even $x-y$ instead of $x+y^{-1}$, keeping with custom. You have to be careful with this, since $(a-b)-c\neq a-(b-c)$. (Try it with a few numbers if you doubt me.) Additive groups are normally abelian.

If the operation is multiplication, we may refer to the group as a **multiplicative group** or a **group under multiplication**.

When the operation is understood from context (and usually it is), we write G rather than $(G,+)$ or $(G,\times)$ or $(G,\circ)$. We still write $(G,+)$ when we want to emphasize that the operation is addition.

Example 2.3. While $\mathbb{N}$ is a monoid under addition, it is not a group, as the inverse property is not satisfied. The failure is not that inverses don't exist at all; they do! Rather, the failure is that *the inverses are not in* $\mathbb{N}$ *itself!* It's not enough to have an inverse in *some* set; the inverse must be in the same set! For this reason, $\mathbb{N}$ is not a group.

Example 2.4. Certainly $\mathbb{Z}$ is an additive group; in fact, it is abelian. Why?

- It is a monoid under addition.
- Every integer has an additive inverse, and that inverse lies *in* $\mathbb{Z}$.
- Addition of integers is commutative.

Example 2.5. Recall the "boolean and" monoid $B = \{F, T\}$ from Example 1.9. Is it a group? We already know it satisfies the three properties of a monoid, so we need merely determine whether every element has an inverse.

When we perform the operation on an element and its inverse, the result is the identity. So deciding whether the inverse property is satisfied requires us firts to ask ourselves, "What is the identity of the 'boolean and' set? Again, we found in Example 1.9 that the identity is T, so we have to show that there exist $x, y \in B$ such that

$$F \wedge x = T \quad \text{and} \quad T \wedge y = T \ .$$

The second one is easy:

$$T \wedge T = T \ ,$$

but the first one is impossible, because

$$F \wedge x = F$$

regardless of the value of x! Hence "boolean and" is not, in fact, a group.

Example 2.6. In addition to $\mathbb{Z}$, the following sets are abelian groups under addition.

- the set $\mathbb{Q}$ of **rational numbers**;
- the set $\mathbb{R}$ of **real numbers**; and
- if $S = \mathbb{Z}, \mathbb{Q}$, or $\mathbb{R}$, the set $S^{m\times n}$ of $m \times n$ matrices whose elements are in S. (It's important here that the operation is *addition*.)

None of these is a group under multiplication; in some cases, multiplicative inverses don't even exist! However, we can sometimes restrict a monoid to a subset, and obtain a group. For example, the set of *invertible* $n \times n$ matrices with elements in $\mathbb{Q}$ or $\mathbb{R}$ is a multiplicative group. We leave the proof to the exercises, but this fact builds on properties you learned in linear algebra, such as those described in Section 0.3.

Definition 2.7. We call the set of invertible $n \times n$ matrices with elements in $\mathbb{R}$ the **general linear group of degree** n, abbreviated $\mathrm{GL}_n(\mathbb{R})$.

Order of a group, Cayley tables

The next few tools are useful to analyze groups.

Definition 2.8. Let S be any set. We write $|S|$ to indicate the number of elements in S, and say that $|S|$ is the **size** or **cardinality** of S. If there is an infinite number of elements in S, then we write $|S| = \infty$. We also write $|S| < \infty$ to indicate that $|S|$ is finite.

If S is also group, we call its size the **order of** G. A group has finite order if $|G| < \infty$ and infinite order if $|G| = \infty$.

Here are three examples of finite groups, all of order 2.

Example 2.9. The sets

$$\{1,-1\}, \quad \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$
$$\text{and} \quad \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

are all groups under multiplication:

- In the first group, the identity is 1, and -1 is its own inverse; verifying closure is straightforward, and you know from arithmetic that the associative property holds.
- In the second and third groups, the identity is the identity matrix; each matrix is its own inverse; verifying closure is straightforward, and you know from linear algebra that the associative property holds.

I will now make an extraordinary claim:

Claim 1. For all intents and purposes, there is only one group of order two.

This claim may seem preposterous; after all, the example above has three completely different groups of order two. In fact, the claim is quite vague, because we're using vague language. After all, what is meant by the phrase, "for all intents and purposes"? Basically, we meant that:

- group theory cannot distinguish between the groups *as groups;* or,
- their multiplication table (or addition table, or whatever-operation table) has the same structure.

If you read the second characterization and think, "he means they're isomorphic!", then pat yourself on the back. We won't look at this notion seriously until Chapter 4, but Chapter 1 gave you a rough idea of what that meant: the groups are identical *as groups*. The precise phrase we will use later is that there is only one group of order two, "up to isomorphism."

We prove this claim in a "brute force" manner, by looking at the table generated by the operation of the group. Now, this phrase, "the table generated by the operation of the group," is an ungainly mouthful. Since the name of the table depends on the operation (multiplication table, addition table, etc.), we describe all of them with the following term.

Definition 2.10. The table listing all results of the operation of a monoid or group is its **Cayley table**.

So we will prove the claim by building a Cayley table for a "generic" group of order two. We will show that there is only one possible way to construct such a table. As a consequence, regardless of the set and its operation, every group of order 2 behaves exactly the same way. *It does not matter one whit* what structure the elements of G have, or the convoluted procedure we use to simplify the group operation. If there are only two elements, and it's a group, then *it always works the same*. Why?

Example 2.11. Let G be an arbitrary group of order two. By definition, it has an identity, so write $G = \{e, a\}$ where e represents the known identity, and a the other element.

We did *not* say that e represents the *only* identity. For all we know, a might also be an identity; is that possible? In fact, it is not, but *why* not? Remember that a group is a monoid. We showed in Proposition 2.14 that the identity of a monoid is unique; thus, the identity of a group is unique; thus, there can be only one identity, e. Notice the fulfillment of our promised power of abstraction: we proved a property of *all* monoids; groups are a monoid; thus, that property holds also for groups!

Now we build the Cayley table. We *have* to assign $a \circ a = e$. *Why?*

- To satisfy the identity property, we must have $e \circ e = e$, $e \circ a = a$, and $a \circ e = a$.
- To satisfy the inverse property, a must have an inverse: some $b \in G$ such that $a \circ b = e$. It can't be e, since $a \circ e = a$; that is, $a \circ e \neq e$. The only inverse possible is a itself! That is, $a^{-1} = a$. (Read that as, "the inverse of a is a.") So $a \circ a^{-1} = a \circ a = e$.

The Cayley table of our group looks like this.

$\circ$	e	a
e	e	a
a	a	e

No room remains for ambiguity, so we have completely determined the table! The only assumption we made was that G is a group of order two. This table must apply to *any* group of order two!

In Definition 2.1 and Example 2.11, the symbol $\circ$ is a placeholder for any operation. We assumed nothing about its actual behavior, so it can represent addition, multiplication, or other operations that we have not yet considered. From the point of group theory, all three groups we showed in Example 2.9 are in fact the same, so whatever we do for one, we do also for the rest. Behold the power of abstraction!

Other elementary properties of groups

Notation 2.12. We adopt the following convention:

- If we know only that G is a group under some operation, we write $\circ$ for the operation and proceed as if the group were multiplicative, so that xy is shorthand for $x \circ y$.
- If we know that G is a group and a symbol is provided for its operation, we *usually* use that symbol for the group, *but not always.* Sometimes we treat the group as if it were multiplicative, writing xy instead of the symbol provided.
- We reserve the symbol $+$ exclusively for additive groups.

The following fact looks obvious—but remember, we're talking about elements of *any* group, not merely the sets you have worked with in the past.

Proposition 2.13. Let G be a group and $x \in G$. Then $\left(x^{-1}\right)^{-1} = x$. If G is additive, we write instead that $-(-x) = x$.

Proposition 2.13 says that the inverse of the inverse of x is x itself; that is, if y is the inverse of x, then x is the inverse of y.

Proof. You prove it! See Exercise 2.21. □

Proposition 2.14. The identity of a group is both two-sided and unique; that is, every group has exactly one identity. Also, the inverse of an element is both two-sided and unique; that is, every element has exactly one inverse element.

We can prove this directly using a technique similar to that of Theorem 1.11 on page 52, but there is a "more visual" way to do this. We saw from Example 2.11 that the structure of a group compels certain assignments for the operation, and fills in parts of the Cayley table. We can take this a little further:

Theorem 2.15. Let G be a group of finite order, and let $a, b \in G$. Then a appears exactly once in the row or column of the Cayley table that is headed by b.

This is *not* necessarily true for a monoid; see Exercise 2.31. So, something special about groups gives us this property. Only one thing is special about groups: the property of containing all elements' inverses. We see this in the proof, which we could also prove directly, but again appeal to a "more visual" approach.

Proof. We proceed by contradiction. Suppose a appears at least twice in the row of the Cayley table headed by b. That means we could find distinct $c, d \in G$ such that the Cayley table looked something like this:

$\circ$	$\cdots$	c	$\cdots$	d	$\cdots$
$\vdots$					
b	$\cdots$	a	$\cdots$	a	$\cdots$
$\vdots$					

That is, $bc = a$ and $bd = a$. By substitution, $bc = bd$. Multiply on both sides by b^{-1} to obtain

$$b^{-1}(bc) = b^{-1}(bd),$$

apply the associative property to rewrite as

$$\left(b^{-1}b\right)c = \left(b^{-1}b\right)d,$$

and apply the property of an inverse to rewrite again as

$$ec = ed.$$

The identity property means that $c = d$.

This can't be right; we chose distinct c and d to be distinct, but they can't be distinct if they're equal. This contradiction suggests that we can't choose c and d in this fashion after all, and a can appear only once in the row of the Cayley table headed by b. A similar argument shows that a can appear only once in the column of the Cayley table headed by b.

We still have to show that a appears in at least one row of the addition table headed by b. This follows from the fact that each row of the Cayley table contains $|G|$ elements. What applies to a above applies to the other elements, so each element of G can appear at most once. Thus, if we do not use a, then only $n-1$ pairs are defined, which contradicts either the definition of an operation (bx must be defined for all $x \in G$) or closure (that $bx \in G$ for all $x \in G$). Hence a must appear at least once. □

You might notice that the "meat" of this proof suggests a more general property:

In any group, if $a = bc$ and $a = bd$, then $c = d$.

You can prove this statement using exactly the same approach as in the proof above; see Exercise 2.26. Again, this is *not* true about monoids.

Let's apply these ideas to our original question: whether identities and inverses are unique. We already pointed out that the identity of G is unique, since G is a monoid, and the identity of a monoid is unique. What about the inverse?

Proof of Proposition 2.14. Let G be a group, and let $x \in G$. Let y be an inverse of x; that is,

$$xy = e \quad \text{and} \quad yx = e\ .$$

(Notice that we use multiplicaiton as the "generic" operation, even though we did not say the operation was multiplication.) By the identity property, $ex = x$. Now suppose we have "another" inverse of x, called w; that is,

$$xw = e \quad \text{and} \quad wx = e\ .$$

By substitution,

$$xy = e = xw,$$

and by linking the first and last expressions in the equality,

$$xy = xw.$$

Multiply both sides by x^{-1} *on the left* (this is important) and we have

$$x^{-1}(xy) = x^{-1}(xw)\,.$$

By the associative property,

$$\left(x^{-1}x\right)y = \left(x^{-1}x\right)w\ ,$$

which we can simplify as

$$ey = ew\ , \quad \text{or,} \quad y = w\ .$$

We have shown that the "second" inverse w is in fact the first inverse, y. We chose w arbitrarily among "all" the inverses of x, so all of them must the same. In short, there is only one inverse of x in G, and that is e. □

As with monoids, we can build new groups from old, using Cartesian products. You will show in the exercises that the direct product of groups is also a group.

Let G be a group, and $x \in G$.
Claim: $\left(x^{-1}\right)^{-1} = x$; or, if the operation is addition, $-(-x) = x$.
Proof:
1. By _____, $x \cdot x^{-1} = e$ and $x^{-1} \cdot x = e$.
2. By _____, $\left(x^{-1}\right)^{-1} = x$.
3. Negative are merely how we express opposites when the operation is addition, so $-(-x) = x$.

Figure 2.1. Material for Exercise 2.21

Definition 2.16. Let $G_1, \ldots, G_n$ be groups. The **direct product** of $G_1, \ldots, G_n$ is the Cartesian product $G_1 \times \cdots \times G_n$ together with the operation $\otimes$ such that for any $(g_1, \ldots, g_n)$ and $(h_1, \ldots, h_n)$ in $G_1 \times \cdots \times G_n$,

$$(g_1, \ldots, g_n) \otimes (h_1, \ldots, h_n) = (g_1 h_1, \ldots, g_n h_n),$$

where each product $g_i h_i$ is performed according to the operation of G_i. In other words, the direct product of *groups* generalizes the direct product of *monoids*.

Exercises.

Exercise 2.17. Explain why every group of 2 elements is abelian.

Exercise 2.18. Show that every group of order 3 has the same structure.

Exercise 2.19. *Not* every group of order 4 has the same structure, because there are two Cayley tables with different structures. One of these groups is the **Klein four-group**, where each element is its own inverse; the other is called a **cyclic group** of order 4, where not every element is its own inverse. Determine the Cayley tables for each group.

Exercise 2.20. Let G be a group, and $x, y \in G$. Show that $xy^{-1} \in G$.

Exercise 2.21.
(a) Fill in each blank of Figure 2.21 with the appropriate justification or statement.
(b) Why should someone think to look at the product of x and x^{-1} in order to show that $\left(x^{-1}\right)^{-1} = x$?

Exercise 2.22. Explain why $(\mathbb{M}, \times)$ is not a group.

Exercise 2.23. We stated that, in general, $(a-b)-c \neq a-(b-c)$, even in an additive group. Give an example of this inequality with precise values of a, b, and c.

Exercise 2.24. Explain why the set $\mathcal{C}_{\prec}(X)$ of all subsets of a set X that are convex with respect to a linear ordering $\prec$ is not a group. (See Exercise 1.26.)

Exercise 2.25. Is $\left(\mathbb{N}^+, \mathrm{lcm}\right)$ a group? (See Exercise 1.27.)

Exercise 2.26. Let G be a group, and $x, y, z \in G$. Show that if $xz = yz$, then $x = y$; or if the operation is addition, that if $x + z = y + z$, then $x = y$.

Exercise 2.27. Show in detail that $\mathbb{R}^{2\times 2}$ is an additive group.

Exercise 2.28. Recall the Boolean-or monoid $(B, \vee)$ from Exercise 1.16. Is it a group? If so, is it abelian? Explain how it justifies each property. If not, explain why not.

Exercise 2.29. Recall the Boolean-xor monoid $(B, \oplus)$ from Exercise 1.17. Is it a group? If so, is it abelian? Explain how it justifies each property. If not, explain why not.

Exercise 2.30. In Section 1.1, we showed that F_S, the set of all functions, is a monoid for any S.
(a) Show that $F_{\mathbb{R}}$, the set of all functions on the real numbers $\mathbb{R}$, is *not* a group.
(b) Describe a subset of $F_{\mathbb{R}}$ that *is* a group. Another way of looking at this question is: what restriction would you have to impose on any function $f \in F_S$ to fix the problem you found in part (a)?

Exercise 2.31. Indicate a monoid you have studied that does not satisfy Theorem 2.15. That is, find a monoid M such that (i) M is finite, and (ii) there exist $a, b \in M$ such that in the the Cayley table, a appears at least twice in a row or column headed by b.

Exercise 2.32. Let $n \in \mathbb{N}^+$. Let $G_1, G_2, \ldots, G_n$ be groups, and consider

$$\begin{aligned}\prod_{i=1}^{n} G_i &= G_1 \times G_2 \times \cdots \times G_n \\ &= \{(a_1, a_2, \ldots, a_n) : a_i \in G_i \ \forall i = 1, 2, \ldots, n\}\end{aligned}$$

with the operation $\dagger$ where if $x = (a_1, a_2, \ldots, a_n)$ and $y = (b_1, b_2, \ldots, b_n)$, then

$$x \dagger y = (a_1 b_1, a_2 b_2, \ldots, a_n b_n),$$

where each product $a_i b_i$ is performed according to the operation of the group G_i. Show that $\prod_{i=1}^{n} G_i$ is a group, and notice that this shows that the direct product of groups is a group, as claimed above. (We used $\otimes$ instead of $\dagger$ there, though.)

Exercise 2.33. Let $m \in \mathbb{N}^+$.
(a) Show in detail that $\mathbb{R}^{m\times m}$ is a group under addition.
(b) Show by counterexample that $\mathbb{R}^{m\times m}$ is *not* a group under multiplication.

Exercise 2.34. Let $m \in \mathbb{N}^+$. Explain why $\mathrm{GL}_m(\mathbb{R})$ satisfies the identity and inverse properties of a group.

Exercise 2.35. Let $\mathbb{R}^{>0} = \{x \in \mathbb{R} : x > 0\}$, and $\times$ the ordinary multiplication of real numbers. Show that $\left(\mathbb{R}^{>0}, \times\right)$ is a group.

Exercise 2.36. Define $\mathbb{Q}^*$ to be the set of non-zero rational numbers; that is,

$$\mathbb{Q}^* = \left\{\frac{a}{b} : a, b \in \mathbb{Z} \text{ where } a \neq 0 \text{ and } b \neq 0\right\}.$$

Claim: Any two elements a, b of any group G satisfy $(ab)^{-1} = b^{-1}a^{-1}$.
Proof:
1. Let _____.
2. By the _____, _____, and _____ properties of groups,

$$(ab)\,b^{-1}a^{-1} = a\left(b\cdot b^{-1}\right)a^{-1} = aea^{-1} = aa^{-1} = e.$$

3. We chose _____ arbitrarily, so this holds for all elements of all groups, as claimed.

Figure 2.2. Material for Exercise 2.37

Show that $\mathbb{Q}^*$ is a multiplicative group.

Exercise 2.37. (a) Let $G = \mathrm{GL}_2(\mathbb{R})$. Show that there exist $a, b \in G$ such that $(ab)^{-1} \neq a^{-1}b^{-1}$.
(b) Suppose H is an arbitrary group.
   (i) Explain why we cannot assume that for every $a, b \in H$, $(ab)^{-1} = a^{-1}b^{-1}$.
   (ii) Fill in the blanks of Figure 2.37 with the appropriate justification or statement.

Exercise 2.38. Let $\circ$ denote the ordinary composition of functions, and consider the following functions that map any point $P = (x, y) \in \mathbb{R}^2$ to another point in $\mathbb{R}^2$:

$$I(P) = P,$$
$$F(P) = (y, x),$$
$$X(P) = (-x, y),$$
$$Y(P) = (x, -y).$$

(a) Let $P = (2,3)$. Label the points P, $I(P)$, $F(P)$, $X(P)$, $Y(P)$, $(F \circ X)(P)$, $(X \circ Y)(P)$, and $(F \circ F)(P)$ on an x-y axis. (Some of these may result in the same point; if so, label the point twice.)
(b) Show that $F \circ F = X \circ X = Y \circ Y = I$.
(c) Show that $G = \{I, F, X, Y\}$ is *not* a group.
(d) Find the smallest group $\overline{G}$ such that $G \subset \overline{G}$. While you're at it, construct the Cayley table for $\overline{G}$.
(e) Is $\overline{G}$ abelian?
(f) With $P = (2,3)$, plot $f(P)$ for every $f \in \overline{G}$. Do you notice anything about the geometry of $\overline{G}$, which was lacking from the geometry of G?

Definition 2.39. 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.

Exercise 2.40. (a) Explain why $[x, y] = e$ iff x and y commute.

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

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

Figure 2.3. Material for Exercise 2.40(c)

(b) Show that $[x,y]^{-1} = [y,x]$; that is, the inverse of $[x,y]$ is $[y,x]$.
(c) Show that $\left(g^z\right)^{-1} = \left(g^{-1}\right)^z$; 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 2.40 with the appropriate justification or statement.

2.2: The symmetries of a triangle

In this section, we show that the symmetries of an equilateral triangle form a group. We call this group D_3. This group *is not abelian.* You already know that groups of order 2, 3, and 4 are abelian; in Section 3.3 you will learn why a group of order 5 must also be abelian. Thus, D_3 is the smallest non-abelian group.

Intuitive development of D_3

To describe D_3, start with an equilateral triangle in $\mathbb{R}^2$, with its center at the origin. Intuitively, a "symmetry" is a transformation of the plane that leaves the *triangle* in the same location, even if its *points* are in different locations. "Transformations" include actions like rotation, reflection

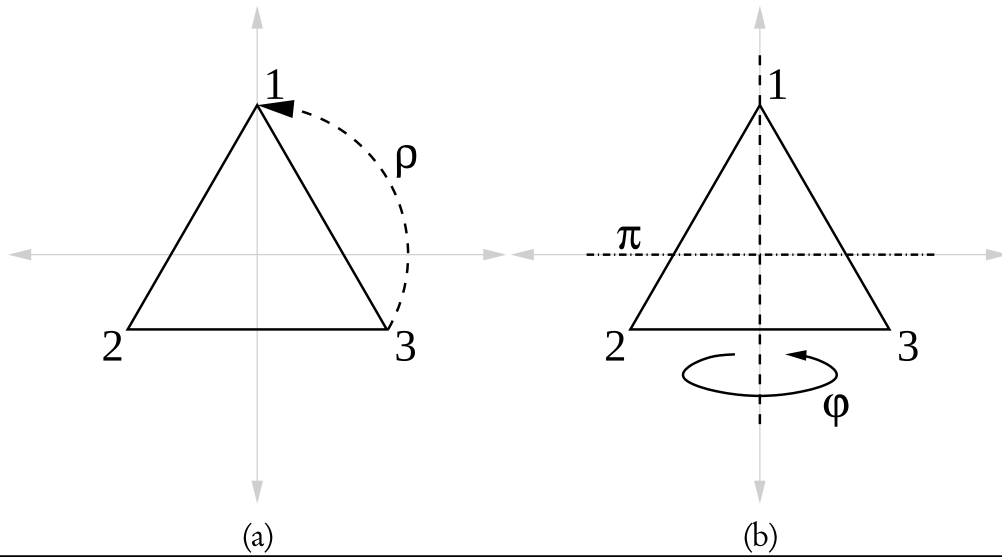


Figure 2.4. Rotation and reflection of the triangle

(flip), and translation (shift). Translating the triangle a nonzero distance won't leave the triangle intact regardless of direction, so translation is not a symmetry.

On the other hand, some rotations and reflections are symmetries. Two obvious symmetries of an equilateral triangle are a 120° clockwise rotation through the origin, and a reflection across the y-axis, which we often call a "flip". We'll call the first of these ρ, and the second φ. See Figure 28.

It is helpful to observe two important properties.

Theorem 2.41. 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 across 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 $\varphi\rho = \rho^2\varphi$. □

Did you 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 more work; the detailed proofs appear later.

And did you notice something interesting about Theorem 2.41? It implies that the operation in D_3 is non-commutative! We have $\varphi\rho = \rho^2\varphi$, and a little logic shows that $\rho^2\varphi \neq \rho\varphi$: thus $\varphi\rho \neq \rho\varphi$.

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.

Theorem 2.42. 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.

Definition 2.43. $D_3 = \{\iota, \varphi, \rho, \rho^2, \rho\varphi, \rho^2\varphi\}$.

Theorem 2.44. D_3 is a group under composition of functions.

Proof. We prove this by showing that all the properties of a group are satisfied. We will start the proof, and leave you to finish it in Exercise 2.45.

Closure: In Exercise 2.45, you will compute the Cayley table of D_3. There, you will see that every composition is also an element of D_3.

Associative: We've already proved this! Way back in Section 1.1, we showed that F_S, the set of functions over a set S, was a monoid under composition for *any* set S. To do that, we had to show that composition of functions was associative. There's no point in repeating that proof here; doing it once ought to be enough for a sane person. Well! The plane of real-valued points is a set, and the symmetries ι, φ, ρ, etc. are themselves functions. After all, 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. What's good for F_S is good for D_3. Behold the power of abstraction!

Identity: We claim that ι is the identity. To see this, let $\sigma \in D_3$ be any symmetry; we need to show that $\iota\sigma = \iota$ and $\sigma\iota = \sigma$. For the first, apply σ to the triangle. Then apply ι, which leaves everything in place, so 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.

Alternately, you could look at the result of Exercise 2.45; you will find that $\iota\sigma = \sigma\iota = \sigma$ for every $\sigma \in D_3$.

Inverse: Intuitively, 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 itself, as well. So, yes, $\sigma^{-1} \in D_3$.

Here, the intuition is a little too imprecise; it isn't *that* obvious that rotation is a one-to-one function. Fortunately, the result of Exercise 2.45 shows that ι, the identity, appears in every row and column. That means that every element has an inverse. □

Exercises.

Unless otherwise specified, ρ and φ refer to the elements of D_3.

Exercise 2.45. The Cayley 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. You may use Theorems 2.41 and 2.42 as well as the same reasoning on the vertices that we used to prove those theorems. *Be explicit about how you determine each product.*

Exercise 2.46. Choose three elements of D_3 that are neither ι, φ, nor ρ. For each σ you have chosen, compute σ^2, σ^3, ... until you arrive at $\sigma^k = \iota$ for some $k \in \mathbb{N}^+$. (You can use your completed Cayley table from the previous exercise to help.) What property do you notice about each value of k?

Exercise 2.47. 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 2.48 answers the first question.

Lemma 2.48. 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 acts on two points in the plane!

We consider $P = (1,0)$ as the first point. Write $\alpha(P) = Q = (q_1, q_2)$; this is the point where α moves Q. The distance between P and the origin is 1. Since $\alpha(0,0) = (0,0)$, the distance between Q and the origin is $\sqrt{q_1^2 + q_2^2}$. Because α preserves distance,

$$1 = \sqrt{q_1^2 + q_2^2},$$

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. Any point on this circle can be parametrized as

$$(\cos t, \sin t)$$

where $t \in [0, 2\pi)$ represents an angle. Hence, $\alpha(P) = (\cos t, \sin t)$.

We consider $R = (0,1)$ as the second point. Write $\alpha(R) = S = (s_1, s_2)$. 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

$$(\cos t - s_1)^2 + (\sin t - s_2)^2 = 2. \tag{8}$$

Properties of trigonometry rewrite (8) as

$$-2(s_1 \cos t + s_2 \sin t) + (s_1^2 + s_2^2) = 1. \tag{9}$$

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

$$\sqrt{s_1^2+s_2^2}=1$$
$$s_1^2+s_2^2=1.$$

Substituting this into (9), we find that

$$\begin{aligned}-2\left(s_1\cos t+s_2\sin t\right)+s_1^2+s_2^2&=1\\-2\left(s_1\cos t+s_2\sin t\right)+1&=1\\-2\left(s_1\cos t+s_2\sin t\right)&=0\\s_1\cos t&=-s_2\sin t.\end{aligned}\tag{10}$$

We can see that $s_1=\sin t$ and $s_2=-\cos t$ is *one* solution to the problem, and $s_1=-\sin t$ and $s_2=\cos t$ 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 (10) 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 (10), this implies that $s_2=\mp\cos t$. Thus there are *two* possible values of s_1 and s_2.

It can be shown (see Exercise 2.52) that α is a linear transformation on the vector space $\mathbb{R}^2$ with the basis $\left\{\vec{P},\vec{R}\right\}=\{(1,0),(0,1)\}$. 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. □

Consider an example of the two basic forms of α, and their effect on the points in the plane.

Example 2.49. Consider the set of points

$$\mathcal{S}=\{(0,2),(\pm2,1),(\pm1,-2)\};$$

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

$$\begin{aligned}
\rho(\mathcal{S}) &= \{\rho(0,2), \rho(-2,1), \rho(2,1), \rho(-1,-2), \rho(1,-2)\} \\
&= \left\{\left(-\sqrt{2}, \sqrt{2}\right), \left(-\sqrt{2} - \frac{\sqrt{2}}{2}, -\sqrt{2} + \frac{\sqrt{2}}{2}\right),\right. \\
&\quad \left(\sqrt{2} - \frac{\sqrt{2}}{2}, \sqrt{2} + \frac{\sqrt{2}}{2}\right), \\
&\quad \left(-\frac{\sqrt{2}}{2} + \sqrt{2}, -\frac{\sqrt{2}}{2} - \sqrt{2}\right), \\
&\quad \left.\left(\frac{\sqrt{2}}{2} + \sqrt{2}, \frac{\sqrt{2}}{2} - \sqrt{2}\right)\right\} \\
&\approx \{(-1.4, 1.4), (-2.1, -0.7), (0.7, 2.1), \\
&\quad (0.7, -2.1), (2.1, -0.7)\}.
\end{aligned}$$

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

$$\begin{aligned}
\varphi(\mathcal{S}) &= \{\varphi(0,2), \varphi(-2,1), \varphi(2,1), \varphi(-1,-2), \varphi(1,-2)\} \\
&\approx \{(1.4, -1.4), (-0.7, -2.1), (2.1, 0.7), \\
&\quad \downarrow(-2.1, 0.7), (-0.7, 2.1)\}.
\end{aligned}$$

This is shown in Figure 2.49 . The line of reflection for φ has slope $\left(1 - \cos\frac{\pi}{4}\right)/\sin\frac{\pi}{4}$. (You will show this in Exercise 2.54.)

The second questions asks which of the matrices described by Lemma 2.48 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 28(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

  $$\rho = \begin{pmatrix} \cos\frac{2\pi}{3} & -\sin\frac{2\pi}{3} \\ \sin\frac{2\pi}{3} & \cos\frac{2\pi}{3} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

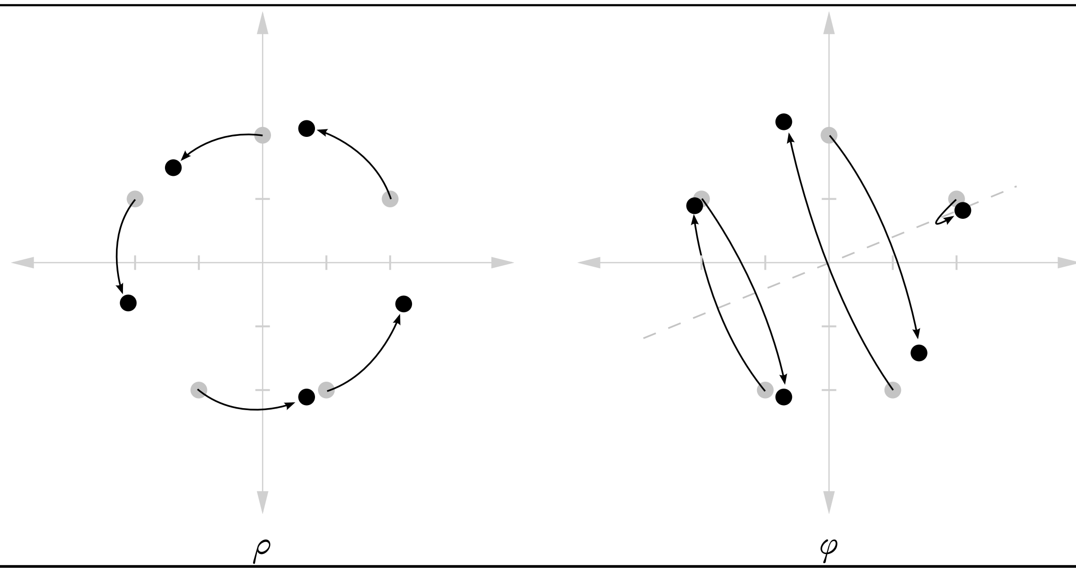


Figure 2.5. Actions of ρ and φ on a pentagon, with $t = \pi/4$

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 28(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 Exercise 2.51.)

We can now give more detailed proofs of Theorems 2.41 and 2.42. We'll prove the first here, and you'll prove the second in the exercises.

Detailed proof of Theorem 2.41. 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

$$\begin{aligned}\rho^2\varphi &= \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.\end{aligned}$$

□

Exercises.

Unless otherwise specified, ρ and φ refer to the elements of D_3.

Exercise 2.50. Show explicitly (by matrix multiplication) that $\rho^3 = \varphi^2 = \iota$.

Exercise 2.51. 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 $\rho^2\varphi$.

Exercise 2.52. Show that any function α satisfying the requirements of Theorem 2.48 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 $\alpha(P + Q) = \alpha(P) + \alpha(Q)$.

Exercise 2.53. 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.)

Exercise 2.54. Fill in each blank of Figure 2.54 with the appropriate justification.

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 2.6. Material for Exercise 2.54

2.3: Cyclic groups and order of elements

Here we re-introduce the familiar notation of exponents, in a manner consistent with what you learned for exponents of real numbers. We use this to describe an important class of groups that recur frequently.

Notation 2.55. Let G be a group, and $g \in G$. If we want to perform the operation on g ten times, we could write

$$g \cdot g \cdot g \cdot g \cdot g \cdot g \cdot g \cdot g \cdot g \cdot g,$$

but this grows tiresome. We could also write

$$\prod_{i=1}^{10} g\ ,$$

which means the same thing. (In mathematics, a capital Π typically means "product.")

That also grows tiresome, especially since a convenient high-school algebra notation means the same thing:

$$g^{10}.$$

We will write things that way from now on. We likewise define g^{-10} to represent

$$\prod_{i=1}^{10} g^{-1} = g^{-1} \cdot g^{-1} \cdot g^{-1} \cdot g^{-1} \cdot g^{-1} \cdot g^{-1} \cdot g^{-1} \cdot g^{-1} \cdot g^{-1} \cdot g^{-1}\ ,$$

which you should understand to mean "the operation on g^{-1} ten times." Indeed, for any $n \in \mathbb{N}^+$ and any $g \in G$ we adopt the following convention:

- g^n means to perform the operation on n copies of g, so $g^n = \prod_{i=1}^{n} g$;
- g^{-n} means to perform the operation on n copies of g^{-1}, so $g^{-n} = \prod_{i=1}^{n} g^{-1} = \left(g^{-1}\right)^n$;
- $g^0 = e$, and if I want to be annoying I can write $g^0 = \prod_{i=1}^{0} g$.

In additive groups we write instead $ng = \sum_{i=1}^{n} g$, $(-n)\,g = \sum_{i=1}^{n} (-g)$, and $0g = 0$.

Notice that this definition assumed n is *positive*. If n is negative, adapt accordingly.

Cyclic groups and generators

Definition 2.56. 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

$$\left\{\ldots, g^{-2}, g^{-1}, e, g^1, g^2, \ldots\right\}.$$

If the group is additive, we would of course write

$$\{\ldots, -2g, -g, 0, g, 2g, \ldots\}.$$

Example 2.57. $\mathbb{Z}$ is cyclic, since any $n \in \mathbb{Z}$ has the form $n \cdot 1$. Thus $\mathbb{Z} = \langle 1 \rangle$. Furthermore, n has the form $(-n) \cdot (-1)$, so $\mathbb{Z} = \langle -1 \rangle$ as well. Both 1 and -1 are generators of $\mathbb{Z}$.

You will show in the exercises that $\mathbb{Q}$ is not cyclic.

In Definition 2.56 we referred to g as *a* generator of G, not as *the* generator. There could in fact be more than one generator, as we see in Example 2.57 from the fact that $\mathbb{Z} = \langle 1 \rangle = \langle -1 \rangle$. Here is another example.

Example 2.58. Let

$$G = \left\{ \begin{array}{c} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \\ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \end{array} \right\} \subsetneq \mathrm{GL}_m(\mathbb{R}).$$

It turns out that G is a cyclic group; both the second and third matrices generate it. For example,

$$\begin{aligned} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^2 &= \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \\ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^3 &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^4 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Every element of G appears in this list of powers of the second element, so the second element generates the group G making it cyclic.

Notice how you prove a group is cyclic: (a) find a potential generator; (b) list all its powers; (c) check that every element of the group appears in the list. This is straightforward enough when the group is finite, but what about a cyclic group? In that case, parts (b) and (c) must be done symbolically, as we did in Example 2.57.

We've elided over an important question arises here. Given a group G and an element $g \in G$, define

$$\langle g \rangle = \left\{ \ldots, g^{-2}, g^{-1}, e, g, g^2, \ldots \right\}.$$

Is $\langle g \rangle$ a group for any $g \in G$? Yes!

Theorem 2.59. For every group G and for every $g \in G$, $\langle g \rangle$ is an abelian group.

To prove Theorem 2.59, we need to make sure we can perform the usual arithmetic on exponents.

Lemma 2.60. Let G be a group, $g \in G$, and $m, n \in \mathbb{Z}$. Each of the following holds:
(A) $g^m g^{-m} = e$; that is, $g^{-m} = (g^m)^{-1}$.
(B) $(g^m)^n = g^{mn}$.
(C) $g^m g^n = g^{m+n}$.

The proof will justify this argument by applying the notation. We have to take care, because the lemma has $m, n \in \mathbb{Z}$, while the notation has $n \in \mathbb{N}^+$. To make this work, we'll consider separate cases for m and n positive or negative. We call this a *case analysis*.

Proof. Each claim follows by case analysis.

(A) If $m = 0$, then $g^{-m} = g^0 = e = e^{-1} = \left(g^0\right)^{-1} = (g^m)^{-1}$.

Otherwise, $m \neq 0$. First assume that $m \in \mathbb{N}^+$. By notation, $g^{-m} = \prod_{i=1}^{m} g^{-1}$. Hence

$$\begin{aligned}
g^m g^{-m} &= \left(\prod_{i=1}^{m} g\right)\left(\prod_{i=1}^{m} g^{-1}\right) && \text{(by definition)}\\
&= \left(\prod_{i=1}^{m-1} g\right)\left(g \cdot g^{-1}\right)\left(\prod_{i=1}^{m-1} g^{-1}\right) && \text{(associative property)}\\
&= \left(\prod_{i=1}^{m-1} g\right) e \left(\prod_{i=1}^{m-1} g^{-1}\right) && \text{(inverse property)}\\
&= \left(\prod_{i=1}^{m-1} g\right)\left(\prod_{i=1}^{m-1} g^{-1}\right) && \text{(identity property)}\\
&\vdots\\
&= e.
\end{aligned}$$

Since the inverse of an element is unique, $g^{-m} = (g^m)^{-1}$.
Now assume that $m \in \mathbb{Z}\backslash\mathbb{N}$. Since m is negative, we cannot express the product using m; the notation discussed on page 93 requires a *positive* exponent. On the other hand, the opposite of a negative number is positive, so $-m \in \mathbb{N}^+$. Let's write $\widehat{m} = -m$. Since $\widehat{m}$ is positive, we can apply the notation directly; $g^{-m} = g^{\widehat{m}} = \prod_{i=1}^{\widehat{m}} g$, while $g^m = g^{-\widehat{m}} = \prod_{i=1}^{\widehat{m}} g^{-1}$. (To see this in a more concrete example, try it with an actual number. If $m = -5$, then $\widehat{m} = -(-5) = 5$, so $g^m = g^{-5} = g^{-\widehat{m}}$ and $g^{-m} = g^5 = g^{\widehat{m}}$.) As above, we have

$$g^m g^{-m} \underset{\text{subs.}}{=} g^{-\widehat{m}} g^{\widehat{m}} \underset{\text{not.}}{=} \left(\prod_{i=1}^{\widehat{m}} g^{-1}\right)\left(\prod_{i=1}^{\widehat{m}} g\right) = e.$$

Hence $g^{-m} = (g^m)^{-1}$.

(B) If $n = 0$, then $(g^m)^n = (g^m)^0 = e$ because *anything* to the zero power is e. On the other hand, $g^{mn} = g^0 = e$, so in this case $(g^m)^n = 0 = g^{mn}$, as desired.
Now we have to show that $(g^m)^n = g^{mn}$ when $n \neq 0$. Assume first that $n \in \mathbb{N}^+$. By notation, $(g^m)^n = \prod_{i=1}^{n} g^m$. We split this into two further subcases.

(B1) If $m \in \mathbb{N}$, we have

$$(g^m)^n \underset{\text{not.}}{=} \prod_{i=1}^{n}\left(\prod_{i=1}^{m} g\right) \underset{\text{ass.}}{=} \prod_{i=1}^{mn} g \underset{\text{not.}}{=} g^{mn}.$$

(B2) Otherwise, repeat what we did above by setting $\widehat{m} = -m \in \mathbb{N}^+$. We have

$$\begin{aligned}
(g^m)^n &\underset{\text{subs.}}{=} \left(g^{-\widehat{m}}\right)^n \underset{\text{not.}}{=} \prod_{i=1}^{n}\left(\prod_{i=1}^{\widehat{m}} g^{-1}\right)\\
&\underset{\text{ass.}}{=} \prod_{i=1}^{\widehat{m}n} g^{-1} \underset{\text{not.}}{=} \left(g^{-1}\right)^{\widehat{m}n}\\
&\underset{\text{not.}}{=} g^{-\widehat{m}n} \underset{\text{assoc.}}{=} g^{(-\widehat{m})n} \underset{\text{subs.}}{=} g^{mn}.
\end{aligned}$$

What if n is negative? Let $\widehat{n} = -n$ and repeat as above.

(C) We consider three cases.
If $m = 0$, then $g^m = g^0 = e$, so

$$g^{m+n} \underset{\text{add. id.}}{=} g^n \underset{\text{gp. ident.}}{=} e \cdot g^n \underset{\text{subst.}}{=} g^m g^n .$$

The case $n = 0$ is proved similarly. It remains to show what happens when $m, n \neq 0$.
If m, n have the same sign (that is, $m, n \in \mathbb{N}^+$ or $m, n \in \mathbb{Z} \setminus \mathbb{N}$), then if they are both positive the argument is straightforward substitution:

$$g^{m+n} = \Big(\underbrace{g \cdot g \cdots g}_{m+n \text{ copies}}\Big) \underset{\text{assoc.}}{=} \Big(\underbrace{g \cdot g \cdots g}_{m \text{ copies}}\Big)\Big(\underbrace{g \cdot g \cdots g}_{n \text{ copies}}\Big) = g^m g^n .$$

If both are negative, however, $m + n$ is also negative, so that

$$\begin{aligned} g^{m+n} &= \Big(\underbrace{g^{-1} \cdot g^{-1} \cdots g^{-1}}_{|m+n| \text{ copies}}\Big) \underset{\text{assoc.}}{=} \Big(\underbrace{g^{-1} \cdot g^{-1} \cdots g^{-1}}_{|m| \text{ copies}}\Big)\Big(\underbrace{g^{-1} \cdot g^{-1} \cdots g^{-1}}_{|n| \text{ copies}}\Big) \\ &= \left(g^{-1}\right)^{-m} \left(g^{-1}\right)^{-n} = g^m g^n . \end{aligned}$$

Now consider the case where m and n have different signs. In the first case, suppose m is negative and $n \in \mathbb{N}^+$. As in (A), let $\widehat{m} = -m \in \mathbb{N}^+$; then

$$g^m g^n = \left(g^{-1}\right)^{-m} g^n = \left(\prod_{i=1}^{\widehat{m}} g^{-1}\right)\left(\prod_{i=1}^{n} g\right).$$

If $\widehat{m} \geq n$, we have more copies of g^{-1} than g, so after cancellation,

$$g^m g^n = \prod_{i=1}^{\widehat{m}-n} g^{-1} = g^{-(\widehat{m}-n)} = g^{m+n}.$$

Otherwise, $\widehat{m} < n$, and we have more copies of g than of g^{-1}. After cancellation,

$$g^m g^n = \prod_{i=1}^{n-\widehat{m}} g = g^{n-\widehat{m}} = g^{n+m} = g^{m+n}.$$

The remaining case ($m \in \mathbb{N}^+$, $n \in \mathbb{Z} \setminus \mathbb{N}$) is similar, and you will prove it for homework.

□

These properties of exponent arithmetic allow us to show that $\langle g \rangle$ is a group.

Proof of Theorem 2.59. 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 2.60 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, $e = g^0 \in \langle g\rangle$.
- 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,

$$xy = g^a g^b = g^{a+b} = g^{b+a} = g^b g^a = yx.$$

□

The order of an element

Given an element and an operation, Theorem 2.59 links them to a group. It makes sense, therefore, to link an element to the order of the group that it generates.

Definition 2.61. Let G be a group, and $g \in G$. We say that the **order** of g is $\operatorname{ord}(g) = |\langle g\rangle|$. If $\operatorname{ord}(g) = \infty$, we say that g has **infinite order**.

(Another way of saying this is that the order of an element is the number of distinct elements in its cyclic group.)

If the order of a group is finite, then we can write an element in different ways.

Example 2.62. Recall Example 2.58; we can write

$$\begin{aligned}\begin{pmatrix}1&0\\0&1\end{pmatrix} = \begin{pmatrix}0&-1\\1&0\end{pmatrix}^0 &= \begin{pmatrix}0&-1\\1&0\end{pmatrix}^4\\ &= \begin{pmatrix}0&-1\\1&0\end{pmatrix}^8 = \cdots.\end{aligned}$$

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 $0 \le r < 4$. 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:

$$\begin{aligned}
\begin{pmatrix}0&-1\\1&0\end{pmatrix}^{4q} &= \begin{pmatrix}1&0\\0&1\end{pmatrix},\\
\begin{pmatrix}0&-1\\1&0\end{pmatrix}^{4q+1} &= \begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}0&-1\\1&0\end{pmatrix} = \begin{pmatrix}0&-1\\1&0\end{pmatrix},\\
\begin{pmatrix}0&-1\\1&0\end{pmatrix}^{4q+2} &= \begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}-1&0\\0&-1\end{pmatrix} = \begin{pmatrix}-1&0\\0&-1\end{pmatrix},\\
\begin{pmatrix}0&-1\\1&0\end{pmatrix}^{4q+3} &= \begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}0&1\\-1&0\end{pmatrix} = \begin{pmatrix}0&1\\-1&0\end{pmatrix}.
\end{aligned}$$

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

$$\langle g \rangle = \{I_2, g, g^2, g^3\}.$$

The example suggests that if the order of an element G is $n \in \mathbb{N}$, then we can write

$$\langle g \rangle = \{e, g, g^2, \ldots, g^{n-1}\}.$$

This explains why we call $\langle g \rangle$ a *cyclic* group: once they reach $\text{ord}(g)$, the powers of g "cycle".To prove this in general, we have to show that for a generic cyclic group $\langle g \rangle$ with $\text{ord}(g) = n$,

- n is the smallest positive power that gives us the identity; that is, $g^n = e$, and
- for any two integers between 0 and n, the powers of g are different; that is, if $0 \le a < b < n$, then $g^a \ne g^b$.

Theorem 2.63 accomplishes that, and a bit more as well.

Theorem 2.63. Let G be a group, $g \in G$, and $\text{ord}(g) = n$. Then
(A) $e, g, g^2, \ldots, g^{n-1}$ are all distinct.
In addition, if $n < \infty$, each of the following holds:
(B) $g^n = e$;
(C) n is the smallest positive integer d such that $g^d = e$; 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 $0 \le a < b < n$ and $g^a = g^b$; then $e = (g^a)^{-1} g^b$. By Lemma 2.60, we can write

$$e = g^{-a} g^b = g^{-a+b} = g^{b-a}.$$

Let $d = b - a$, and notice that $d < n$. 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 $0 \le r < d$. Applying Lemma 2.60 again, we have

$$g^m = g^{qd+r} = \left(g^d\right)^q g^r = e^q g^r = g^r,$$

so *any* power of g can be written as a remainder after division by d. In other words,

$$\langle g \rangle = \{e, g, g^2, \ldots, g^{d-1}\}.$$

This implies that $|\langle g \rangle| = d$, which contradicts the assumption that $n = \text{ord}(g) = |\langle d \rangle|$.

For the remainder of the proof, we assume that $n < \infty$.

(B) We know that $\text{ord}(g) = n$, 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 2.60, we find that $g^{n-d} = e$. Recall that $0 \leq d < n$, so $0 < n - d \leq n$. Since $g^a \neq e$ for $a = 1, 2, \ldots, n-1$, we must have $n - d = n$, so $d = 0$. By substitution, $g^n = g^d = g^0 = e$.

(C) In (B), S is the set of all positive integers m such that $g^m = e$; we let the smallest element be d, and thus $d \leq n$. On the other hand, (A) tells us that we cannot have $d < n$; otherwise, $g^d = g^0 = e$. Hence, $n \leq d$. We already had $d \leq n$, 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 2.60 and some arithmetic tell us that

$$\begin{aligned} g^b &= g^b \cdot e = g^b \cdot e^q \\ &= g^b \cdot (g^n)^q = g^b \cdot g^{nq} \\ &= g^b \cdot g^{a-b} = g^{b+(a-b)} = g^a. \end{aligned}$$

Conversely, if we assume that $g^b = g^a$, then Lemma 2.60 implies that $g^{b-a} = e$. Use the Division Theorem to choose $q, r \in \mathbb{Z}$ such that $b - a = nq + r$ and $0 \leq r < n$. By substitution and Lemma 2.60,

$$e = g^{b-a} = g^{qn+r} = (g^n)^q g^r = e^q g^r = g^r.$$

Recall that $0 \leq r < n$. 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 $g^n = e$ and we "cycle around."

Exercises.

Exercise 2.64. Recall from Example 2.58 the matrix

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Express A as a power of the other non-identity matrices of the group.

Exercise 2.65. In each of the following, compute the cyclic group, and then the order, of the element a of the group G.

(a) $a = \varphi$, $G = D_3$
(b) $a = \rho^2$, $G = D_3$
(c) $a = \rho\varphi$, $G = D_3$
(d) $a = \mathbf{i}$, $G = Q_8$
(e) $a = -\mathbf{1}$, $G = Q_8$

Let G be a group, and $g \in G$. Let $d, n \in \mathbb{Z}$ and assume $\operatorname{ord}(g) = d$.
Claim: $g^n = e$ if and only if $d \mid n$.
Proof:

1. Assume that $g^n = e$.
   (a) By _____, there exist $q, r \in \mathbb{Z}$ such that $n = qd + r$ and $0 \le r < d$.
   (b) By _____, $g^{qd+r} = e$.
   (c) By _____, $g^{qd} g^r = e$.
   (d) By _____, $\left(g^d\right)^q g^r = e$.
   (e) By _____, $e^q g^r = e$.
   (f) By _____, $eg^r = e$. By the identity property, $g^r = e$.
   (g) By _____, d is the *smallest* positive integer such that $g^d = e$.
   (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 2.60, the right hand side of that equation can be rewritten as _____.
   (d) Recall that $\operatorname{ord}(g) = d$. By Theorem 2.63, $g^d = e$, 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, $g^n = e$.
3. We showed first that if $g^n = e$, then $d|n$; we then showed that _____. This proves the claim.

Figure 2.7. Material for Exercise 2.67

(f) $a = \mathbf{1}, G = Q_8$

Exercise 2.66. Complete the proof of Lemma 2.60(C).

Exercise 2.67. Fill in each blank of Figure 2.67 with the justification or statement.

Exercise 2.68. Show that any group of 3 elements is cyclic.

Exercise 2.69. Is the Klein 4-group (Exercise 2.19 on page 81) cyclic? What about the cyclic group of order 4?

Exercise 2.70. Show that $\mathbb{Q}$ is not cyclic.

Exercise 2.71. Use a fact from linear algebra to explain why $\mathrm{GL}_m(\mathbb{R})$ is not cyclic.

2.4: The roots of unity

Recall from Section 0.4 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 $f(x) = x^4 - 1$, which had four roots, ± 1 and $\pm i$, where $i^2 = -1$.

Now replace the 4 by an n, so that we are considering $x^n - 1$. The roots of these polynomials are both intruiguing and important! We conclude this chapter with some attention to why they are intriguing, while the exercises contain hints as to why they are important.

The algebraic structure of $\mathbb{C}$

Recall the complex numbers from Section 0.4. They have some of the structures we have been studying in the past two chapters.

> **Theorem 2.72.** $\mathbb{C}$ is a monoid under multiplication, and an abelian group under addition. In addition, multiplication **distributes** over addition; that is, for any $x, y, z \in \mathbb{C}$ we have $x(y+z) = xy + xz$.

This structure of multiplication and addition pops up a lot; we actually built $\mathbb{C}$ to make sure it *had* this structure. This is the structure of a **ring** — in fact, it is a special kind of ring, called a **field**. We discuss the ring structure in some detail in Chapter 7. But for now we content ourselves with monoid under multiplication and abelian group under addition.

Proof. Let $x, y, z \in \mathbb{C}$. Write $x = a + bi$, $y = c + di$, and $z = e + fi$, for some $a, b, c, d, e, f \in \mathbb{R}$. Let's look at multiplication first.

closure? We built $\mathbb{C}$ to be closed under multiplication, so the discussion above suffices.

associative? We need to show that

$$(xy)z = x(yz). \tag{11}$$

Expanding the product on the left, we have

$$(xy)z = [(a+bi)(c+di)](e+fi) = [(ac-bd)+(ad+bc)i](e+fi) .$$

Expand again, and we get

$$(xy)z = [(ac-bd)e-(ad+bc)f]+[(ac-bd)f+(ad+bc)e]i .$$

Now let's look at the product on the right of equation (11). Expanding it, we have

$$x(yz) = (a+bi)[(c+di)(e+fi)] = (a+bi)[(ce-df)+(cf+de)i] .$$

Expand again, and we get

$$x(yz) = [a(ce-df)-b(cf+de)]+[a(cf+de)+b(ce-df)]i .$$

If you look carefully, you will see that both expansions resulted in the same complex number:

$$(ace-bde-adf-bcf)+(acf-bdf+ade+bce)i .$$

Thus, multiplication in $\mathbb{C}$ is associative.

identity? We claim that $1 \in \mathbb{R}$ is the multiplicative identity even for $\mathbb{C}$. Recall that we can write $1 = 1 + 0i$. Then,

$$1x = (1+0i)(a+bi) = (1a-0b) + (1b+0a)i = a+bi = x.$$

Since x was arbitrary in $\mathbb{C}$, it must be that 1 is, in fact, the identity.

distributive? We need to show that

$$x(y+z) = xy + xz\ . \tag{12}$$

Simplifying the product on the left, we have

$$x(y+z) = (a+bi)[(c+di)+(e+fi)] = (a+bi)[(c+e)+(d+f)i]\ .$$

Expand that product to get

$$x(y+z) = [a(c+e) - b(d+f)] + [a(d+f) + b(c+e)]i\ .$$

Now let's look at the product on the right of equation 12. Expanding it, we have

$$\begin{aligned} xy + xz &= (a+bi)(c+di) + (a+bi)(e+fi) \\ &= [(ac-bd) + i(ad+bc)] + [(ae-bf) + i(af+be)] \\ &= [(ac-bd) + (ae-bf)] + [(ad+bc) + (af+be)]i\ . \end{aligned}$$

If you look carefully, you will see that both expansions resulted in the same complex number:

$$(ac+ae-bd-bf) + i(ad+ad+af+bc+be)\ .$$

Thus, multiplication in $\mathbb{C}$ is distributive.

We have shown that $\mathbb{C}$ is a monoid under multiplication, and that multiplication distributes over addition. What about addition; is it an abelian group? We leave that to the exercises. □

There are a *lot* of wonderful properties of $\mathbb{C}$ that we could discuss. For example, you can see that the roots of x^2+1 lie in $\mathbb{C}$, but what of the roots of x^2+2? They're in there, too. In fact, *every* polynomial of degree n with real coefficients has n roots in $\mathbb{C}$! We can't prove that without a lot more theory, though.

Roots of unity

Any root of the polynomial $f(x) = x^n - 1$ is called a **root of unity**. We will see in Chapter 9 that these are very important in the study of polynomial roots, in part because of their elegant form.

Back on p. 41 we showed how any nth root of i is a complex number. This required some trigonometry, which you should go back and review if you forgot it, because we're going to use it again here with roots of unity.

Theorem 2.73. 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 2.73, we need a different property of ω. We could prove it as part of the proof of Theorem 2.73, but it's quite useful on its own, so we separate it as Lemma 2.74.

Lemma 2.74 (Powers of ω)**.** If ω is defined as in Theorem 2.73, 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 Lemma 2.60 and the inductive hypothesis to rewrite ω^{m+1} as,

$$\omega^{m+1} \underset{\text{Lem. 2.60}}{=} \omega^m \cdot \omega \underset{\substack{\text{ind.}\\\text{hyp.}}}{=} \left[\cos\left(\frac{2\pi m}{n}\right) + i \sin\left(\frac{2\pi}{n}\right)\right] \cdot \left[\cos\left(\frac{2\pi}{n}\right) + i \sin\left(\frac{2\pi}{n}\right)\right].$$

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

$$\begin{aligned}\omega^{m+1} = &\left[\cos\left(\frac{2\pi m}{n}\right)\cos\left(\frac{2\pi}{n}\right) - \sin\left(\frac{2\pi m}{n}\right)\sin\left(\frac{2\pi}{n}\right)\right] \\ &+ i\left[\sin\left(\frac{2\pi}{n}\right)\cos\left(\frac{2\pi m}{n}\right) + \sin\left(\frac{2\pi m}{n}\right)\cos\left(\frac{2\pi}{n}\right)\right].\end{aligned}$$

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 2.74, proving Theorem 2.73 is quite easy.

Proof of Theorem 2.73. Substitution and Lemma 2.74 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 $x^n - 1$. □

The phenomenon described by Theorem 2.73 means that the n roots of unity form a cyclic group!

Theorem 2.75. The nth roots of unity are $\Omega_n = \{1, \omega, \omega^2, \ldots, \omega^{n-1}\}$, where ω is defined as in Theorem 2.73. They form a cyclic 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 α_1, α_2, …, α_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 true for all polynomials — but it really is here, honest! We discuss this in Chapter 10, specifically Section 10.2; the idea is called *unique factorization.*
- Second, f will factor as

  $$f(x) = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n)\ .$$

  By the Factor Theorem (Theorem 7.45), each root α_i of f corresponds to a linear factor $x - \alpha_i$. At the very least, then, f must factor as

  $$f(x) = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n)\cdot g(x)\ ,$$

  where g is yet to be determined. Each linear factor has degree one, so $\deg f = n + \deg g$, but $\deg f = n$, which forces $\deg g = 0$. We see that f can have no more roots, and the *only* roots of f are α_1, …, α_n.

(You can see this already in $x^4 - 1$ above. You should have seen the Factor 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 $\mathbb{N}^+$:

$$(\omega^m)^n - 1 = \omega^{mn} - 1 = \omega^{nm} - 1 = (\omega^n)^m - 1 = 1^m - 1 = 0.$$

Hence ω^m is a root of unity for any $m \in \mathbb{N}^+$. However, most of these overlap, just as $(-1)^2 = (-1)^4 = (-1)^6 = \cdots$. If $\omega^m = \omega^\ell$, then

$$\cos\left(\frac{2\pi m}{n}\right) = \cos\left(\frac{2\pi \ell}{n}\right) \quad \text{and} \quad \sin\left(\frac{2\pi m}{n}\right) = \sin\left(\frac{2\pi \ell}{n}\right),$$

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 Exercise 2.85.) 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.

closed? Let $x, y \in \Omega_n$; you will show in Exercise 2.82 that $xy \in \Omega_n$.

associative? 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 all $n \in \mathbb{N}^+$.

inverses? Let $x \in \Omega_n$; you will show in Exercise 2.83 that $x^{-1} \in \Omega_n$.

cyclic? Theorem 2.73 tells us that $\omega \in \Omega_n$; the remaining elements are powers of ω. Hence $\Omega_n = \langle \omega \rangle$.

□

Combined with the explanation we gave earlier of the complex plane, Theorem 2.75 gives us a wonderful symmetry for the roots of unity.

Example 2.76. 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 2.74,

$$\omega^m = \cos\left(\frac{2\pi m}{7}\right) + i \sin\left(\frac{2\pi m}{7}\right),$$

where $m = 0, 1, \ldots, 6$. By substitution, the angles we are looking at are

$$0, \frac{2\pi}{7}, \frac{4\pi}{7}, \frac{6\pi}{7}, \frac{8\pi}{7}, \frac{10\pi}{7}, \frac{12\pi}{7}.$$

Recall that in the complex plane, any complex number $a + bi$ corresponds to the point (a, b) on $\mathbb{R}^2$. The Pythagorean identity $\cos^2\alpha + \sin^2\alpha = 1$ tells us that the coordinates of the roots of

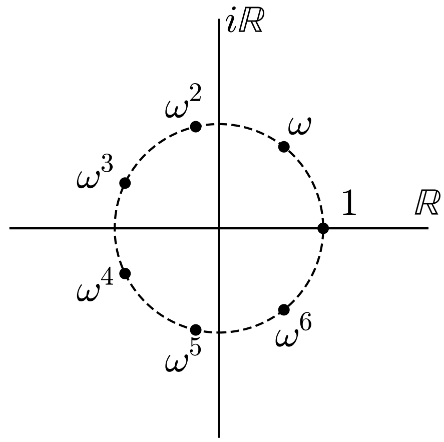


Figure 2.8. The seventh roots of unity, on the complex plane

unity lie on the unit circle. Since the angles are at equal intervals, they divide the unit circle into seven equal arcs! See Figure 2.8.

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 generator of Ω_n? 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 2.77. We call any generator of Ω_n a **primitive *n*th root of unity**.

Example 2.78. Consider $\Omega_{10} = \left\{1, \omega, \omega^2, \ldots, \omega^9\right\}$. You know that $\omega^{10} = 1$. Besides ω, what is another primitive root of unity?

If we look at powers of ω^5, we'd find that $\left(\omega^5\right)^2 = \omega^{10} = 1$ and $\left(\omega^5\right)^3 = \omega^5$. The powers have begun to cycle, so $\left\langle \omega^5 \right\rangle = \left\{1, \omega^5\right\}$. By definition, ω^5 does *not* generate Ω_{10}, and is not a primitive root of unity.

If we look at powers of ω^7, we'd find that

$$\begin{array}{ll}
\left(\omega^7\right)^1 = \omega^7 & \left(\omega^7\right)^2 = \omega^{14} = \omega^4 \\
\left(\omega^7\right)^3 = \omega^{21} = \omega & \left(\omega^7\right)^4 = \omega^{28} = \omega^8 \\
\left(\omega^7\right)^5 = \omega^{35} = \omega^5 & \left(\omega^7\right)^6 = \omega^{42} = \omega^2 \\
\left(\omega^7\right)^7 = \omega^{49} = \omega^9 & \left(\omega^7\right)^8 = \omega^{56} = \omega^6 \\
\left(\omega^7\right)^9 = \omega^{63} = \omega^3 & \left(\omega^7\right)^{10} = \omega^{70} = 1\,.
\end{array}$$

With the 10th power of ω^7 we have generated ten distinct elements of Ω_{10}, which is all of them! So ω^7 generates Ω_{10}.

Exercises.

Unless stated otherwise, $n \in \mathbb{N}^+$ and ω is a primitive nth root of unity.

Exercise 2.79. Show that $\mathbb{C}$ is a group under addition.

Exercise 2.80.
(a) Find all the primitive square roots of unity, all the primitive cube roots of unity, and all the primitive quartic (fourth) roots of unity.
(b) Sketch *all* the square roots of unity on a complex plane. (Not just the primitive ones, but all.) Repeat for the cube and quartic roots of unity, each on a separate plane.
(c) Are any cube roots of unity *not* primitive? what about quartic roots of unity?

Exercise 2.81. Recall Ω_{10} from Example 2.78.
(a) Determine which values of k make the element ω^k of Ω_{10} a primitive 10th root of unity.
(b) What pattern do you notice in the answers to part (a)?

Exercise 2.82. Let $n \in \mathbb{N}^+$, and suppose that a and b are both positive powers of ω. Show that $ab \in \Omega_n$.

Exercise 2.83.
(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.

Exercise 2.84. 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 2.75 to explain how we can use β and Ω_n to list all n roots of $x^n - b$.

Exercise 2.85.
(a) For each $\omega \in \Omega_6$, find $x, y \in \mathbb{R}$ such that $\omega = x + yi$. Plot all the points (x, y) on a graph.
(b) Do you notice any pattern to the points? If not, repeat part (a) for Ω_7, Ω_8, etc., until you see the pattern.

Exercise 2.86.
(a) Show that $\mathbb{C}$ and $\mathbb{R}^2$ are isomorphic as monoids under addition, where we view $\mathbb{R}^2$ as a direct product of $\mathbb{R}$.
(b) Explain why $\mathbb{C}$ and $\mathbb{R}^2$ are *not* isomorphic as monoids under *multiplication*, where again we view $\mathbb{R}^2$ as a direct product of $\mathbb{R}$.

Exercise 2.87. Recall from Exercise 0.94 the set of quaternions $\{\pm\mathbf{1}, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\}$, 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) Use the properties of these matrices that you proved in Exercise 0.94 to build the Cayley table of Q_8. (In this case, the Cayley table is the multiplication table.)
(c) Show that Q_8 is a group under matrix multiplication.
(d) Explain why Q_8 is not an abelian group.

Exercise 2.88. List the elements of $\langle -\mathbf{1} \rangle$ and $\langle \mathbf{j} \rangle$ in Q_8.

Exercise 2.89. In Exercise 2.87 you showed that the quaternions form a group under matrix multiplication. Verify that $H = \{\mathbf{1}, -\mathbf{1}, \mathbf{i}, -\mathbf{i}\}$ is a cyclic group. What elements generate H?

Exercise 2.90. Show that Q_8 is not cyclic.

Chapter 3: Subgroups

A subset of a group is not necessarily a group; for example, $\{2,4\} \subset \mathbb{Z}$, but $\{2,4\}$ doesn't satisfy at least three properties of an additive group unless we change the definition of addition. (Take a moment to answer this question: which three properties does $\{2,4\}$ fail to satisfy?)

Of course, some subsets of groups *are* groups:

Definition 3.1. 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.

One of the keys 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 3.1). We then explore how subgroups create *cosets,* equivalence classes within a group that perform a role similar to division of integers (Section 3.2). It turns out that in finite groups, we can count the number of these equivalence classes quite easily (Section 3.3).

Cosets open the door to a special class of groups called *quotient groups*, (Section 3.4), one of which is a very natural, very useful tool (Section 3.5) that will allow us to devise some "easy" solutions for problems in Number Theory (Chapter 6).

3.1: Subgroups

Notation 3.2. If H is a subgroup of G, then we write $H < G$.

A group always has at least one subgroup, and often two.

Lemma 3.3. Let $(G, \circ)$ be any group.

- $G < G$; that is, any group is a subgroup of itself.
- $\{e\} < G$, called the **trivial subgroup**.

Proof. By hypothesis, G is a group, and certainly $G \subseteq G$, so by definition $G < G$. As for the trivial subgroup,

- it is closed because $e \circ e = e \in G$;
- it inherits the associative property from G;
- it clearly has an identity; and
- e is its own inverse ($e \circ e = e$ so $e^{-1} = e$).

□

We are often interested in other subgroups.

Definition 3.4. If $\{e\} \subsetneq H \subsetneq G$, then H is a **proper subgroup** of G. If we are feeling naughty, we refer to G and $\{e\}$ as **improper subgroups**, but this is highly nonstandard (not even Wikipedia's page on subgroups sinks to such depths) and should not be used in conversations with mathematicians who suffer from an inflated sense of convention.

Example 3.5. Check that the following statements are true by verifying that the properties of a group are satisfied: that is, the accused subgroup is in fact a subset of the named group, and is a group under the same operation.
(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) $\langle \mathbf{i} \rangle$ is a subgroup of Q_8.
Which of them are proper subgroups?

Checking all four properties of a group is cumbersome. In that case, which properties *must* we check to decide whether a subset is a subgroup?

We can eliminate the associative and abelian properties from consideration. In fact, the operation remains associative and commutative for any sub*set*.

Lemma 3.6. 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 3.6 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 $H = \emptyset$, then the lemma is true vacuously.[17]

Otherwise, $H \neq \emptyset$. Let $a, b, c \in H$. Since $H \subseteq G$, by inclusion we have $a, b, c \in G$. Since the operation is associative in G, $a(bc) = (ab)c$. Likewise, if G is abelian, then $ab = ba$. □

Lemma 3.6 has reduced the number of requirements for a subgroup from four to three. Amazingly, we can simplify this further, to *only one criterion*.

Theorem 3.7 (The Subgroup Theorem)**.** Let $H \subseteq G$ be nonempty. The following are equivalent:
(A) $H < G$;
(B) for every $x, y \in H$, we have $xy^{-1} \in H$.

Notation 3.8. If G were an additive group, we would write $x - y$ instead of xy^{-1}.

Proof. By Exercise 2.20 on page 81, (A) implies (B).

Conversely, assume (B). By Lemma 3.6, we need to show only that H satisfies the closure, identity, and inverse properties. We do this slightly out of order:

identity? Let $x \in H$. By (B), $e = x \cdot x^{-1} \in H$.[18]
inverse? Let $x \in H$. Since H satisfies the identity property, $e \in H$. By (B), $x^{-1} = e \cdot x^{-1} \in H$.

[17] A statement is true "vacuously" whenever you're dealing with "elements" of the empty set. The idea is that the empty set has no elements, so you can't find any *counterexamples* to the statement, so the statement is true. Here's another example: for every dog in the set $\emptyset$, the dog meows.
[18] Notice that here we are replacing the y in (B) with x. This is fine, since nothing in (B) requires x and y to be distinct.

closure? Let $x,y \in H$. Since H satisfies the inverse property, $y^{-1} \in H$. By (B), $xy = x \cdot \left(y^{-1}\right)^{-1} \in H$.

Since H satisfies the closure, identity, and inverse properties, $H < G$. □

Let's take a look at the Subgroup Theorem in action.

Example 3.9. Let $d \in \mathbb{Z}$. We claim that $d\mathbb{Z} < \mathbb{Z}$. (Here $d\mathbb{Z}$ is the set defined in Example 3.5.) *Why?* Let's use the Subgroup Theorem.

Let $x,y \in d\mathbb{Z}$. If we can show that $x-y \in d\mathbb{Z}$, 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. That's what we want, so let's try to show that $x-y \in d\mathbb{Z}$; that is, $x-y$ is an integer multiple of d.

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

$$\begin{aligned} x-y &= x+(-y) = dm + d(-n) \\ &= d(m+(-n)) = d(m-n). \end{aligned}$$

Recall that $\mathbb{Z}$ is an abelian group, so $m-n \in \mathbb{Z}$, 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, $d\mathbb{Z} < \mathbb{Z}$.

The following geometric example gives a visual image of what a subgroup "looks" like.

Example 3.10. 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 intuitive 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

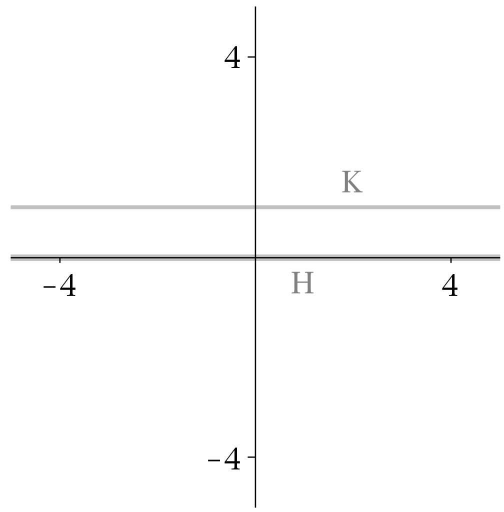


Figure 3.1. H and K from Example 3.10

groups. In this case, the identity is $(0,0)$, which is *not* in K. It isn't always that easy to tell when a subset is not a subgroup, though; the Subgroup Theorem usually makes it much easier.

Figure 3.10 diagrams H and K. You will diagram another subgroup of G in Exercise 3.18.

Examples 3.9 and 3.10 give us examples of how the Subgroup Theorem verifies subgroups of *abelian* groups. Two interesting examples of nonabelian subgroups appear in D_3. We don't even need the Subgroup Theorem to demonstrate this!

Example 3.11. Recall D_3 from Section 2.2. Both $H = \{\iota, \varphi\}$ and $K = \{\iota, \rho, \rho^2\}$ are subgroups of D_3. *Why?* Certainly $H, K \subsetneq G$, and Theorem 2.59 on page 95 tells us that H and K are groups, since $H = \langle \varphi \rangle$, and $K = \langle \rho \rangle$.

If a group satisfies a given property, a natural question to ask is whether its subgroups also satisfy this property. Cyclic groups are a good example: is every subgroup of a cyclic group also cyclic? The answer relies on the Division Theorem (Theorem 0.41 on page 17).

Theorem 3.12. 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. For the trivial subgroup $H = \{e\}$ we have $H = \langle e \rangle = \langle g^0 \rangle$, and we are done.

Assume, therefore, that H is nontrivial. We claim we can find an element $h \in H$ with the form $h = g^k$ where $k \in \mathbb{N}^+$. Since H is nontrivial, we can choose $x \in H \backslash \{e\}$. By inclusion, x is also an element of G, and g generates G, so g generates x, so $x = g^k$ for some $k \in \mathbb{Z}$. We chose $x \neq e$, so $k \neq 0$. If $k > 0$, then, let $h = x$ are we are done: $h = g^k$ with $k \in \mathbb{N}^+$. That leaves the case where $k < 0$, but closure of H implies that $x^{-1} = g^{-k} \in H$, and $-k \in \mathbb{N}^+$, so let $h = x^{-1}$ and we have $h = g^{-k}$ in the correct form.

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, $S = \left\{i \in \mathbb{N}^+ : g^i \in H\right\}$. From the well-ordering of $\mathbb{N}$, there exists a smallest element of S; 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
- $0 \leq r < d$.

Let $y = g^r$; by Lemma 2.60, 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 $0 \leq r < d$. If $0 < r$, then $r \in S$; after all, S is the set of all positive powers of g. But we have $r < d$, so $0 < r$ would contradicts the choice of d as the smallest element of S. The combination $0 \leq r$ and $0 \not< r$ implies $r = 0$ 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.

Example 3.13. Recall from Example 2.57 on page 94 that $\mathbb{Z}$ is cyclic; in fact $\mathbb{Z} = \langle 1 \rangle$. By Theorem 3.12, $d\mathbb{Z}$ is cyclic. In fact, $d\mathbb{Z} = \langle d \rangle$. Can you find another generator of $d\mathbb{Z}$?

Exercises.

Exercise 3.14. Recall that Ω_n, the nth roots of unity, is the cyclic group $\langle \omega \rangle$.
(a) Compute Ω_2 and Ω_4, and explain why $\Omega_2 < \Omega_4$.
(b) Compute Ω_8, and explain why both $\Omega_2 < \Omega_8$ and $\Omega_4 < \Omega_8$.
(c) Explain why, if $d \mid n$, then $\Omega_d < \Omega_n$.

Exercise 3.15. Show that even though the Klein 4-group is not cyclic, each of its proper subgroups is cyclic (see Exercises 2.19 on page 81 and 2.69 on page 101).

Exercise 3.16.
(a) Fill in each blank of Figure 3.2 with the appropriate justification or expression.
(b) Why would someone take this approach, rather than use the definition of a subgroup?

Exercise 3.17.

(a) Let $D_n(\mathbb{R}) = \{aI_n : a \in \mathbb{R}\} \subseteq \mathbb{R}^{n\times n}$; that is, $D_n(\mathbb{R})$ is the set of all diagonal matrices whose values along the diagonal is constant. Show that $D_n(\mathbb{R}) < \mathbb{R}^{n\times n}$. (In case you've forgotten Exercise 2.33, the operation here is addition.)
(b) Let $D_n^*(\mathbb{R}) = \{aI_n : a \in \mathbb{R}\backslash\{0\}\} \subseteq \mathrm{GL}_n(\mathbb{R})$; that is, $D_n^*(\mathbb{R})$ is the set of all non-zero diagonal matrices whose values along the diagonal is constant. Show that $D_n^*(\mathbb{R}) < \mathrm{GL}_n(\mathbb{R})$. (In case you've forgotten Definition 2.7, the operation here is multiplication.)

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 3.2. Material for Exercise 3.16

Exercise 3.18. Let $G = \mathbb{R}^2$, with addition defined as in Example 3.10. 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.

Exercise 3.19. Let G be an abelian group. Let H, K be subgroups of G. Let

$$H + K = \{x + y : x \in H, y \in K\}.$$

Show that $H + K < G$.

Exercise 3.20. Let $H = \{\iota, \varphi\} < D_3$.
(a) Find a different subgroup J of D_3 with only two elements.
(b) Let $HJ = \{hj : h \in H, j \in J\}$. Show that $HJ \not< D_3$.
(c) Why does the result of (b) not contradict the result of Exercise 3.19?

Exercise 3.21. Explain why $\mathbb{R}$ cannot be cyclic.

Exercise 3.22. Fill each blank of Figure 3.3 with the appropriate justification or expression.

Exercise 3.23. Find a group G and subgroups H, K of G such that $A = H \cup K$ *is not* a subgroup of G.

Exercise 3.24. Recall the set of orthogonal matrices from Exercise 0.95.
(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 $\mathrm{SO}(n) < \mathrm{O}(n)$.

3.2: Cosets

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 3.3. Material for Exercise 3.22

One of the most powerful tools in group theory is that of 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

Recall the illustration of how the Division Theorem partitions the integers according to their remainder (Section 0.2). Two aspects of division were critical for this:
- *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 that $A = 4\mathbb{Z} < \mathbb{Z}$ (page 111). All the elements of B have the form $1 + a$ for some $a \in A$. For example, $-3 = 1 + (-4)$. Likewise, all the elements of C have the form $2 + a$ for some $a \in A$, and all the 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.

Equivalence classes

In the context of an equivalence relation, related elements of a set are considered "equivalent".

Example 3.25. 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 3.26.** 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 3.27. Continuing our example above, $3, 19 \in [7]$ but $6 \notin [7]$.

It turns out that equivalence relations partition a set! We will prove this in a moment, but we should look at a concrete example first.

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.

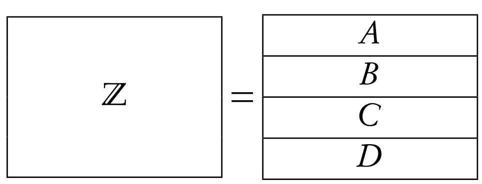
Another way to look at division by d is that it sorts the integers into one of d sets, according to its remainder after division. An illustration using $n = 4$:

$\mathbb{Z}$:	…	-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

$$\begin{aligned} A &= \{\dots,-4,0,4,8,\dots\} = [0] \\ B &= \{\dots,-3,1,5,9,\dots\} = [1] \\ C &= \{\dots,-2,2,6,10,\dots\} = [2] \\ D &= \{\dots,-1,3,7,11,\dots\} = [3]. \end{aligned} \tag{13}$$

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,

$$A \cap B = A \cap C = A \cap D = B \cap C = B \cap D = C \cap D = \emptyset.$$

This should remind you of a partition! (Definition 0.5)

Example 3.28. Let $\mathcal{B} = \{A, B, C, D\}$ where A, B, C, and D are defined as in (13). Since the elements of $\mathcal{B}$ are disjoint, and they cover $\mathbb{Z}$, we conclude that $\mathcal{B}$ is a partition of $\mathbb{Z}$.

A more subtle property is at work here: division has actually produced for us an equivalence relation on the integers.

Theorem 3.29. Let $d \in \mathbb{Z} \setminus \{0\}$, and define a relation $\equiv_d$ in the following way: for any $m, n \in \mathbb{Z}$, we say that $m \equiv_d n$ if and only if they have the same remainder after division by d. This is an equivalence relation.

Proof. We have to prove that $\equiv_d$ is reflexive, symmetric, and transitive.

Reflexive? Let $n \in \mathbb{Z}$. The Division Theorem tells us that the remainder of division of n by d is unique, so $n \equiv_d n$.

Symmetric? Let $m, n \in \mathbb{Z}$, and assume that $m \equiv_d n$. This tells us that m and n have the same remainder after division by d. It doesn't matter whether we state m first or n first; we can just as well say that n and m have the same remainder after division by d. That is, $n \equiv_d m$.

Transitive? Let $\ell, m, n \in \mathbb{Z}$, and assume that $\ell \equiv_d m$ and $m \equiv_d n$. This tells us that ℓ and m have the same remainder after division by d, and m and n have the same remainder after division by d. The Division Theorem tells us that the remainder of division of n by d is unique, so ℓ and n have the same remainder after division by d. That is, $\ell \equiv_d n$. □

We have seen that division induces both a partition and an equivalence relation. Do equivalence relations always coincide with partitions? Surprisingly, yes!

Theorem 3.30. An equivalence relation partitions a set, and any partition of a set defines an equivalence relation.

Actually, it isn't so surprising if you understand the proof, or even 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 other two properties help ensure that no element can be related to two elements that are not themselves related. The proof provides some detail.

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 ($X \cap Z \neq \emptyset$). By the definition of a partition, X and Z are not distinct.

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, so assume S is non-empty. Let $x \in S$. Notice that $[x] \neq \emptyset$, since the reflexive property of an equivalence relation guarantees that $x \sim x$, which implies that $x \in [x]$.

Let $\mathcal{B}$ be the set of all equivalence classes of elements of x; that is, $\mathcal{B} = \{[x] : x \in S\}$. We have already seen that every $x \in S$ 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 $X \cap Y \neq \emptyset$; 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. □

From partitions to cosets

So the partition by remainders of division by four is related to the subgroup A of multiples of 4. This will become very important in Chapter 6. How can we generalize this phenomenon to other groups, even nonabelian ones?

Definition 3.31. 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

$$Ag = \{ag : a \in A\}.$$

As usual, if A is an additive subgroup, we write the left and right cosets of A with g as $g + A$ and $A + g$.

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 means “left cosets”.

Example 3.32. Recall the group D_3 from Section 2.2 and the subgroup $H = \langle\varphi\rangle = \{\iota, \varphi\}$ from Example 3.11. 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 $\rho H \neq H\rho$.

Sometimes, the left coset and the right coset *are* equal. This is always true in abelian groups, as illustrated by Example 3.33.

Example 3.33. Consider the subgroup $H = \{(a,0) : a \in \mathbb{R}\}$ of $\mathbb{R}^2$ from Exercise 3.18. 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 Exercise 3.18. 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 $\mathbb{R}^2$?

The group does not *have* to be abelian in order to have the left and right cosets 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 3.34. 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.)

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 nother 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 a partition (Definition 0.5). Does it hold true in general? What other properties might cosets contain?

Properties of Cosets

We could forgive you for concluding from this that cosets are not especially useful, even if they do generalize division; after all, you don't realize how powerful division is. The rest of this chapter should correct any such misapprehension; for now, we present some properties of cosets that illustrate further their similarities to division.

Theorem 3.35. The cosets of a subgroup partition the group.

Putting this together with Theorem 3.30 implies another nice result.

Corollary 3.36. Let $A < G$. Define a relation $\sim$ on $x, y \in G$ by

$$x \sim y \quad \Longleftrightarrow \quad x \text{ is in the same coset of } A \text{ as } y.$$

This relation is an equivalence relation.

We will make use of this result, in due course.

Proof of Theorem 3.35. 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 $g = ge$. Since $e \in A$ by definition of subgroup, $g = ge \in gA$. 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, $X \cap Y = \emptyset$. By way of contradiction, assume that $X \neq Y$ but $X \cap Y \neq \emptyset$. Since $X \neq Y$, one of the two cosets contains an element that does not appear in the other; without loss of generality, assume that $z \in X$ but $z \notin Y$. By definition, there exist $x, y \in G$ such that $X = xA$ and $Y = yA$; we can write $z = xa$ for some $a \in A$. Since $X \cap Y \neq \emptyset$, there exists some $w \in X \cap Y$; by definition, we can find $b, c \in A$ such that $w = xb = yc$. Solve this last equation for x, and we have $x = (yc)\, b^{-1}$. Substitute this into the equation for z, and we have

$$z = xa = \left[(yc)\, b^{-1}\right] a \underset{\text{ass.}}{=} y\left(cb^{-1}a\right).$$

Since A is a subgroup, hence a group, it is closed under inverses and multiplication, so $cb^{-1}a \in A$. But then $z = y\left(cb^{-1}a\right) \in yA$, which contradicts the choice of z! The assumption that we could find distinct cosets that are not disjoint must have been false, and since X and Y were arbitrary, this holds for all cosets of A.

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 3.37 (Equality of cosets)**.** Let G be a group and $H < G$. All of the following hold:
(CE1) $eH = H$.
(CE2) For all $a \in G$, $a \in H$ iff $aH = H$.
(CE3) For all $a, b \in G$, $aH = bH$ if and only if $a^{-1}b \in H$.
(CE4) For all $a, b \in G$, $aH = bH$ if and only if $a \in bH$.

As usual, you should keep in mind that in additive groups the first three conditions translate to
(CE1) $0 + H = H$.
(CE2) For all $a \in G$, if $a \in H$ then $a + H = H$.
(CE3) For all $a, b \in G$, $a + H = b + H$ if and only if $a - b \in H$.
(CE4) For all $a, b \in G$, $a + H = b + H$ if and only if $a \in b + H$.

Proof. We only sketch the proof here. You will fill in the details in Exercise 3.46. Remember that part of this problem involves proving that two sets are equal, and to prove that, you should prove that each is a subset of the other.

(CE1) is "obvious" (but fill in the details anyway).

We'll skip (CE2) for the moment, and move to (CE3). Since (CE3) is also an equivalence, we have to prove two directions. Let $a, b \in G$. First, assume that $aH = bH$. By the identity property, $e \in H$, so $b = be \in bH$. Hence, $b \in aH$; that is, we can find $h \in H$ such that $b = ah$. By substitution and the properties of a group, $a^{-1}b = a^{-1}(ah) = h$, so $a^{-1}b \in H$.

Conversely, assume that $a^{-1}b \in H$. We must show that $aH = bH$, which requires us to show that $aH \subseteq bH$ and $aH \supseteq bH$. Since $a^{-1}b \in H$, we have

$$b = a\left(a^{-1}b\right) \in aH.$$

We can thus write $b = ah$ for some $h \in H$. Let $y \in bH$; then $y = b\widehat{h}$ for some $\widehat{h} \in H$, and we have $y = (ah)\widehat{h} \in H$. Since y was arbitrary in bH, we now have $aH \supseteq bH$.

Although we could build a similar argument to show that $aH \subseteq bH$, instead we point out that $aH \supseteq bH$ implies that $aH \cap bH \neq \emptyset$. The cosets are not disjoint, so by Theorem 3.35, they are not distinct: $aH = bH$.

Now we turn to (CE2). Let $a \in G$, and assume $a \in H$. By the inverse property, $a^{-1} \in H$. We know that $e \in H$, so by closure, $a^{-1}e \in H$. We can now use (CE3) and (CE1) to determine that $aH = eH = H$.

Finally, we look at (CE4). Let $a, b \in G$. Since $e \in H$, we know that $a = ae \in aH$ and $b = be \in bH$. If $aH = bH$, then $a \in bH$. Conversely, if $a \in bH$, then aH and bH are not disjoint; the partition property implies that $aH = bH$. □

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.

Exercises.

Exercise 3.38. Show explicitly why left and right cosets are equal in abelian groups.

Exercise 3.39. In Exercise 3.14, you showed that $\Omega_2 < \Omega_8$. Compute the left and right cosets of Ω_2 in Ω_8.

Exercise 3.40. Let $\{e, a, b, a+b\}$ be the Klein 4-group. (See Exercises 2.19 on page 81, 2.69 on page 101, and 3.15 on page 114.) Compute the left cosets of $\langle a \rangle$.

Exercise 3.41. Let $\{e, a, b, a+b\}$ be the Klein 4-group. Compute the left cosets of $\{e\}$.

Exercise 3.42. Compute the left cosets of $\langle -1 \rangle$ in Q_8. Repeat this for $\langle \mathbf{j} \rangle$.

Exercise 3.43. In Exercise 3.20 on page 115, you found another subgroup K of order 2 in D_3. Does K satisfy the property $\alpha K = K\alpha$ for all $\alpha \in D_3$?

Exercise 3.44. Recall the subgroup L of $\mathbb{R}^2$ from Exercise 3.18 on page 114.

Let G be a group and $H < G$.
Claim: $eH = H$.

1. First we show that _____. Let $x \in eH$.
   (a) By definition, _____.
   (b) By the identity property, _____.
   (c) By definition, _____.
   (d) We had chosen an arbitrary element of eH , so by inclusion, _____.
2. Now we show the converse. Let _____.
   (a) By the identity property, _____.
   (b) By definition, _____ $\in eH$.
   (c) We had chosen an arbitrary element, so by inclusion, _____.

Figure 3.4. Material for Exercise 3.46

(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 geometric interpretation of the cosets of L in $\mathbb{R}^2$ to explain why the cosets of L partition $\mathbb{R}^2$.

Exercise 3.45. Recall $D_n(\mathbb{R})$ from Exercise 3.17 on page 114. Give a description in set notation for

$$\begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} + D_2(\mathbb{R}).$$

List some elements of the coset.

Exercise 3.46. Fill in each blank of Figure 3.46 with the appropriate justification or statement.

Exercise 3.47. Recall the relation $\equiv_d$ described in Theorem 3.29.
(a) Explain why $2 \cdot 3 \equiv_6 0$.
(b) Explain how (a) is related to Exercise 0.87.
(c) Integer equations are equivalence relations. Explain how part (a) shows 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.

3.3: Lagrange's Theorem

This section introduces an important result describing the number of cosets a subgroup can have. This leads to some properties regarding the order of a group and any of its elements.

Notation 3.48. 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:

$$A \backslash G = \{Ag : g \in G\}.$$

Example 3.49. Let $G = \mathbb{Z}$ and $A = 4\mathbb{Z}$. We saw in Example 3.28 that

$$G/A = \mathbb{Z}/4\mathbb{Z} = \{A, 1+A, 2+A, 3+A\}.$$

We actually "waved our hands" in Example 3.28. 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 $0 \leq r < 4$. 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 3.50. Let $G = D_3$ and $K = \left\{\iota, \rho, \rho^2\right\}$ as in Example 3.34, then

$$G/K = D_3/\langle\rho\rangle = \{K, \varphi K\}.$$

Example 3.51. Let $H < \mathbb{R}^2$ be as in Example 3.10 on page 112; 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

$$(1,1) + H,\ (4,-2) + H.$$

Here's a question for you to think about. Speaking *geometrically*, what do the elements of G/H look like? This question is similar to Exercise 3.44.

It is important to keep in mind that G/A is a set whose elements are also sets. As a result, showing equality of two elements of G/A requires one to show that two sets are equal.

When G is finite, a simple formula gives us the size of G/A.

Theorem 3.52 (Lagrange's Theorem)**.** Let G be a group of finite order, and $A < G$. Then
$$|G/A| = \frac{|G|}{|A|}.$$

Lagrange's Theorem states that the number of elements in G/A is the same as the quotient of the order of G by the order of A. The notation of cosets is somewhat suggestive of the relationship we illustrated at the beginning of Section 3.2 between cosets and division of the integers. Nevertheless, 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 dividing, or partitioning, if you will, the group G by by the cosets of its subgroup A, obtaining the set of cosets G/A.

Proof. From Theorem 3.35 we know that the cosets of A partition G. How many such cosets are there? $|G/A|$, by definition! Each coset has the same size, $|A|$. 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 3.53. The order of a subgroup divides the order of a group.

Example 3.54. Let G be the Klein 4-group (see Exercises 2.19 on page 81, 2.69 on page 101, and 3.15 on page 114). Every subgroup of the Klein 4-group has order 1, 2, or 4. As predicted by Corollary 3.53, 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 Exercise 3.20 on page 115 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 3.55. In a finite group G, the order of any element divides the order of a group.

Proof. You do it! See Exercise 3.57. □

Exercises.

Exercise 3.56. Recall from Exercise 3.14 that if $d \mid n$, then $\Omega_d < \Omega_n$. How many cosets of Ω_d are there in Ω_n?

Exercise 3.57. Fill in each blank of Figure 3.57 with the appropriate justification or expression.

Exercise 3.58. Suppose that a group G has order 8, but is not cyclic. Show that $g^4 = e$ for all $g \in G$.

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 _____, $\text{ord}(x)$ 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 3.5. Material for Exercise 3.57

Exercise 3.59. Let G be a finite group, and $g \in G$. Show that $g^{|G|} = e$.

Exercise 3.60. Suppose that a group has five elements. Why *must* it be abelian?

Exercise 3.61. Find a sufficient (but not necessary) condition on the order of a group of order at least two that guarantees that the group is cyclic.

3.4: Quotient Groups

Let $A < G$. Is there a natural generalization of the operation of G that makes G/A a group? By a "natural" generalization, we mean something like

$$(gA)(hA) = (gh)A.$$

"Normal" subgroups

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 3.62 shows how it can go wrong.

Example 3.62. Recall $H = \langle\varphi\rangle < D_3$ from Example 3.32. 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 3.63. 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 3.37, 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 $4m$.

Does that mean we cannot make a group out of cosets of nonabelian groups? Not quite. The key in Example 3.63 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* $ga = a'g$.

Think about this in terms of sets: for every $g \in G$ and for *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* $g \in G$.

This property merits a definition.

Definition 3.64. Let $A < G$. If

$$gA = Ag$$

for every $g \in G$, then A is a **normal subgroup** of G.

Notation 3.65. We write $A \triangleleft G$ to indicate that A is a normal subgroup of 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 3.66. Let $A < G$. Then (CO1) implies (CO2).
(CO1) $A \triangleleft G$.
(CO2) Let $X, Y \in G/A$ and $x, y \in G$ such that $X = xA$ and $Y = yA$. The operation $\cdot$ on G/A defined by

$$XY = (xy)A$$

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 3.37(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 $w\check{\alpha}$ as $x\hat{\alpha}$ and $y\dot{\alpha}$ as $z\ddot{\alpha}$, where $\check{\alpha}, \hat{\alpha}, \dot{\alpha}, \ddot{\alpha} \in A$.

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 $\check{a} \in A$ such that $ya = \check{a}y$. By substitution, $t = w(\check{a}y)$. By the associative property, $t = (w\check{a})y$. By definition of a coset, $w\check{a} \in wA$. By hypothesis, A is normal, so $wA = Aw$. Thus $w\check{a} \in Aw$. By hypothesis, $W = X$; that is, $wA = xA$. Thus $w\check{a} \in xA$, and by definition of a coset, $w\check{a} = x\hat{a}$ for some $\hat{a} \in A$. By substitution, $t = (x\hat{a})y$. The associative property again gives us $t = x(\hat{a}y)$; since A is normal we can write $\hat{a}y = y\dot{a}$ for some $\dot{a} \in A$. Hence $t = x(y\dot{a})$. Now,

$$y\dot{a} \in yA = Y = Z = zA,$$

so we can write $y\dot{a} = z\ddot{a}$ for some $\ddot{a} \in A$. By substitution and the definition of coset arithmetic,

$$t = x(z\ddot{a}) = (xz)\ddot{a} \in (xz)A = (xA)(zA) = XZ.$$

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 3.63 shows the following Theorem.

Theorem 3.67. Let G be an abelian group, and $H < G$. Then $H \triangleleft G$.

Proof. You do it! See Exercise 3.81. □

We said before that we don't need an abelian group to have a normal subgroup. Here's a *great* example.

Example 3.68. Let

$$A_3 = \{\iota, \rho, \rho^2\} < D_3.$$

We call A_3 the **alternating group** on three elements. We claim that $A_3 \triangleleft D_3$. 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 3.34, 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 3.69. Let G be a group. If $A \triangleleft G$, then G/A is a group.

Proof. Assume $A \triangleleft G$. By Lemma 3.66, 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 $A \triangleleft G$, as shown in Lemma 3.66: 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,

$$(XY)Z = ((g_1 g_2) g_3)A.$$

(Note the parentheses grouping $g_1 g_2$.) Now apply the associative property of G and reverse the previous steps to obtain

$$\begin{aligned}(XY)Z &= (g_1(g_2 g_3))A \\ &= (g_1 A)((g_2 g_3)A) \\ &= (g_1 A)[(g_2 A)(g_3 A)] \\ &= X(YZ).\end{aligned}$$

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 $e \in A$, Lemma 3.37 on page 121 implies that $A = eA$, so

$$XA = (gA)(eA) = (ge)A = gA = X.$$

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,

$$XY = (gA)\left(g^{-1}A\right) = \left(gg^{-1}\right)A = eA = A.$$

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 3.70. Let G be a group, and $A \triangleleft G$. 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 3.71. 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 composition 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 3.5.

Conjugation

It can be hard to show that the left and right cosets of a subgroup of infinite order are equal. For example, how do we decide whether $\mathrm{SO}_n(\mathbb{R}) \triangleleft \mathrm{O}_n(\mathbb{R})$? An alternative approach does the trick.

Definition 3.72. Let G be a group, $g \in G$, and $H < G$. Define the **conjugation** of H by g as

$$gHg^{-1} = \{h^g : h \in H\}.$$

(The notation h^g is the definition of conjugation from Exercise 2.40 on page 83; that is, $h^g = ghg^{-1}$.)

Theorem 3.73. $H \triangleleft G$ if and only if $H = gHg^{-1}$ for all $g \in G$.

Proof. You do it! See Exercise 3.82. □

Example 3.74. We posed the question of whether $\mathrm{SO}_n(\mathbb{R}) \triangleleft \mathrm{O}_n(\mathbb{R})$. 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 3.73, $\mathrm{SO}_n(\mathbb{R}) \triangleleft \mathrm{O}_n(\mathbb{R})$.

Example 3.75. 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 $AMA^{-1} \notin \mathrm{O}_2(\mathbb{R})$.

Exercises.

Exercise 3.76. Show that for any group G, $\{e\} \triangleleft G$ and $G \triangleleft G$.

Exercise 3.77. Recall from Exercise 3.14 that if $d \mid n$, then $\Omega_d < \Omega_n$.

(a) Explain how we know that, in fact, $\Omega_d \triangleleft \Omega_n$.

(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?

Exercise 3.78. In Exercise 3.42, you computed the left cosets of $\langle \mathbf{j} \rangle$ in Q_8.

(a) Compute the right cosets, as well.

(b) Is $\langle \mathbf{j} \rangle$ a normal subgroup of Q_8? If so, compute the Cayley table of $Q_8 / \langle \mathbf{j} \rangle$.

Exercise 3.79. Let $H = \langle \mathbf{i} \rangle < Q_8$.

(a) Show that $H \triangleleft Q_8$ by computing all the cosets of H.

(b) Compute the Cayley table of Q_8/H.

Exercise 3.80. In Exercise 3.42, 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?

Exercise 3.81. Let G be an abelian group. Explain why for any $H < G$ we know that $H \triangleleft G$.

Let G be a group, and $H < G$.
Claim: $H \triangleleft G$ if and only if $H = gHg^{-1}$ for all $g \in G$.
Proof:

1. First, we show that if $H \triangleleft G$, 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 $H \triangleleft G$.
   (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 $h = g^{-1}h'g$.
      (A key point here is that this is true 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, $gH = Hg$.

Figure 3.6. Material for Exercise 3.82

Exercise 3.82. Fill in each blank of Figure 39 with the appropriate justification or statement.[19]

[19] 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: $Z(G) \triangleleft G$.

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 $Z(G)$, $(gh^{-1})x =$_____=_____$= \ldots = x(gh^{-1})$.
   (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 3.7. Material for Exercise 3.86

Exercise 3.83. Recall the subgroup L of $\mathbb{R}^2$ from Exercises 3.18 on page 114 and 3.44 on page 122.
(a) Explain how we know that $L \triangleleft \mathbb{R}^2$ *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.

Exercise 3.84. Explain why every subgroup of $D_m(\mathbb{R})$ is normal.

Exercise 3.85. Show that Q_8 is not a normal subgroup of $\mathrm{GL}_m(\mathbb{C})$.

Exercise 3.86. Fill in every blank of Figure 3.86 with the appropriate justification or statement.

Exercise 3.87. Let G be a group, and $H < G$. Define the **normalizer** of H as

$$N_G(H) = \{g \in G : gH = Hg\}.$$

Show that $H \triangleleft N_G(H)$.

Exercise 3.88. 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:

- $A \triangleleft G$; and
- G/A is abelian.

Exercise 3.89. Recall from Exercise 2.40 on page 83 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 39 with the appropriate justification or statement.

Definition 3.90. We call $[G,G]$ the **commutator subgroup** of G, and make use of it in Section 9.4.

Claim: For any group G, $[G,G]$ is a normal subgroup of G.

Proof:

1. Let _____.
2. We will use Exercise 3.82 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 _____, $h = ex^{-1}ey^{-1}exeye$.
   (d) By _____, $h = \left(gg^{-1}\right)x^{-1}\left(gg^{-1}\right)y^{-1}\left(gg^{-1}\right)x\left(gg^{-1}\right)y\left(gg^{-1}\right)$.
   (e) By _____, $h = g\left(g^{-1}x^{-1}g\right)\left(g^{-1}y^{-1}g\right)\left(g^{-1}xg\right)\left(g^{-1}yg\right)g^{-1}$.
   (f) By _____, $h = g\left(x^{-1}\right)^{g^{-1}}\left(y^{-1}\right)^{g^{-1}}\left(x^{g^{-1}}\right)\left(y^{g^{-1}}\right)g^{-1}$.
   (g) By Exercise 2.40 on page 83(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 3.8. Material for Exercise 3.89

3.5: "Clockwork" groups

By Theorem 3.67, every subgroup H of $\mathbb{Z}$ is normal. Let $n \in \mathbb{Z}$; since $n\mathbb{Z} < \mathbb{Z}$, it follows that $n\mathbb{Z} \triangleleft \mathbb{Z}$. Thus $\mathbb{Z}/n\mathbb{Z}$ is a quotient group.

We used $n\mathbb{Z}$ in many examples of subgroups. One reason is that you are accustomed to working with $\mathbb{Z}$, so it should be conceptually easy. Another reason is that the quotient group $\mathbb{Z}/n\mathbb{Z}$ has a vast array of applications in number theory and computer science. You will see some of these in Chapter 6. Because this group is so important, we give it several special names.

Definition 3.91. Let $n \in \mathbb{Z}$. We call the quotient group $\mathbb{Z}/n\mathbb{Z}$
- $\mathbb{Z}$ **mod** n, or
- the **linear residues modulo** n.

Notation 3.92. It is common to write $\mathbb{Z}_n$ instead of $\mathbb{Z}/n\mathbb{Z}$.

Example 3.93. You already saw a bit of $\mathbb{Z}_4 = \mathbb{Z}/4\mathbb{Z}$ at the beginning of Section 3.2 and again in Example 3.63. Recall that $\mathbb{Z}_4 = \{4\mathbb{Z}, 1+4\mathbb{Z}, 2+4\mathbb{Z}, 3+4\mathbb{Z}\}$. Addition in this group will always give us one of those four representations of the cosets:

$$\begin{aligned}(2+4\mathbb{Z})+(1+4\mathbb{Z}) &= 3+4\mathbb{Z};\\ (1+4\mathbb{Z})+(3+4\mathbb{Z}) &= 4+4\mathbb{Z} = 4\mathbb{Z};\\ (2+4\mathbb{Z})+(3+4\mathbb{Z}) &= 5+4\mathbb{Z} = 1+4\mathbb{Z};\end{aligned}$$

and so forth.

Reasoning similar to that used at the beginning of Section 3.2 would show that

$$\mathbb{Z}_{31} = \mathbb{Z}/31\mathbb{Z} = \{31\mathbb{Z}, 1+31\mathbb{Z}, \ldots, 30+31\mathbb{Z}\}.$$

We show this explicitly in Theorem 3.97.

Before looking at some properties of $\mathbb{Z}_n$, let's look for an easier way to talk about its elements. It is burdensome to write $a+n\mathbb{Z}$ whenever we want to discuss an element of $\mathbb{Z}_n$, so we adopt the following convention.

Notation 3.94. Let $A \in \mathbb{Z}_n$ and choose $r \in \mathbb{Z}$ such that $A = r + n\mathbb{Z}$.
- If it is clear from context that A is an element of $\mathbb{Z}_n$, then we simply write r instead of $r+n\mathbb{Z}$.
- If we want to emphasize that A is an element of $\mathbb{Z}_n$ (perhaps there are a lot of integers hanging about) then we write $[r]_n$ instead of $r+n\mathbb{Z}$.
- If the value of n is obvious from context, we simply write $[r]$.

To help you grow accustomed to the notation $[r]_n$, we use it for the rest of this chapter, even when n is mindbogglingly obvious.

The first property is that, for most values of n, $\mathbb{Z}_n$ has finitely many elements. To show that there are finitely many elements of $\mathbb{Z}_n$, we rely on the following fact, which is important enough to highlight as a separate result.

Lemma 3.95. Let $n \in \mathbb{Z}\backslash\{0\}$ and $[a]_n \in \mathbb{Z}_n$. Use the Division Theorem to choose $q, r \in \mathbb{Z}$ such that $a = qn + r$ and $0 \le r < |n|$. Then $[a]_n = [r]_n$.

The proof of Lemma 3.95 on the preceding page is similar to the discussion in Example 3.49 on page 124, so you might want to reread that.

Proof. We give two different proofs. Both are based on the fact that $[a]_n$ and $[r]_n$ are *cosets;* so showing that they are equal is tantamount to showing that a and r are different elements of the same set.

(1) By definition and substitution,

$$\begin{aligned}[a]_n &= a + n\mathbb{Z} \\ &= (qn + r) + n\mathbb{Z} \\ &= \{(qn + r) + nd : d \in \mathbb{Z}\} \\ &= \{r + n(q + d) : d \in \mathbb{Z}\} \\ &= \{r + nm : m \in \mathbb{Z}\} \\ &= r + n\mathbb{Z} \\ &= [r]_n .\end{aligned}$$

(2) Rewrite $a = qn + r$ as $a - r = qn$. By definition, $a - r \in n\mathbb{Z}$. The immensely useful Lemma 3.37 shows that $a + n\mathbb{Z} = r + n\mathbb{Z}$, and the notation implies that $[a]_n = [r]_n$. □

Definition 3.96. On account of Lemma 3.95, we can designate the remainder of division of a by n, whose value is between 0 and $|n| - 1$, inclusive, as the **canonical representation** of $[a]_n$ in $\mathbb{Z}_n$.

Theorem 3.97. $\mathbb{Z}_n$ is finite for every nonzero $n \in \mathbb{Z}$. In fact, if $n \neq 0$ then $\mathbb{Z}_n$ has $|n|$ elements corresponding to the remainders from division by n: 0, 1, 2, ..., $n - 1$.

Proof. Lemma 3.95 on the preceding page states that every element of such $\mathbb{Z}_n$ can be represented by $[r]_n$ for some $r \in \mathbb{Z}$ where $0 \leq r < |n|$. But there are only $|n|$ possible choices for such a remainder. □

Let's look at how we can perform arithmetic in $\mathbb{Z}_n$.

Lemma 3.98. Let $d, n \in \mathbb{Z}$ and $[a]_n, [b]_n \in \mathbb{Z}_n$. Then

$$[a]_n + [b]_n = [a + b]_n$$

and

$$d\,[a]_n = [da]_n .$$

For example, $[3]_7 + [9]_7 = [3 + 9]_7 = [12]_7 = [5]_7$ and $-4\,[3]_5 = [-4 \cdot 3]_5 = [-12]_5 = [3]_5$.

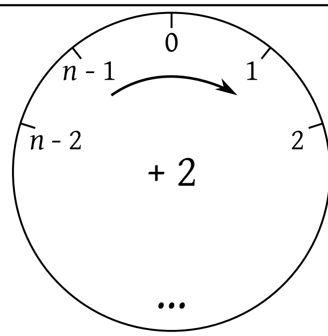


Figure 3.9. Addition in $\mathbb{Z}_n$ is "clockwork": $[n-1]_n + [2]_n = [1]_n$.

Proof. The proof really amounts to little more than manipulating the notation. By the definitions of coset addition and of $\mathbb{Z}_n$,

$$\begin{aligned}[a]_n + [b]_n &= (a + n\mathbb{Z}) + (b + n\mathbb{Z}) \\ &= (a + b) + n\mathbb{Z} \\ &= [a + b]_n .\end{aligned}$$

For $d\,[a]_n$, we consider three cases.

If $d = 0$, then $d\,[a]_n = [0]_n$ by Notation 2.55 on page 93, and $[0]_n = [0 \cdot a]_n = [da]_n$. By substitution, then, $d\,[a]_n = [da]_n$.

If d is positive, then the expression $d\,[a]_n$ is the sum of d copies of $[a]_n$, which the Lemma's first claim (now proved) implies to be

$$\begin{aligned}\underbrace{[a]_n + [a]_n + \cdots + [a]_n}_{d \text{ times}} &= [2a]_n + \underbrace{[a]_n + \cdots + [a]_n}_{d-2 \text{ times}} \\ &\vdots \\ &= [da]_n .\end{aligned}$$

If d is negative, then Notation 2.55 again tells us that $d\,[a]_n$ is the sum of $|d|$ copies of $-[a]_n$. So, what is the additive inverse of $[a]_n$? Using the first claim, $[a]_n + [-a]_n = [a + (-a)]_n = [0]_n$, so $-[a]_n = [-a]_n$. By substitution,

$$\begin{aligned}d\,[a]_n &= |d|\,(-[a_n]) = |d|\,[-a]_n \\ &= [|d| \cdot (-a)]_n = [-d \cdot (-a)]_n = [da]_n .\end{aligned}$$

□

Lemmas 3.95 and 3.98 imply that each $\mathbb{Z}_n$ acts as a **"clockwork" group**. Why?

- To add $[a]_n$ and $[b]_n$, let $c = a + b$.
- If $0 \leq c < |n|$, then you are done. After all, division of c by n gives $q = 0$ and $r = c$.
- Otherwise, $c < 0$ or $c \geq |n|$, so we divide c by n, obtaining q and r where $0 \leq r < |n|$. The sum is $[r]_n$.

We call this "clockwork" because it counts like a clock: if you sit down at 5 o'clock and wait two hours, you rise at not at 13 o'clock, but at $13 - 12 = 1$ o'clock. See Figure 3.5.

It should be clear from Example 2.11 on page 78 as well as Exercise 2.18 on page 81 that $\mathbb{Z}_2$ and $\mathbb{Z}_3$ have precisely the same structure as the groups of order 2 and 3. On the other hand, we saw in Exercise 2.19 on page 81 that there are two possible structures for a group of order 4: the Klein 4-group, and a cyclic group. Which structure does $\mathbb{Z}_4$ have?

Example 3.99. Use Lemma 3.98 to observe that

$$\langle [1]_4 \rangle = \{[0]_4, [1]_4, [2]_4, [3]_4\}$$

since $[2]_4 = [1]_4 + [1]_4$, $[3]_4 = [2]_4 + [1]_4$, and $[0]_4 = 0 \cdot [1]_4$ (or $[0]_4 = [3]_4 + [1]_4$).

The fact that $\mathbb{Z}_4$ was cyclic makes one wonder: is $\mathbb{Z}_n$ always cyclic? Yes!

Theorem 3.100. $\mathbb{Z}_n$ is cyclic for every $n \in \mathbb{Z}$.

This theorem has a more general version, which you will prove in the homework.

Proof. Let $n \in \mathbb{Z}$ and $[a]_n \in \mathbb{Z}_n$. By Lemma 3.98,

$$[a]_n = [a \cdot 1]_n = a\,[1]_n \in \langle [1]_n \rangle .$$

Since $[a]_n$ was arbitrary in $\mathbb{Z}_n$, $\mathbb{Z}_n \subseteq \langle [1]_n \rangle$. Closure implies that $\mathbb{Z}_n \supseteq \langle [1]_n \rangle$, so in fact $\mathbb{Z}_n = \langle [1]_n \rangle$, and $\mathbb{Z}_n$ is therefore cyclic. □

Not *every* non-zero element necessarily generates $\mathbb{Z}_n$. We know that $[2]_4 + [2]_4 = [4]_4 = [0]_4$, so in $\mathbb{Z}_4$, we have

$$\langle [2]_4 \rangle = \{[0]_4, [2]_4\} \subsetneq \mathbb{Z}_4.$$

A natural and interesting followup question is, which non-zero elements *do* generate $\mathbb{Z}_n$? You need a bit more background in number theory before you can answer that question, but in the exercises you will build some more addition tables and use them to formulate a hypothesis.

The following important lemma gives an "easy" test for whether two integers are in the same coset of $\mathbb{Z}_n$.

Lemma 3.101. Let $a, b, n \in \mathbb{Z}$ and assume that $n \neq 0$. The following are equivalent.
(A) $a + n\mathbb{Z} = b + n\mathbb{Z}$.
(B) $[a]_n = [b]_n$.
(C) $n \mid (a - b)$.

Proof. You do it! See Exercise 3.108. □

Exercises.

Exercise 3.102. We showed that $\mathbb{Z}_n$ is finite for $n \neq 0$. What if $n = 0$? How many elements would it have? Illustrate a few additions and subtractions, and indicate whether you think that $\mathbb{Z}_0$ is an interesting or useful group.

Exercise 3.103. In the future, we won't consider $\mathbb{Z}_n$ when $n < 0$. Show that this is because $\mathbb{Z}_n = \mathbb{Z}_{|n|}$.

Exercise 3.104. Write out the Cayley tables for $\mathbb{Z}_2$ and $\mathbb{Z}_3$. Remember that the operation is addition.

Exercise 3.105. Write out the Cayley table for $\mathbb{Z}_5$. Remember that the operation is addition. Which elements generate $\mathbb{Z}_5$?

Exercise 3.106. Write down the Cayley table for $\mathbb{Z}_6$. Remember that the operation is addition. Which elements generate $\mathbb{Z}_6$?

Exercise 3.107. Compare the results of Example 3.99 and Exercises 3.104, 3.105, and 3.106. Formulate a conjecture as to which elements generate $\mathbb{Z}_n$. Do not try to prove your example.

Exercise 3.108. Prove Lemma 3.101.

Exercise 3.109. Prove the following generalization of Theorem 3.100: If G is a cyclic group and $A \triangleleft G$, then G/A is cyclic.

Chapter 4:
Isomorphisms

We have on occasion observed that different groups have the same Cayley table. We have also talked about different groups having the same structure: regardless of whether a group of order two is additive or multiplicative, its elements behave in exactly the same fashion. The groups may consist of elements whose construction was quite different, and the definition of the operation may also be different, but the "group behavior" is nevertheless identical.

We saw in Chapter 1 that algebraists describe such a relationship between two monoids as *isomorphic*. Isomorphism for groups has the same intuitive meaning as isomorphism for monoids:

If two groups G and H have identical group structure,
we say that G and H are *isomorphic*.

We want to study isomorphism of groups in quite a bit of detail, so to define isomorphism precisely, we start by reconsidering another topic that you studied in the past, functions. There we will also introduce the related notion of *homomorphism.* Despite the same basic intuitive definition, the precise definition of group homomorphism turns out simpler than for monoids. This is the focus of Section 4.1. Section 4.2 lists some results that should help convince you that the existence of an isomorphism does, in fact, show that two groups have an identical group structure. Section 4.3 describes how we can create new isomorphisms from a homomorphism's *kernel,* a special subgroup defined by a homomorphism. Section 4.4 introduces a class of isomorphism that is important for later applications, an *automorphism*.

4.1: Homomorphisms

Groups have more structure than monoids. Just as a monoid homomorphism would require that we preserve both identities and the operation (page 57), you might infer that the requirements for a group isomorphism are stricter than those for a monoid isomorphism. After all, you have to preserve not only identities and the operation, but inverses as well.

In fact, the additional structure of groups allows us to have *fewer* requirements for a group homomorphism.

Group isomorphisms

Definition 4.1. Let $(G,\times)$ and $(H,+)$ be groups. If there exists a function $f: G \to H$ that preserves the operation, which is to say that

$$f(xy) = f(x) + f(y) \quad \text{for every } x, y \in G,$$

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

This definition requires the preservation of neither inverses nor identities! You might conclude from this that group homomorphism aren't even monoid homomorphisms; we will see in a moment that this is quite untrue!

Notation 4.2. As with monoids, you have to be careful with the fact that different groups have different operations. Depending on the context, the proper way to describe the homomorphism property may be

- $f(xy) = f(x) + f(y)$;
- $f(x+y) = f(x)f(y)$;
- $f(x \circ y) = f(x) \odot f(y)$;
- etc.

Example 4.3. A trivial example of a homomorphism, but an important one, is the identity function $\iota : G \to G$ by $\iota(g) = g$ for all $g \in G$. It should be clear that this is a homomorphism, since for all $g, h \in G$ we have

$$\iota(gh) = gh = \iota(g)\iota(h).$$

For a non-trivial homomorphism, let $f : \mathbb{Z} \to 2\mathbb{Z}$ by $f(x) = 4x$. Then f is a group homomorphism, since for any $x \in \mathbb{Z}$ we have

$$f(x) + f(y) = 4x + 4y = 4(x+y) = f(x+y).$$

Hopefully, the homomorphism property reminds you of certain special functions and operations that you studied in Linear Algebra or Calculus. Recall from Exercise 2.35 that $\mathbb{R}^{>0}$, the set of all positive real numbers, is a multiplicative group.

Example 4.4. Let $f : (\mathrm{GL}_m(\mathbb{R}), \times) \to (\mathbb{R}\backslash\{0\}, \times)$ by $f(A) = \det A$. By Theorem 0.85, $\det A \cdot \det B = \det(AB)$. Thus

$$f(A) \cdot f(B) = \det A \cdot \det B = \det A \cdot \det B = \det(AB) = f(AB),$$

implying that f is a homomorphism of groups.

Let's look at a clockwork group.

Example 4.5. Let $n \in \mathbb{Z}$ such that $n > 1$, and let $f : (\mathbb{Z}, +) \to (\mathbb{Z}_n, +)$ by the assignment $f(x) = [x]_n$. We claim that f is a homomorphism. *Why?* From Lemma 3.98, we know that for any $x, y \in \mathbb{Z}_n$, $f(x+y) = [x+y]_n = [x]_n + [y]_n = f(x) + f(y)$.

By preserving the operation, we preserve an enormous amount of information about a group. If there is a homomorphism f from G to H, then elements of the **image of** G,

$$f(G) = \{h \in H : \exists g \in G \text{ such that } f(g) = h\}$$

act the same way as their **preimages** in G.

This does *not* imply that the *group structure* is the same. In Example 4.5, for example, f is a homomorphism from an infinite group to a finite group; even if the group operations behave in a similar way, the groups themselves are inherently different. If we can show that the groups have the same "size" in addition to a similar operation, then the groups are, for all intents and purposes, identical.

How do we decide that two groups have the same size? For finite groups, this is "easy": count the elements. We can't do that for infinite groups, so we need something a little more general.[20]

[20]The standard method in set theory of showing that two sets are the same "size" is to show that there exists a one-

Definition 4.6. Let $f : G \to H$ be a homomorphism of groups. If f is also a bijection, then we say that G is **isomorphic** to H, write $G \cong H$, and call f an **isomorphism**.

Example 4.7. Recall the homomorphisms of Example 4.3,

$$\iota : G \to G \quad \text{by} \quad \iota(g) = g \qquad \text{and} \qquad f : \mathbb{Z} \to 2\mathbb{Z} \quad \text{by} \quad f(x) = 4x.$$

First we show that ι is an isomorphism. We already know it's a homomorphism, so we need only show that it's a bijection.

one-to-one: Let $g, h \in G$. Assume that $\iota(g) = \iota(h)$. By definition of ι, $g = h$. Since g and h were arbitrary in G, ι is one-to-one.

onto: Let $g \in G$. We need to find $x \in G$ such that $\iota(x) = g$. Using the definition of ι, $x = g$ does the job. Since g was arbitrary in G, ι is onto.

Now we show that f is not a bijection, and hence not an isomorphism.

***not** onto:* There is no element $a \in \mathbb{Z}$ such that $f(a) = 2$. If there were, $4a = 2$. The only possible solution to this equation is $a = 1/2 \notin \mathbb{Z}$.

This is despite the fact that f is one-to-one:

one-to-one: Let $a, b \in \mathbb{Z}$. Assume that $f(a) = f(b)$. By definition of f, $4a = 4b$. Then $4(a - b) = 0$; by the zero product property of the integers, $4 = 0$ or $a - b = 0$. Since $4 \neq 0$, we must have $a - b = 0$, or $a = b$. We assumed $f(a) = f(b)$ and showed that $a = b$. Since a and b were arbitrary, f is one-to-one.

Example 4.8. Recall the homomorphism of Example 4.4,

$$f : \mathrm{GL}_m(\mathbb{R}) \to \mathbb{R}^{>0} \quad \text{by} \quad f(A) = |\det A|.$$

We claim that f is onto, but not one-to-one.

That f is not one-to-one: Observe that f maps both of the following two diagonal matrices to 2, even though the matrices are unequal:

$$A = \begin{pmatrix} 2 & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & & & & \\ & 2 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & \ddots \end{pmatrix}.$$

(Unmarked entries are zeroes.)

That f is onto: Let $x \in \mathbb{R}^{>0}$; then $f(A) = x$ where A is the diagonal matrix

$$A = \begin{pmatrix} x & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots \end{pmatrix}.$$

to-one, onto function between the sets. For example, one can use this definition to show that $\mathbb{Z}$ and $\mathbb{Q}$ are the same size, but $\mathbb{Z}$ and $\mathbb{R}$ are not. So an isomorphism is a homomorphism that also shows that two sets are the same size.

(Again, unmarked entries are zeroes.)

We *cannot* conclude from these examples that $\mathbb{Z} \ncong 2\mathbb{Z}$ and that $\mathbb{R}^{>0} \ncong \mathbb{R}^{m\times n}$. *Why not?* In each case, we were considering only one of the (possibly many) homomorphisms. It is quite possible that we could find different homomorphisms that *would* be bijections, showing that $\mathbb{Z} \cong 2\mathbb{Z}$ and that $\mathbb{R}^{>0} \cong \mathbb{R}^{m\times n}$. The first assertion is in fact true, while the second is not; you will explain why in the exercises.

Properties of group homomorphism

We turn now to three important properties of group *homomorphism*. For the rest of this section, we assume that $(G, \times)$ and $(H, \circ)$ are groups. Notice that the operations are both "multiplicative".

We still haven't explored the relationship between group homomorphisms and monoid homomorphisms. If a group homomorphism has fewer criteria, can it actually guarantee more structure? Theorem 4.9 answers in the affirmative.

Theorem 4.9. Let $f : G \to H$ be a homomorphism of groups. Denote the identity of G by e_G, and the identity of H by e_H. Then f
preserves identities: $f(e_G) = e_H$; and
preserves inverses: for every $x \in G$, $f\left(x^{-1}\right) = f(x)^{-1}$.

Read the proof below carefully, and identify precisely why this theorem holds for groups, but not for monoids.

Proof. *That f preserves identities:* Let $x \in G$, and $y = f(x)$. By the property of homomorphisms,

$$e_H y = y = f(x) = f(e_G x) = f(e_G) f(x) = f(e_G) y.$$

By the transitive property of equality,

$$e_H y = f(e_G) y.$$

Multiply both sides of the equation *on the right* by y^{-1} to obtain

$$e_H = f(e_G).$$

This shows that f, an arbitrary homomorphism of arbitrary groups, maps the identity of the domain to the identity of the range.

That f preserves inverses: Let $x \in G$. By the property of homomorphisms and by the fact that f preserves identity,

$$e_H = f(e_G) = f\left(x \cdot x^{-1}\right) = f(x) \cdot f\left(x^{-1}\right).$$

Thus

$$e_H = f(x) \cdot f\left(x^{-1}\right).$$

Pay careful attention to what this equation says! "The product of $f(x)$ and $f\left(x^{-1}\right)$ is the identity," which means that those two elements must be inverses! Hence, $f\left(x^{-1}\right)$ is the inverse of $f(x)$, which we write as

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

□

The trick, then, is that the property of inverses guaranteed to groups allows us to do more than we can do in a monoid. In this case, *more* structure in the group led to *fewer* conditions for equivalence. This is not true in general; we we discuss rings, we will see that more structure can lead to more conditions.

If homomorphisms preserve the inverse after all, it makes sense that "the inverse of the image is the image of the inverse." Corollary 4.10 affirms this.

Corollary 4.10. Let $f : G \to H$ be a homomorphism of groups. Then $f\left(x^{-1}\right)^{-1} = f(x)$ for every $x \in G$.

Proof. You do it! See Exercise 4.23. □

It will probably not surprise you that homomorphisms preserve powers of an element.

Theorem 4.11. Let $f : G \to H$ be a homomorphism of groups. Then f preserves powers of elements of G. That is, if $f(g) = h$, then $f(g^n) = f(g)^n = h^n$.

Proof. You do it! See Exercise 4.28. □

Naturally, if homomorphisms preserve powers of an element, they must also preserve cyclic groups.

Corollary 4.12. Let $f : G \to H$ be a homomorphism of groups. If $G = \langle g \rangle$ is a cyclic group, then $f(g)$ determines f completely. In other words, the image $f(G)$ is a cyclic group, and $f(G) = \langle f(g) \rangle$.

Proof. Assume that $G = \langle g \rangle$; that is, G is cyclic. We have to show that two sets are equal. By definition, for any $x \in G$ we can find $n \in \mathbb{Z}$ such that $x = g^n$.

First we show that $f(G) \subseteq \langle f(g) \rangle$. Let $y \in f(G)$ and choose $x \in G$ such that $y = f(x)$. Since G is a cyclic group generated by g, we can choose $n \in \mathbb{Z}$ such that $x = g^n$. By substitution and Theorem 4.11, $y = f(x) = f(g^n) = f(g)^n$. By definition, $y \in \langle f(g) \rangle$. Since y was arbitrary in $f(G)$, $f(G) \subseteq \langle f(g) \rangle$.

Now we show that $f(G) \supseteq \langle f(g) \rangle$. Let $y \in \langle f(g) \rangle$, and choose $n \in \mathbb{Z}$ such that $y = f(g)^n$. By Theorem 4.11, $y = f(g^n)$. Since $g^n \in G$, $f(g^n) \in f(G)$, so $y \in f(G)$. Since y was arbitrary in $\langle f(g) \rangle$, $f(G) \supseteq \langle f(g) \rangle$.

We have shown that $f(G) \subseteq \langle f(g) \rangle$ and $f(G) \supseteq \langle f(g) \rangle$. By equality of sets, $f(G) = \langle f(g) \rangle$. □

The final property of homomorphism that we check here is an important algebraic property of functions; it should remind you of a topic in Section 0.3. It will prove important in subsequent sections and chapters.

Definition 4.13. Let G and H be groups, and $f : G \to H$ a homomorphism. Let

$$K = \{g \in G : f(g) = e_H\};$$

that is, K is the set of all elements of G that f maps to the identity of H. We call K the **kernel** of f, written $\ker f$.

Theorem 4.14. Let $f : G \to H$ be a homomorphism of groups. Then $\ker f \triangleleft G$.

Proof. You do it! See Exercise 4.25. □

Exercises.

Exercise 4.15.
(a) Show that $f : \mathbb{Z} \to 2\mathbb{Z}$ by $f(x) = 2x$ is an isomorphism. Hence $\mathbb{Z} \cong 2\mathbb{Z}$.
(b) Show that $\mathbb{Z} \cong n\mathbb{Z}$ for every nonzero integer n.

Exercise 4.16. Let $n \geq 1$ and $f : \mathbb{Z} \longrightarrow \mathbb{Z}_n$ by $f(a) = [a]_n$.
(a) Show that f is a homomorphism.
(b) Explain why f cannot possibly be an isomorphism.
(c) Determine $\ker f$. (It might help to use a specific value of n first.)
(d) 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.)

Exercise 4.17. Show that $\mathbb{Z}_2$ is isomorphic to the group of order two from Example 2.11 on page 78. *Caution!* Remember to denote the operations properly: $\mathbb{Z}_2$ is additive, but we used $\circ$ for the operation of the group of order two.

Exercise 4.18. Show that $\mathbb{Z}_2$ is isomorphic to the Boolean xor group of Exercise 2.29 on page 82. *Caution!* Remember to denote the operation in the Boolean xor group correctly.

Exercise 4.19. Show that $\mathbb{Z}_n \cong \Omega_n$ for $n \in \mathbb{N}^+$.

Exercise 4.20. Suppose we try to define $f : Q_8 \longrightarrow \Omega_4$ by $f(\mathbf{i}) = f(\mathbf{j}) = f(\mathbf{k}) = i$, and $f(\mathbf{xy}) = f(\mathbf{x}) f(\mathbf{y})$ for all other $\mathbf{x}, \mathbf{y} \in Q_8$. Show that f is *not* a homomorphism.

Exercise 4.21. Show that $\mathbb{Z}$ is isomorphic to $\mathbb{Z}_0$. (Because of this, people generally don't pay attention to $\mathbb{Z}_0$. See also Exercise 3.102 on page 139.)

Exercise 4.22. Recall the subgroup L of $\mathbb{R}^2$ from Exercises 3.18 on page 114, 3.44 on page 122, and 3.83 on page 134. Show that $L \cong \mathbb{R}$.

Exercise 4.23. Prove Corollary 4.10.

Exercise 4.24. Suppose f is an isomorphism. How many elements does $\ker f$ contain?

Claim: $\ker\varphi \triangleleft G$.

Proof:

1. By _____, it suffices to show that for any $g \in G$, $\ker\varphi = g(\ker\varphi)g^{-1}$. So, let $g \in$_____.
2. First we show that $(\ker\varphi) \supseteq g(\ker\varphi)g^{-1}$. Let $x \in g(\ker\varphi)g^{-1}$.
   (a) By _____, there exists $k \in \ker\varphi$ such that $x = gkg^{-1}$.
   (b) By _____, $\varphi(x) = \varphi\left(gkg^{-1}\right)$.
   (c) By _____, $\varphi(x) = \varphi(g)\varphi(k)\varphi(g)^{-1}$.
   (d) By _____, $\varphi(x) = \varphi(g)e_H\varphi(g)^{-1}$.
   (e) By _____, $\varphi(x) = e_H$.
   (f) By definition of the kernel, _____.
   (g) Since _____, $g(\ker\varphi)g^{-1} \subseteq \ker\varphi$.
3. Now we show the converse; that is, _____. Let $k \in \ker\varphi$.
   (a) Let $x = g^{-1}kg$. Notice that if $x \in \ker\varphi$, then we would have what we want, since in this case _____.
   (b) In fact, $x \in \ker\varphi$. After all, _____.
   (c) Since _____, $\ker\varphi \subseteq g(\ker\varphi)g^{-1}$.
4. By _____, $\ker\varphi = g(\ker\varphi)g^{-1}$.

Figure 4.1. Material for Exercise 4.25

Exercise 4.25. Let G and H be groups, and $\varphi : G \to H$ a homomorphism.
(a) Show that $\ker\varphi < G$.
(b) Fill in each blank of Figure 4.1 with the appropriate justification or statement.

Exercise 4.26. Let φ be a homomorphism from a finite group G to a group H. Recall from Exercise 4.25 that $\ker\varphi \triangleleft G$. Explain why $|\ker\varphi| \cdot |\varphi(G)| = |G|$. (This is sometimes called **the Homomorphism Theorem**.)

Exercise 4.27. Let $f : G \to H$ be an isomorphism. Isomorphisms are by definition one-to-one functions, so f has an inverse function f^{-1}. Show that $f^{-1} : H \to G$ is also an isomorphism.

Exercise 4.28. Prove Theorem 4.11.

Exercise 4.29. Let $f : G \to H$ be a homomorphism of groups. Assume that G is abelian.
(a) Show that $f(G)$ is abelian.
(b) Must H be abelian? Explain why or why not.

Exercise 4.30. Let $f : G \to H$ be a homomorphism of groups. Let $A < G$. Show that $f(A) < H$.

Exercise 4.31. Let $f : G \to H$ be a homomorphism of groups. Let $A \triangleleft G$.
(a) Show that $f(A) \triangleleft f(G)$.
(b) Do you think that $f(A) \triangleleft H$? Justify your answer.

Exercise 4.32. Show that if G is a group, then $G/\{e\} \cong G$ and $G/G \cong \{e\}$.

Exercise 4.33. Recall the orthogonal group and the special orthogonal group from Exercise 3.24. 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 $\ker \varphi = \mathrm{SO}(n)$.

Exercise 4.34. In Chapter 1, the definition of an isomorphism for *monoids* required that the function map the identity to the identity (Definition 1.31 on page 57). By contrast, Theorem 4.9 shows that the preservation of the operation guarantees that a group homomorphism maps the identity to the identity, so we don't need to require this in the definition of an isomorphism for *groups* (Definition 4.6).

The difference between a group and a monoid is the existence of an inverse. Give a concrete example that shows where Theorem 4.9 is false for monoids.

4.2: Consequences of isomorphism

Throughout this section, $(G, \times)$ and $(H, \circ)$ are groups.

The purpose of this section is to show why we use the name *isomorphism:* if two groups are isomorphic, then they are indistinguishable *as groups*. The elements of the sets are different, and the operation may be defined differently, but as groups the two are identical. Suppose that two groups G and H are isomorphic. We will show that

- isomorphism is an equivalence relation;
- G is abelian iff H is abelian;
- G is cyclic iff H is cyclic;
- every subgroup A of G corresponds to a subgroup A' of H (in particular, if A is of order n, so is A');
- every normal subgroup N of G corresponds to a normal subgroup N' of H;
- the quotient group G/N corresponds to a quotient group H/N'.

All of these depend on the existence of an isomorphism $f : G \to H$. In particular, uniqueness is guaranteed only for any one isomorphism; if two different isomorphisms f, f' exist between G and H, then a subgroup A of G may well correspond to two distinct subgroups B and B' of H.

Isomorphism is an equivalence relation

The fact that isomorphism is an equivalence relation will prove helpful with the equivalence properties; for example, "G is cyclic iff H is cyclic." So, we start with that one first.

Theorem 4.35. Isomorphism is an equivalence relation. That is, $\cong$ satisfies the reflexive, symmetric, and transitive properties.

Proof. First we show that $\cong$ is reflexive. Let G be any group, and let ι be the identity homomorphism from Example 4.3. We showed in Example 4.7 that ι is an isomorphism. Since $\iota : G \to G$, $G \cong G$. Since G was an arbitrary group, $\cong$ is reflexive.

Next, we show that $\cong$ is symmetric. Let G, H be groups and assume that $G \cong H$. By definition, there exists an isomorphism $f : G \to H$. By Exercise 4.27, f^{-1} is also a isomorphism. Hence $H \cong G$.

Finally, we show that $\cong$ is transitive. Let G, H, K be groups and assume that $G \cong H$ and $H \cong K$. By definition, there exist isomorphisms $f : G \to H$ and $g : H \to K$. Define $h : G \to K$ by

$$h(x) = g(f(x)).$$

We claim that h is an isomorphism. We show each requirement in turn:

That h is a homomorphism, let $x, y \in G$. By definition of h, $h(x \cdot y) = g(f(x \cdot y))$. Applying the fact that g and f are both homomorphisms,

$$h(x \cdot y) = g(f(x \cdot y)) = g(f(x) \cdot f(y)) = g(f(x)) \cdot g(f(y)) = h(x) \cdot h(y).$$

Thus h is a homomorphism.

That h is one-to-one, let $x, y \in G$ and assume that $h(x) = h(y)$. By definition of h,

$$g(f(x)) = g(f(y)).$$

By hypothesis, g is an isomorphism, so by definition it is one-to-one, so if its outputs are equal, so are its inputs. In other words,

$$f(x) = f(y).$$

Similarly, f is an isomorphism, so $x = y$. Since x and y were arbitrary in G, h is one-to-one.

That h is onto, let $z \in K$. We claim that there exists $x \in G$ such that $h(x) = z$. Since g is an isomorphism, it is by definition onto, so there exists $y \in H$ such that $g(y) = z$. Since f is an isomorphism, there exists $x \in G$ such that $f(x) = y$. Putting this together with the definition of h, we see that

$$z = g(y) = g(f(x)) = h(x).$$

Since z was arbitrary in K, h is onto.

We have shown that h is a one-to-one, onto homomorphism. Thus h is an isomorphism, and $G \cong K$. □

Isomorphism preserves basic properties of groups

We now show that isomorphism preserves two basic properties of groups that we introduced in Chapter 2: abelian and commutative. Both proofs make use of the fact that isomorphism is an equivalence relation; in particular, that the relation is symmetric.

Theorem 4.36. Suppose that $G \cong H$. Then G is abelian iff H is abelian.

Proof. Let $f : G \to H$ be an isomorphism. Assume that G is abelian. We must show that H is abelian. By Exercise 4.29, $f(G)$ is abelian. Since f is an isomorphism, and therefore onto, $f(G) = H$. Hence H is abelian.

We turn to the converse. Assume that H is abelian. Since isomorphism is symmetric, $H \cong G$. Along with the above argument, this implies that if H is abelian, then G is, too.

Hence, G is abelian iff H is abelian. □

Theorem 4.37. Suppose $G \cong H$. Then G is cyclic iff H is cyclic.

Proof. Let $f : G \to H$ be an isomorphism. Assume that G is cyclic. We must show that H is cyclic; that is, we must show that every element of H is generated by a fixed element of H.

Since G is cyclic, by definition $G = \langle g \rangle$ for some $g \in G$. Let $h = f(g)$; then $h \in H$. We claim that $H = \langle h \rangle$.

Let $x \in H$. Since f is an isomorphism, it is onto, so there exists $a \in G$ such that $f(a) = x$. Since G is cyclic, there exists $n \in \mathbb{Z}$ such that $a = g^n$. By Theorem 4.11,

$$x = f(a) = f(g^n) = f(g)^n = h^n.$$

Since x was an arbitrary element of H and x is generated by h, all elements of H are generated by h. Hence $H = \langle h \rangle$ is cyclic.

Since isomorphism is symmetric, $H \cong G$. Along with the above argument, this implies that if H is cyclic, then G is, too.

Hence, G is cyclic iff H is cyclic. □

Isomorphism preserves the structure of subgroups

Theorem 4.38. Suppose $G \cong H$. Every subgroup A of G is isomorphic to a subgroup B of H. Moreover, each of the following holds.
(A) $|A| = |B|$.
(B) A is normal iff B is normal.

Proof. Let $f : G \to H$ be an isomorphism. Let A be a subgroup of G. By Exercise 4.30, $f(A) < H$.

We claim that f is one-to-one and onto from A to $f(A)$. Onto is immediate from the definition of $f(A)$. The one-to-one property holds because f is one-to-one in G and $A \subseteq G$. We have shown that $f(A) < H$ and that f is one-to-one and onto from A to $f(A)$. Hence $A \cong f(A)$.

Claim (A) follows from the fact that f is a bijection: this is the definition of when two sets have equal size.

For claim (B), assume $A \triangleleft G$. We want to show that $B \triangleleft H$; that is, $xB = Bx$ for every $x \in H$. Let $x \in H$ and $y \in B$; since f is an isomorphism, it is onto, so $f(g) = x$ and $f(a) = y$ for some $g \in G$ and some $a \in A$. By substitution and the homomorphism property,

$$xy = f(g)f(a) = f(ga).$$

Since $A \triangleleft G$, $gA = Ag$, so there exists $a' \in A$ such that $ga = a'g$. Let $y' = f(a')$. By substitution and the homomorphism property,

$$xy = f(a'g) = f(a')f(g) = y'x.$$

By definition and substitution, we have $y' = f(a') \in f(A) = B$. We conclude that, $xy = y'x \in Bx$.

We have shown that for arbitrary $x \in H$ and arbitrary $y \in B$, there exists $y' \in B$ such that $xy = y'x$. Hence $xB \subseteq Bx$. A similar argument shows that $xB \supseteq Bx$, so $xB = Bx$. This is the definition of a normal subgroup, so $B \triangleleft H$.

Since isomorphism is symmetric, $B \cong A$. Along with the above argument, this implies that if $B \triangleleft H$, then $A \triangleleft G$, as well.

Hence, A is normal iff B is normal. □

Theorem 4.39. Suppose $G \cong H$ as groups. Every quotient group of G is isomorphic to a quotient group of H.

We use Lemma 3.37(CE3) on page 121 on coset equality heavily in this proof; you may want to go back and review it.

Proof. Let $f : G \to H$ be an isomorphism. Consider an arbitrary quotient group of G defined as G/A, where $A \triangleleft G$. Let $B = f(A)$; by Theorem 4.38 $B \triangleleft H$, so H/B is a quotient group. We want to show that $G/A \cong H/B$.

To that end, define a new function $f_A : G/A \to H/B$ by

$$f_A(X) = f(g)B \qquad \text{where} \qquad X = gA \in G/A.$$

Keep in mind that f_A maps cosets to cosets, using the relation f from group elements to group elements.

We claim that f_A is an isomorphism. You probably expect that we "only" have to show that f_A is a bijection and a homomorphism, but this is not true. We have to show first that f_A is well-defined. Do you remember what this means? If not, reread page 126. Once you understand the definition, ask yourself, why do we have to show f_A is well-defined?

Just we must define the operation for cosets to give the same result regardless of two cosets' representation, a function on cosets must give the same result regardless of that coset's representation. Let X be any coset in G/A. It is usually the case that X can have more than one representation; that is, we can find $g \neq \widehat{g}$ where $X = gA = \widehat{g}A$. For example, suppose you want to build a function from $\mathbb{Z}_5$ to another set. Suppose that we want $f([2]) = x$. Recall that in $\mathbb{Z}_5$, $\cdots = [-3] = [2] = [7] = [12] = \cdots$. If f is defined in such a way that we would think $f([-3]) \neq x$, we would have a problem, since we need to ensure that $f([-3]) = f([2])$! For another example, consider D_3. We know that $\varphi A_3 = (\rho\varphi) A_3$, even though $\varphi \neq \rho\varphi$; see Example 3.68 on page 129. If $f(g) \neq f(\widehat{g})$, then $f_A(X)$ would have more than one possible value, since

$$f_A(X) = f_A(gA) = f(g) \neq f(\widehat{g}) = f_A(\widehat{g}A) = f(X).$$

In other words, f_A would not be a function, since at least one element of the domain (X) would correspond to at least two elements of the range ($f(g)$ and $f(\widehat{g})$). See Figure 4.2. A homomorphism must first be a function, so if f_A is not even a function, then it is not well-defined.

That f_A is well-defined: Let $X \in G/A$ and consider two representations g_1A and g_2A of X. Let $Y_1 = f_A(g_1A)$ and $Y_2 = f_A(g_2A)$. By definition of f_A,

$$Y_1 = f(g_1)B \quad \text{and} \quad Y_2 = f(g_2)B.$$

To show that f_A is well-defined, we must show that $Y_1 = Y_2$. By hypothesis, $g_1A = g_2A$. Lemma 3.37(CE3) implies that $g_2^{-1}g_1 \in A$. Recall that $f(A) = B$; by definition of the image, $f\left(g_2^{-1}g_1\right) \in B$. The homomorphism property implies that

$$f(g_2)^{-1} f(g_1) = f\left(g_2^{-1}\right) f(g_1) = f\left(g_2^{-1}g_1\right) \in B.$$

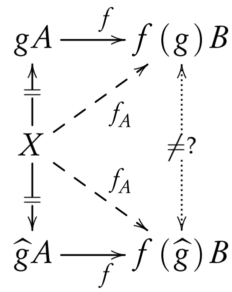


Figure 4.2. When defining a mapping whose domain is a quotient group, we must be careful to ensure that a coset with different representations has the same value. In the diagram above, X has the two representations gA and $\widehat{g}A$, and f_A is defined using f. In this case, is $f(g) = f(\widehat{g})$? If not, then $f_A(X)$ would have two different values, and f_A would not be a function.

Lemma 3.37(CE3) again implies that $f(g_1)B = f(g_2)B$, or $Y_1 = Y_2$, so there is no ambiguity in the definition of f_A as the image of X in H/B; the function is well-defined.

That f_A is a homomorphism: Let $X, Y \in G/A$ and write $X = g_1A$ and $Y = g_2A$ for appropriate $g_1, g_2 \in G$. Now

$$\begin{aligned} f_A(XY) &= f_A((g_1A)\cdot(g_2A)) && \text{(substitution)}\\ &= f_A(g_1g_2\cdot A) && \text{(coset multiplication in } G/A)\\ &= f(g_1g_2)B && \text{(definition of } f_A)\\ &= (f(g_1)f(g_2))\cdot B && \text{(homomorphism property)}\\ &= f(g_1)A'\cdot f(g_2)B && \text{(coset multiplication in } H/B)\\ &= f_A(g_1A)\cdot f_A(g_2A) && \text{(definition of } f_A)\\ &= f_A(X)\cdot f_A(Y) && \text{(substitution).} \end{aligned}$$

By definition, f_A is a homomorphism.

That f_A is one-to-one: Let $X, Y \in G/A$ and assume that $f_A(X) = f_A(Y)$. Let $g_1, g_2 \in G$ such that $X = g_1A$ and $Y = g_2A$. The definition of f_A implies that

$$f(g_1)B = f_A(X) = f_A(Y) = f(g_2)B,$$

so by Lemma 3.37(CE3) $f(g_2)^{-1}f(g_1) \in B$. Recall that $B = f(A)$, so there exists $a \in A$ such that $f(a) = f(g_2)^{-1}f(g_1)$. The homomorphism property implies that

$$f(a) = f\left(g_2^{-1}\right)f(g_1) = f\left(g_2^{-1}g_1\right).$$

Recall that f is an isomorphism, hence one-to-one. The definition of one-to-one implies that

$$g_2^{-1}g_1 = a \in A.$$

Applying Lemma 3.37(CE3) again gives us $g_1A = g_2A$, and

$$X = g_1A = g_2A = Y.$$

We took arbitrary $X, Y \in G/A$ and showed that if $f_A(X) = f_A(Y)$, then $X = Y$. It follows that f_A is one-to-one.

That f_A is onto: You do it! See Exercise 4.40. □

Exercises.

Exercise 4.40. Show that the function f_A defined in the proof of Theorem 4.39 is onto.

Exercise 4.41. Recall from Exercise 2.89 on page 109 that $\langle \mathbf{i} \rangle$ is a cyclic group of Q_8.
(a) Show that $\langle \mathbf{i} \rangle \cong \mathbb{Z}_4$ by giving an explicit isomorphism.
(b) Let A be the proper subgroup of $\langle \mathbf{i} \rangle$. Find the corresponding subgroup of $\mathbb{Z}_4$.
(c) Use the proof of Theorem 4.39 to determine the quotient group of $\mathbb{Z}_4$ to which $\langle \mathbf{i} \rangle / A$ is isomorphic.

Exercise 4.42. Recall from Exercise 4.22 on page 146 that the set

$$L = \left\{ x \in \mathbb{R}^2 : \ x = (a, a) \ \exists a \in \mathbb{R} \right\}$$

defined in Exercise 3.18 on page 114 is isomorphic to $\mathbb{R}$.
(a) Show that $\mathbb{Z} \triangleleft \mathbb{R}$.
(b) Give the precise definition of $\mathbb{R}/\mathbb{Z}$.
(c) Explain why we can think of $\mathbb{R}/\mathbb{Z}$ as the set of classes $[a]$ such that $a \in [0, 1)$. Choose one such $[a]$ and describe the elements of this class.
(d) Find the subgroup H of L that corresponds to $\mathbb{Z} < \mathbb{R}$. What do this section's theorems imply that you can conclude about H and L/H?
(e) Use the homomorphism f_A defined in the proof of Theorem 4.39 to find the images $f_{\mathbb{Z}}(\mathbb{Z})$ and $f_{\mathbb{Z}}(\pi + \mathbb{Z})$.
(f) Use the answer to (c) to describe L/H intuitively. Choose an element of L/H and describe the elements of this class.

4.3: The Isomorphism Theorem

In this section, we identify an important relationship between a subgroup $A < G$ that has a special relationship to a homomorphism, and the image of the quotient group $f(G/A)$. First, an example.

Motivating example

Example 4.43. Recall $A_3 = \{\iota, \rho, \rho^2\} \triangleleft D_3$ from Example 3.68. 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 (Exercise **??** on page ??) 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 $\mathbb{Z}_2$:

$$(\mu \circ \eta)(\iota) = \mu(\eta(\iota)) = \mu(A_3) = 0;$$
$$(\mu \circ \eta)(\rho) = \mu(\eta(\rho)) = \mu(A_3) = 0;$$
$$(\mu \circ \eta)\left(\rho^2\right) = \mu\left(\eta\left(\rho^2\right)\right) = \mu(A_3) = 0;$$
$$(\mu \circ \eta)(\varphi) = \mu(\eta(\varphi)) = \mu(\varphi A_3) = 1;$$
$$(\mu \circ \eta)(\rho\varphi) = \mu(\eta(\rho\varphi)) = \mu(\varphi A_3) = 1;$$
$$(\mu \circ \eta)\left(\rho^2\varphi\right) = \mu\left(\eta\left(\rho^2\varphi\right)\right) = \mu(\varphi A_3) = 1.$$

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
  - $f(g) = e_H$ for all $g \in A$, and
  - $f(g) \neq e_H$ for all $g \notin A$.

If we can do this, then $H \cong G/A$, and as we saw in Section 4.2 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 realize 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 4.44. Let G be a group and $A \triangleleft G$. The function $\eta : G \to G/A$ by

$$\eta(g) = gA$$

is a homomorphism.

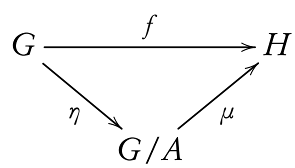
Proof. You do it! See Exercise 4.47. □

Definition 4.45. We call the homomorphism η of Lemma 4.44 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 it was the kernel of f. We use this to formalize the observation of Example 4.43.

Theorem 4.46 (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 4.46 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 3.37(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 $f(a) = e_H$. Substitution gives

$$\mu(g'A) = f(g) \cdot e_H = f(g) = \mu(gA).$$

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

$$\begin{aligned}
f(g^{-1}g') &= f(g^{-1})f(g') && \text{(homomorphism)}\\
&= f(g)^{-1}f(g') && \text{(homomorphism)}\\
&= \mu(gA)^{-1}\mu(g'A) && \text{(definition of } \mu)\\
&= \mu(X)^{-1}\mu(Y) && \text{(substitution)}\\
&= \mu(Y)^{-1}\mu(Y) && \text{(substitution)}\\
&= e_H, && \text{(inverses)}
\end{aligned}$$

so $g^{-1}g' \in \ker f$. By hypothesis, $\ker f = A$, so $g^{-1}g' \in A$. Lemma 3.37(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 $g \in G$,

$$(\mu \circ \eta)(g) = \mu(\eta(g)) = \mu(gA) = f(g).$$

□

Exercises

Exercise 4.47. Prove Lemma 4.44.

Exercise 4.48. Use Exercise 4.33 to explain why $\Omega_2 \cong \mathrm{O}(n)/\mathrm{SO}(n)$.

Exercise 4.49. Recall the normal subgroup L of $\mathbb{R}^2$ from Exercises 3.18, 3.44, and 3.83 on pages 114, 122, and 134, respectively. In Exercise 4.22 on page 146 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 $\mathbb{R}$.

Let G and H be groups, and $A \triangleleft G$.
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 _____, $f(A) = e_H$.
   (c) Let $a \in A$. By definition of the natural homomorphism, $\eta(a) =$ _____.
   (d) By _____, $f(\eta(a)) = e_H$.
   (e) By _____, $\varphi(a) = e_H$.
   (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 _____, $f(\eta(g)) \neq e_H$.
   (c) By _____, $\varphi(g) \neq e_H$.
   (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 4.3. Material for Exercise 4.52

Exercise 4.50. Recall the normal subgroup $\langle -1 \rangle$ of Q_8 from Exercises 2.87 on page 108 and 3.80 on page 132.

(a) Use Lagrange's Theorem to explain why $Q_8 / \langle -1 \rangle$ has order 4.
(b) We know from Exercise 2.19 on page 81 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 $Q_8 / \langle -1 \rangle$.

Exercise 4.51. Recall the kernel of a monoid homomorphism from Exercise **??** on page ??, 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 $y^{-1}z = x$.
(c) Explain how this shows that Exercise **??** "lays the groundwork" for a "monoid generalization" of the Isomorphism Theorem.
(d) Formulate and prove a "Monoid Isomorphism Theorem."

Exercise 4.52. Fill in each blank of Figure 4.3 with the appropriate justification or statement.

4.4: Automorphisms and groups of automorphisms

In this section, we use isomorphisms to build a new kind of group, useful for analyzing roots of polynomial equations. We will discuss the applications of these groups in Chapter 9, but they are of independent interest, as well.

Definition 4.53. Let G be a group. If $f : G \to G$ is an isomorphism, then we call f an **automorphism**.

An automorphism[21] is an isomorphism whose domain and range are the same set. Thus, to show that some function f is an automorphism, you must show first that the domain and the range of f are the same set. Afterwards, you show that f satisfies the homomorphism property, and then that it is both one-to-one and onto.

Example 4.54.

(a) An easy automorphism for any group G is the identity isomorphism $\iota(g) = g$:
- its range is by definition G;
- it is a *homomorphism* because $\iota(g \cdot g') = g \cdot g' = \iota(g) \cdot \iota(g')$;
- it is *one-to-one* because $\iota(g) = \iota(g')$ implies (by evaluation of the function) that $g = g'$; and
- it is *onto* because for any $g \in G$ we have $\iota(g) = g$.

(b) An automorphism in $(\mathbb{Z}, +)$ is $f(x) = -x$:
- its range is $\mathbb{Z}$ because of closure;
- it is a *homomorphism* because $f(x + y) = -(x + y) = -x - y = f(x) + f(y)$;
- it is *one-to-one* because $f(x) = f(y)$ implies that $-x = -y$, so $x = y$; and
- it is *onto* because for any $x \in \mathbb{Z}$ we have $f(-x) = x$.

(c) An automorphism in D_3 is $f(x) = \rho^2 x \rho$:
- its range is D_3 because of closure;
- it is a *homomorphism* because $f(xy) = \rho^2 (xy) \rho = \rho^2 (x \cdot \iota \cdot y) \rho = \rho^2 (x \cdot \rho^3 \cdot y) \rho = (\rho^2 x \rho) \cdot (\rho^2 y \rho) = f(x) \cdot f(y)$;
- it is *one-to-one* because $f(x) = f(y)$ implies that $\rho^2 x \rho = \rho^2 y \rho$, and multiplication on the left by ρ and on the right by ρ^2 gives us $x = y$; and
- it is *onto* because for any $y \in D_3$, choose $x = \rho y \rho^2$ and then $f(x) = \rho^2 (\rho y \rho^2) \rho = (\rho^2 \rho) \cdot y \cdot (\rho^2 \rho) = \iota \cdot y \cdot \iota = y$.

The automorphism of Example 4.54(c) generalizes to an important way. Recall the conjugation of one element of a group by another, introduced in Exercise 2.40 on page 83. By fixing the second element, we can turn this into a function on a group.

Definition 4.55. Let G be a group and $a \in G$. Define the function of **conjugation by** a to be $\operatorname{conj}_a(x) = a^{-1} x a$.

In Example 4.54(c), we had $a = \rho$ and $\operatorname{conj}_a(x) = a^{-1} x a = \rho^2 x \rho$.

You have already worked with conjugation in previous exercises, such as showing that it can provide an alternate definition of a normal subgroup (Exercises 2.40 on page 83 and 3.82 on page 133). Beyond that, conjugating a subgroup *always* produces another subgroup:

[21] The word comes Greek words that mean *self* and *shape*.

Lemma 4.56. Let G be a group, and $a \in G$. Then conj_a is an automorphism. Moreover, for any $H < G$,

$$\{\mathrm{conj}_a(h) : h \in H\} < G.$$

Proof. You do it! See Exercise 4.64. □

The subgroup $\{\mathrm{conj}_a(h) : h \in H\}$ is important enough to identify by a special name.

Definition 4.57. Suppose $H < G$, and $a \in G$. We say that $\{\mathrm{conj}_a(h) : h \in H\}$ is the **group of conjugations of H by a**, and denote it by $\mathrm{Conj}_a(H)$.

Conjugation of a subgroup H by an arbitrary $a \in G$ is *not* necessarily an automorphism; there can exist $H < G$ and $a \in G \backslash H$ such that $H \neq \{\mathrm{conj}_a(h) : h \in H\}$. On the other hand, if H is a *normal* subgroup of G, then we *do* have $H = \{\mathrm{conj}_a(h) : h \in H\}$; this property can act as an alternate definition of a normal subgroup. You will explore this in the exercises.

Now it is time to identify the new group that we promised at the beginning of the section.

The automorphism group

Notation 4.58. Write $\mathrm{Aut}(G)$ for the set of all automorphisms of G. We typically denote elements of $\mathrm{Aut}(G)$ by Greek letters ($\alpha, \beta, \ldots$), rather than Latin letters ($f, g, \ldots$).

Example 4.59. We compute $\mathrm{Aut}(\mathbb{Z}_4)$. Let $\alpha \in \mathrm{Aut}(\mathbb{Z}_4)$ be arbitrary; what do we know about α? By definition, its range is $\mathbb{Z}_4$, and by Theorem 4.9 on page 144 we know that $\alpha(0) = 0$. Aside from that, we consider all the possibilities that preserve the isomorphism properties.

Recall from Theorem 3.100 on page 139 that $\mathbb{Z}_4$ is a cyclic group; in fact $\mathbb{Z}_4 = \langle 1 \rangle$. Corollary 4.12 on page 145 tells us that $\alpha(1)$ will tell us everything we want to know about α. So, what can $\alpha(1)$ be?

Case 1. Can we have $\alpha(1) = 0$? If so, then $\alpha(1) = \alpha(0)$. This is not one-to-one, so we cannot have $\alpha(1) = 0$.

Case 2. Can we have $\alpha(1) = 1$? Certainly $\alpha(1) = 1$ if α is the identity homomorphism ι, so we can have $\alpha(1) = 1$.

Case 3. Can we have $\alpha(1) = 2$? If so, then the homomorphism property implies that

$$\alpha(2) = \alpha(1+1) = \alpha(1) + \alpha(1) = 4 = 0 = \alpha(0).$$

This is not one-to-one, so we cannot have $\alpha(1) = 2$.

Case 4. Can we have $\alpha(1) = 3$? If so, then the homomorphism property implies that

$$\alpha(2) = \alpha(1+1) = \alpha(1) + \alpha(1) = 3 + 3 = 6 = 2; \text{ and}$$
$$\alpha(3) = \alpha(2+1) = \alpha(2) + \alpha(1) = 2 + 3 = 5 = 1.$$

In this case, α is both one-to-one *and* onto. We were careful to observe the homomorphism property when determining α, so we know that α is a homomorphism. So we *can* have $\alpha(1) = 2$.

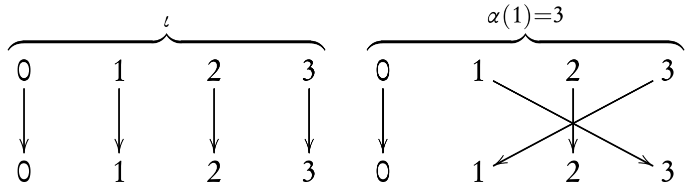


Figure 4.4. The elements of $\operatorname{Aut}(\mathbb{Z}_4)$.

We found only two possible elements of $\operatorname{Aut}(\mathbb{Z}_4)$: the identity automorphism and the automorphism determined by $\alpha(1) = 3$. Figure 4.4 illustrates the two mappings.

If $\operatorname{Aut}(\mathbb{Z}_4)$ were a group, then the fact that it contains only two elements would imply that $\operatorname{Aut}(\mathbb{Z}_4) \cong \mathbb{Z}_2$. But *is* it a group?

Lemma 4.60. For any group G, $\operatorname{Aut}(G)$ is a group under the operation of composition of functions.

On account of this lemma, we can justifiably refer to $\operatorname{Aut}(G)$ as **the automorphism group**.

Proof. Let G be any group. We show that $\operatorname{Aut}(G)$ satisfies each of the group properties from Definition 2.1.

(closed) Let $\alpha, \theta \in \operatorname{Aut}(G)$. We must show that $\alpha \circ \theta \in \operatorname{Aut}(G)$ as well:

- the domain and range of $\alpha \circ \theta$ are both G because the domain and range of both α and θ are both G;
- $\alpha \circ \theta$ is a *homomorphism* because for any $g, g' \in G$ we have,

$$\begin{aligned}(\alpha\circ\theta)(g\cdot g') &= \alpha\left(\theta\left(g\cdot g'\right)\right) && \text{(def. of comp.)}\\ &= \alpha\left(\theta(g)\cdot\theta\left(g'\right)\right) && (\theta \text{ a homom.})\\ &= \alpha(\theta(g))\cdot\alpha\left(\theta\left(g'\right)\right) && (\alpha \text{ a homom.})\\ &= (\alpha\circ\theta)(g)\cdot(\alpha\circ\theta)\left(g'\right); && \text{(def. of comp.)}\end{aligned}$$

- $\alpha \circ \theta$ is *one-to-one* because
  - if $(\alpha\circ\theta)(g) = (\alpha\circ\theta)(g')$, then by the definition of composition, $\alpha(\theta(g)) = \alpha(\theta(g'))$;
  - since α is one-to-one, $\theta(g) = \theta(g')$;
  - since θ is one-to-one, $g = g'$; and
- $\alpha \circ \theta$ is *onto* because for any $z \in G$,
  - α is onto, so there exists $y \in G$ such that $\alpha(y) = z$, and
  - θ is onto, so there exists $x \in G$ such that $\theta(x) = y$, so
  - $(\alpha\circ\theta)(x) = \alpha(\theta(x)) = \alpha(y) = z$.

We have shown that $\alpha \circ \theta$ satisfies the properties of an automorphism; hence, $\alpha \circ \theta \in \operatorname{Aut}(G)$, and $\operatorname{Aut}(G)$ is closed under the composition of functions.

(associative) The associative property is satisfied because the operation is composition of functions, which is associative.

(identity) Denote by ι the identity homomorphism; that is, $\iota(g) = g$ for all $g \in G$. We showed in Example 4.54(a) that ι is an automorphism, so $\iota \in \operatorname{Aut}(G)$. Let $\alpha \in \operatorname{Aut}(G)$; we claim that $\iota \circ \alpha = \alpha \circ \iota = \alpha$. Let $x \in G$ and write $y = \alpha(x)$. We have

$$(\iota \circ \alpha)(x) = \iota(\alpha(x)) = \iota(y) = y = \alpha(x),$$

and likewise $(\alpha \circ \iota)(x) = \alpha(x)$. Since x was arbitrary in G, we have $\iota \circ \alpha = \alpha \circ \iota = \alpha$.

(inverse) Let $\alpha \in \operatorname{Aut}(G)$. Since α is an automorphism, it is an isomorphism. You showed in Exercise 4.27 that α^{-1} is also an isomorphism. The domain and range of α are both G, so the domain and range of α^{-1} are also both G. Hence $\alpha^{-1} \in \operatorname{Aut}(G)$.

□

Since $\operatorname{Aut}(G)$ is a group, we can compute $\operatorname{Aut}(\operatorname{Aut}(G))$, and the same theory holds, so we can compute $\operatorname{Aut}(\operatorname{Aut}(\operatorname{Aut}(G)))$, and so forth. In the exercises, you will compute $\operatorname{Aut}(G)$ for some other groups.

Exercises.

Exercise 4.61. Show that $f(x) = x^2$ is an automorphism on the group $\left(\mathbb{R}^{>0}, \times\right)$, but not on the group $(\mathbb{R}, \times)$.

Exercise 4.62. Recall the subgroup $A_3 = \left\{\iota, \rho, \rho^2\right\}$ of D_3.
(a) List the elements of $\operatorname{Conj}_\rho(A_3)$.
(b) List the elements of $\operatorname{Conj}_\varphi(A_3)$.
(c) In both (a) and (b), we saw that $\operatorname{Conj}_a(A_3) = A_3$ for $a = \rho, \varphi$. This makes sense, since $A_3 \triangleleft D_3$. Find a subgroup K of D_3 and an element $a \in D_3$ where $\operatorname{Conj}_a(K) \neq K$.

Exercise 4.63. Let $H = \langle \mathbf{i} \rangle < Q_8$. List the elements of $\operatorname{Conj}_{\mathbf{j}}(H)$.

Exercise 4.64. Prove Lemma 4.56 on page 160 in two steps:
(a) Show first that conj_a is an automorphism.
(b) Show that $\{\operatorname{conj}_a(h) : h \in H\}$ is a group.

Exercise 4.65. Determine the automorphism group of $\mathbb{Z}_5$.

Exercise 4.66. Determine the automorphism group of D_3.

Chapter 5: Groups of permutations

This chapter introduces groups of permutations. Now, what is a permutation, and why are they so important?

Certain applications of mathematics involve the rearrangement of a list of n elements. It is common to refer to such rearrangements as *permutations*.

> **Definition 5.1.** A **list** is a sequence. Let V be any finite list. A **permutation** is a one-to-one function whose domain and range are both V.

We require V to be a list rather than a set because for a permutation, the order of the elements matters: 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.

Let's take a concrete example. Suppose you have a list of numbers, $(1,3,2,7)$, and you rearrange them by switching the first two entries in the list, $(3,1,2,7)$. The action of switching those first two numbers is a permutation. There is no doubt as to the outcome of the action, so this action is a function. Thus, permutations are special kinds of functions.

The importance of permutations is twofold. First, group theory is a pretty neat and useful thing in itself, and we will see in this chapter that all finite groups can be modeled by groups of permutations. Anything that can model every possible group is by that very fact important.

The second reason permutations are important has to do with the factorization of polynomials. The polynomial $x^4 - 1$ can be factored as

$$(x+1)(x-1)(x+i)(x-i),$$

but it can also be factored as

$$(x-1)(x+1)(x-i)(x+i).$$

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.

Section 5.1 introduces you to groups of permutations, while Section 5.2 describes a convenient way to write permutations. Sections 5.3 and 5.5 introduce you to two special classes of groups of permutation. The main goal of this chapter is to show that groups of permutations are, in some sense, "all there is" to group theory, which we accomplish in Section 5.4. We conclude with a great example of an application of symmetry groups in Section 5.6.

5.1: Permutations

In this first section, we consider some basic properties of permutations.

Permutations as functions

Example 5.2. 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	$\to$	2
2	$\to$	4
3	$\to$	3
4	$\to$	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 2.2.

Proposition 5.3. Let V be a set of n elements. The function $\iota : V \to V$ by $\iota(x) = x$ is a permutation on V. In addition, for any $\alpha \in S_n$, $\iota \circ \alpha = \alpha$ and $\alpha \circ \iota = \alpha$.

Proof. You do it! See Exercise 5.13. □

Permutations have a convenient property.

Lemma 5.4. The composition of two permutations is a permutation.

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. The definition of α and β imply that 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 5.2, we wrote a permutation as a piecewise function. This is burdensome; we would like a more efficient way to denote permutations.

Notation 5.5. 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 5.6. Recall V and π from Example 5.2. 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 5.2.

We can also use π to permute different lists, so long as the new lists have four elements:

$$\pi(3,2,1,4) = (4,3,1,2);$$
$$\pi(2,4,3,1) = (1,2,3,4);$$
$$\pi(a,b,c,d) = (d,a,c,b).$$

Groups of permutations

It comes as a pleasant revelation that sets of permutations form groups in a very natural way. In particular, consider the following set.

Definition 5.7. For $n \geq 2$, denote by S_n the set of all permutations of a list of n elements.

Example 5.8. For $n = 1,2,3$ we have

$$S_1 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$
$$S_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}$$
$$S_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}.$$

Is there some structure to S_n? By definition, a permutation is a one-to-one function. In Example 1.10 on page 51, we found that for any set, the set of functions on that set was a monoid under the operation of 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 5.9. For all $n \geq 2$ $(S_n, \circ)$ is a group.

Notation 5.10. Normally we just write S_n, understanding from context that the operation is composition of functions. It is common to refer to S_n as the **symmetric group** of n elements.

Proof. Let $n \geq 2$. We have to show that S_n satisfies the properties of a group under the operation of composition of functions. Proposition 5.3 tells us that the identity function acts as an identity in S_n, and Lemma 5.4 tells us that S_n is closed under composition.

We still have to show that S_n satisfies the inverse and associative properties. Let V be a finite list with n elements. The fact that $S_n \subseteq F_V$ implies that S_n satisfies the associative property. Let $\alpha \in S_n$. By definition of a permutation, α is one-to-one; since V is finite, α is onto. By Exercise 0.39 on page 16, α 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 $\iota(v) = v$ for every $v \in V$, we have shown that $\alpha^{-1} \circ \alpha = \alpha \circ \alpha^{-1} = \iota$. Again, Exercise 0.39 indicates that α^{-1} is 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. □

A final question: *how large is each S_n?* To answer this, we must count the number of permutations 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 5.11. For each $n \in \mathbb{N}^+$, $|S_n| = n!$

Exercises

Exercise 5.12. 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 $\alpha(6,3,5,2,1,4)$.

Exercise 5.13. Prove Proposition 5.3.

Exercise 5.14. How many elements are there of S_4?

Exercise 5.15. Find an explicit isomorphism from S_2 to $\mathbb{Z}_2$.

5.2: Cycle notation

Tabular notation of permutations is rather burdensome; a simpler notation is possible.

Cycles

Definition 5.16. 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

$$\iota = (1).$$

The fact that the permutation α moves the entry in position α_n to position α_1 is the reason this is called a *cycle;* applying it repeatedly cycles the list of elements around, and on the nth application the list returns to its original order.

Example 5.17. Recall π from Example 5.6. 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 we like. However, the convention is to start with the smallest position that changes. Since π moves elements out of position 1, we start with

$$\pi = (1\,?)\,.$$

The second entry in cycle notation tells us where π moves the element whose position is that of the first entry. The first entry indicates position 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 whose position is that of 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 whose position is that of 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 there is no need. Since we have now returned to the first element in the cycle, we close it:

$$\pi = (1\,2\,4)\,.$$

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.

Not all permutations can be written as one cycle.

Example 5.18. 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,

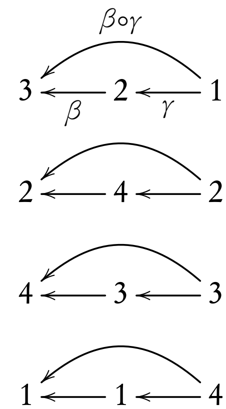


Figure 5.1. Diagram of how $\beta \circ \gamma$ modifies a list of four elements, for $\beta = (2\,3\,4)$ and $\gamma = (1\,2\,4)$.

and the element in the fourth position is moved to the fourth.

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 5.19. Consider the cycles

$$\beta = (2\,3\,4) \quad \text{and} \quad \gamma = (1\,2\,4).$$

What is the cycle notation for

$$\beta \circ \gamma = (2\,3\,4) \circ (1\,2\,4)?$$

Let's think about this. Since cycles represent permutations, and permutations are closed under composition, $\beta \circ \gamma$ must be 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

$$\beta \circ \gamma = (1\,?).$$

Where is this first element moved? Let's apply the definition of composition: $\beta \circ \gamma$ means, "first apply γ; then apply β." Figure 5.1 gives us the basic idea; we will refer to it throughout the example. Since γ moves an element in the first position to the second, and β moves an element in the second position to the third, it must be that $\beta \circ \gamma$ moves an element from the first position to the third. We see this in the top row of Figure 5.1. We now know that

$$\beta \circ \gamma = (1\,3\,?).$$

The next entry should tell us where $\beta \circ \gamma$ moves an element that starts in the third position. 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 see this in the *third* row of Figure 5.1. We now know that

$$\beta \circ \gamma = (1\,3\,4\,?) .$$

Time to look at elements in the fourth position, then. Since γ moves elements 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, we conclude that $\beta \circ \gamma$ 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 5.1.

Another phenomenon occurs when each permutation moves elements that the other does not.

Example 5.20. 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 5.21. We say that two cycles are **disjoint** if none of their entries are common.

Disjoint cycles enjoy an important property: their permutations commute under composition.

Lemma 5.22. 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 5.23. 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 5.22, 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 write cycles with smaller first entries before cycles with larger first entries. Thus, 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. We simplify the composition of cycles that are not disjoint, discarding all cycles of length 1.
3. The remaining cycles will be disjoint. From Lemma 5.22, we know that they commute; write them so that the first cycle's first entry is smallest, the second cycle's first entry is second-smallest, and so forth.

Example 5.24. We return to Example 5.18, 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 5.25. Every permutation can be written as a composition of cycles.

The proof is constructive; we build the cycle notation for the permutation.

Proof. Let π be a permutation; denote its domain by V. Without loss of generality, we write $V = (1,2,\ldots,n)$.

Let i_1 be the smallest element of V such that $\pi(i_1) \neq i_1$. Recall that the range of π has at most n elements, so the sequence $\pi(i_1)$, $\pi(\pi(i_1)) = \pi^2(i_1)$, $\ldots$ cannot continue indefinitely; eventually, we must have $\pi^{k+1}(i_1) = i_1$ for some $k \leq n$. Let

$$\alpha^{(1)} = \left(i_1\ \pi(i_1)\ \pi(\pi(i_1))\ \cdots\ \pi^k(i_1)\right).$$

Is there is some $i_2 \in V$ that is not stationary with respect to π and not an entry of $\alpha^{(1)}$? If so, then generate the cycle $\alpha^{(2)}$ by $\left(i_2\ \pi(i_2)\ \pi(\pi(i_2))\ \cdots\ \pi^\ell(i_2)\right)$, where, as before $\pi^{\ell+1}(i_2) = i_2$.

Repeat this process until every non-stationary element of V corresponds to a cycle, generating $\alpha^{(3)}$, $\ldots$, $\alpha^{(m)}$ for non-stationary $i_3 \notin \alpha^{(1)},\alpha^{(2)}$, $i_4 \notin \alpha^{(1)},\alpha^{(2)},\alpha^{(3)}$, and so on until $i_m \notin \alpha^{(1)},\ldots,\alpha^{(m-1)}$. Since the list is finite, this process will not continue indefinitely, and we have a finite list of cycles.

The remainder of the proof consists of two claims.

Claim 1: Each of the cycles we created is disjoint from any of the rest.

By way of contradiction, assume that two cycles $\alpha^{(i)}$ and $\alpha^{(j)}$ are not disjoint. By construction, the first elements of these cycles are different; let r be the first entry in $\alpha^{(j)}$ that also appears in $\alpha^{(i)}$. Let a be the entry that precedes r in $\alpha^{(i)}$, and b the entry that precedes r in $\alpha^{(j)}$. By construction, we have $\alpha(a) = r = \alpha(b)$. Since r is the *first* entry of each cycle that is the same, $a \neq b$. This contradicts the hypothesis that α is a permutation, as permutations are one-to-one. 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 $\alpha^{(j)}$. 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 $\alpha^{(j)}$.

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 $\alpha^{(k)}(i) = i$ for $k \neq j$. By Claim 1, both i and $\pi(i)$ appear in *only* one of the α. By substitution, the expression $\left(\alpha^{(1)}\alpha^{(2)}\cdots\alpha^{(m)}\right)(i)$ simplifies to

$$\begin{aligned}\left(\alpha^{(1)}\alpha^{(2)}\cdots\alpha^{(m)}\right)(i) &= \alpha^{(1)}\left(\alpha^{(2)}\left(\cdots\alpha^{(m-1)}\left(\alpha^{(m)}(i)\right)\right)\right)\\ &= \alpha^{(1)}\left(\alpha^{(2)}\left(\cdots\alpha^{(j-1)}\left(\alpha^{(j)}(i)\right)\right)\right)\\ &= \alpha^{(1)}\left(\alpha^{(2)}\left(\cdots\alpha^{(j-1)}(\pi(i))\right)\right)\\ &= \pi(i).\end{aligned}$$

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 5.26. 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 5.25 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 5.27. 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:

   $$\gamma = (\;1\quad ?\;).$$

   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:

$$\beta \circ \gamma = (2\,4\,3)\,.$$

Exercises.

Exercise 5.28. For the permutation

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 5 & 2 & 4 & 6 & 3 \end{pmatrix},$$

(a) Write α in cycle notation.
(b) Write α as a piecewise function.

Exercise 5.29. For the permutation

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

(a) Evaluate $\alpha\,(1,2,3,4)$.
(b) Evaluate $\alpha\,(1,4,3,2)$.
(c) Evaluate $\alpha\,(3,1,4,2)$.
(d) Write α in tabular notation.
(e) Write α as a piecewise function.

Exercise 5.30. Let $\alpha = (1\,2\,3\,4)$, $\beta = (1\,4\,3\,2)$, and $\gamma = (1\,3)$. Compute $\alpha \circ \beta$, $\alpha \circ \gamma$, $\beta \circ \gamma$, $\beta \circ \alpha$, $\gamma \circ \alpha$, $\gamma \circ \beta$, α^2, β^2, and γ^2. (Here $\alpha^2 = \alpha \circ \alpha$.) What are the inverses of α, β, and γ?

Exercise 5.31. Compute the order of

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

Exercise 5.32. Show that all the elements of S_3 can be written as compositions of the cycles $\alpha = (1\,2\,3)$ and $\beta = (2\,3)$.

Exercise 5.33. For α and β as defined in Exercise 5.32, 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 $n > 2$.)

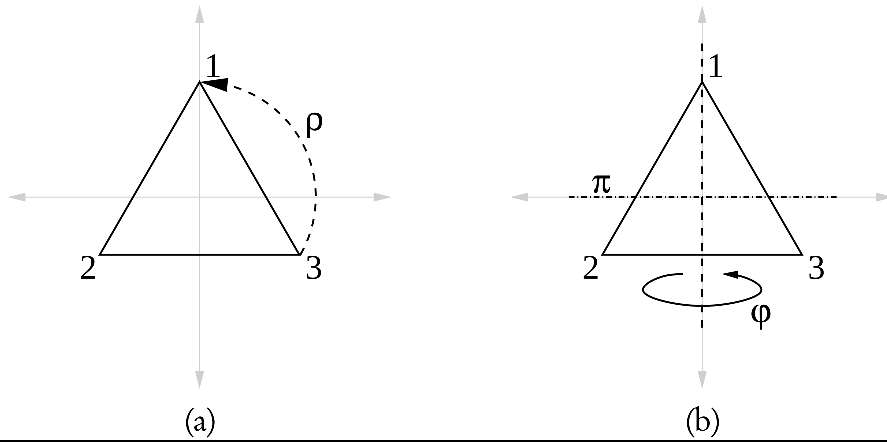


Figure 5.2. Rotation and reflection of an equilateral triangle centered at the origin

Exercise 5.34. Write the Cayley table for S_3.

Exercise 5.35. Show that $D_3 \cong S_3$ by showing that the function $f : D_3 \to S_3$ by $f\left(\rho^a \varphi^b\right) = \alpha^a \beta^b$ is an isomorphism.

Exercise 5.36. Identify at least one normal subgroup of S_3, and at least one subgroup that is not normal.

Exercise 5.37. List the elements of S_4 using cycle notation.

Exercise 5.38. Compute the cyclic subgroup of S_4 generated by $\alpha = (1\,3\,4\,2)$. Compare your answer to that of Exercise 5.31.

Exercise 5.39. Let $\alpha = (\alpha_1\ \alpha_2\ \cdots\ \alpha_m) \in S_n$. (Note $m \leq n$.) Show that we can write α^{-1} as

$$\beta = \left(\alpha_1\ \alpha_m\ \alpha_{m-1}\ \cdots\ \alpha_2\right).$$

For example, if $\alpha = (2\,3\,5\,6)$, $\alpha^{-1} = (2\,6\,5\,3)$.

5.3: Dihedral groups

In Section 2.2 we studied the symmetries of a triangle; we presented the group as the products of matrices ρ and φ, derived from the symmetries of *rotation* and *reflection about the y-axis*. Figure 5.2, a copy of Figure 28 on page 85, shows how ρ and φ correspond to the symmetries of an equilateral triangle centered at the origin. In Exercises 5.32–5.35 you showed that D_3 and S_3 are isomorphic.

From symmetries to permutations

We now turn to the symmetries of a regular n-sided polygon.

Definition 5.40. The **dihedral set** D_n is the set of symmetries of a regular polygon with n sides.

We have two goals in introducing the dihedral group: first, to give you another concrete and interesting group; and second, to serve as a bridge to Section 5.4. The next example starts starts us in that directions.

Example 5.41. Another way to represent the elements of D_3 is to consider how they re-arrange the vertices of the triangle. We can represent the vertices of a triangle as the list $V = (1,2,3)$. Application of ρ to the triangle moves

- vertex 1 to vertex 2;
- vertex 2 to vertex 3; and
- vertex 3 to vertex 1.

This is equivalent to the permutation $(1\,2\,3)$. Application of φ to the triangle moves

- vertex 1 to itself—that is, vertex 1 does not move;
- vertex 2 to vertex 3; and
- vertex 3 to vertex 2.

This is equivalent to the permutation $(2\,3)$.

In the context of the symmetries of the triangle, it looks as if ρ and φ correspond to $(1\,2\,3)$ and $(2\,3)$, respectively. Recall that ρ and φ generate all the symmetries of a triangle; likewise, these two cycles generate all the permutations of a list of three elements! (See Example 5.8 and Exercise **??** on page ??.)

We can do this with D_4 and S_4 as well.

Example 5.42. Using the tabular notation for permutations, we identify some elements of D_4, the set of symmetries of a square. As with the triangle, we can represent the vertices of a square as the list $V = (1,2,3,4)$. The identity symmetry ι, which moves the vertices back onto themselves, is thus the cycle (1). We also have a 90° rotation which moves vertex 1 to vertex 2, vertex 2 to vertex 3, and so forth. As a permutation, we can write that as

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

The other rotations are clearly powers of ρ. We can visualize three kinds of flips: one across the y-axis,

$$\varphi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \end{pmatrix}\begin{pmatrix} 3 & 4 \end{pmatrix};$$

one across the x-axis,

$$\vartheta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \end{pmatrix}\begin{pmatrix} 2 & 3 \end{pmatrix};$$

and one across a diagonal,

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

See Figure 5.3 on the next page. We can also imagine other diagonals; but they can be shown to be superfluous, just as we show shortly that ϑ and ψ are superfluous. There may be other symmetries of the square, but we'll stop here for the time being.

Is it possible to write ψ as a composition of φ and ρ? It turns out that $\psi = \varphi \circ \rho$. We can show

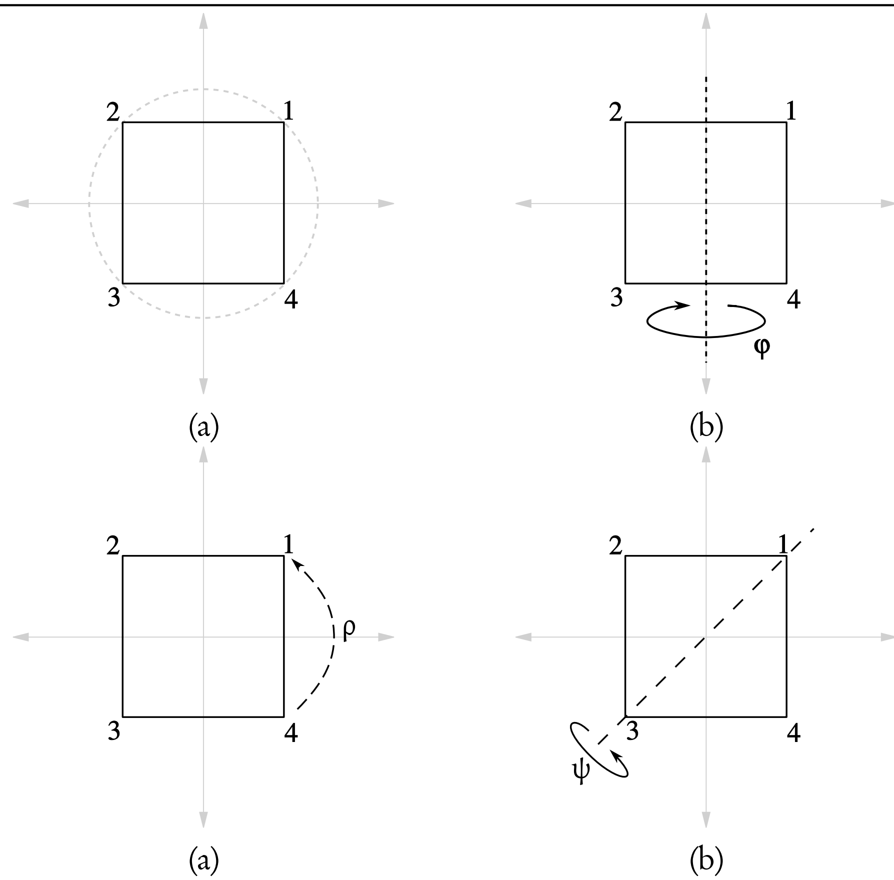


Figure 5.3. Rotation and reflection of a square centered at the origin

this by observing that

$$\varphi \circ \rho = (\ 1\ \ 2\ )(\ 3\ \ 4\ )(\ 1\ \ 2\ \ 3\ \ 4\ ) = (\ 2\ \ 4\ ) = \psi.$$

We can see this geometrically; see Figure 5.4. First, ρ moves $(1,2,3,4)$ to $(4,1,2,3)$. Subsequently, φ moves $(4,1,2,3)$ to $(1,4,3,2)$. Likewise, ψ would permute $(1,2,3,4)$ directly to $(1,4,3,2)$. Either way, we see that $\psi = \varphi \circ \rho$. A similar argument shows that $\vartheta = \varphi \circ \rho^2$, so it looks as if we need only φ and ρ to generate D_4.

Similar arguments verify that the reflection and the rotation have a property similar to that in S_3:

$$\varphi \circ \rho = \rho^3 \circ \varphi,$$

so unless there is some symmetry of the square that cannot be described by rotation or reflection on the y-axis, we can list all the elements of D_4 using a composition of some power of ρ after some power of φ. There are four unique 90° rotations and two unique reflections on the y-axis, implying that D_4 has at least eight elements:

$$D_4 \supseteq \left\{\iota, \rho, \rho^2, \rho^3, \varphi, \rho\varphi, \rho^2\varphi, \rho^3\varphi\right\}.$$

Can D_4 have other elements? There are in fact $|S_4| = 4! = 24$ possible permutations of the vertices, but are they all symmetries of a square? Consider the permutation from $(1,2,3,4)$ to $(2,1,3,4)$: in the basic square, the distance between vertices 1 and 3 is $\sqrt{2}$, but in the config-

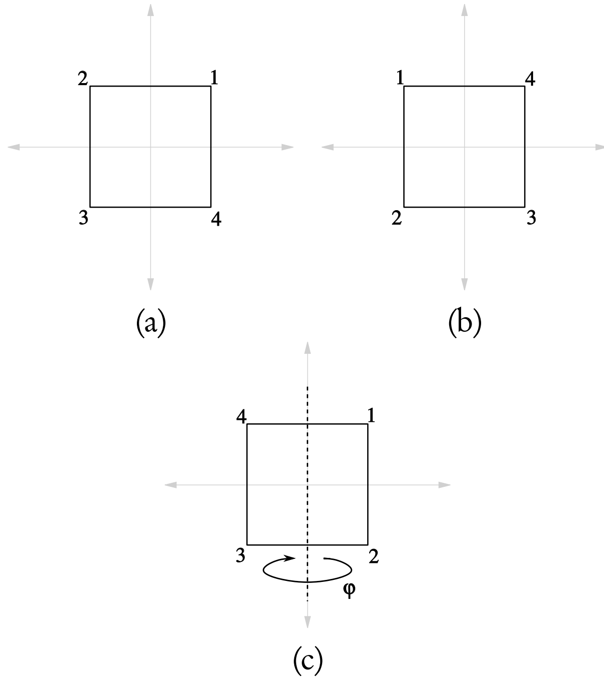


Figure 5.4. Rotation and reflection of a square centered at the origin

uration $(2,1,3,4)$ vertices 1 and 3 are adjacent on the square, so the distance between them has diminished to 1. Meanwhile, vertices 2 and 3 are no longer adjacent, so the distance between them has increased from 1 to $\sqrt{2}$. Since the distances between points on the square was not preserved, the permutation described, $\begin{pmatrix} 1 & 2 \end{pmatrix}$, is *not* an element of D_4. The same can be shown for the other fifteen permutations of four elements.

Hence D_4 has eight elements, making it smaller than S_4, which has $4! = 24$.

Is D_n always a group?

Theorem 5.43. Let $n \in \mathbb{N}^+$. If $n \geq 3$, then $(D_n, \circ)$ is a group with $2n$ elements, called the **dihedral group**.

It is possible to prove Theorem 5.43 using the following proposition, which could be proved using an argument from matrices, as in Section 2.2.

Proposition 5.44. All the symmetries of a regular n-sided polygon can be generated by a composition of a power of the rotation ρ of angle $2\pi/n$ and a power of the flip φ across the y-axis. In addition, $\varphi^2 = \rho^n = \iota$ (the identity symmetry) and $\varphi\rho = \rho^{n-1}\varphi$.

However, that would be a colossal waste of time. Instead, we prove the theorem *by turning symmetries of the polygon into permutations.*

D_n and S_n

Our strategy is as follows. For arbitrary $n \in \mathbb{N}^+$, we consider a list $(1,2,\ldots,n)$ of vertices of the n-sided polygon, imagine how they can move without violating the rules of symmetry,

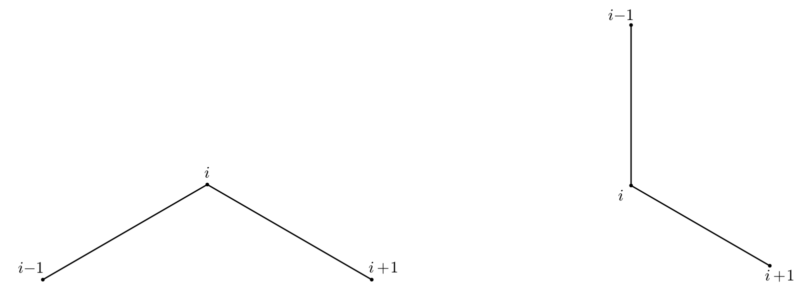


Figure 5.5. To preserve distance between vertices, a permutation of a regular polygon must move vertex i and its neighbors in such a way that they remain neighbors.

and then count how many possible permutations that gives us. We then show that this set of permutations satisfies the requirements of a group.

Proof of Theorem 5.43. Let $n \in \mathbb{N}^+$ and assume $n \geq 3$. Let $V = (1, 2, \ldots, n)$ be a list of the vertices of the n-sided polygon, *in order.* Thus, the distance from vertex $i-1$ to vertex i is precisely the distance from vertex i to vertex $i+1$.

What must be true after we apply any symmetry? While vertices $i-1$, i, and $i+1$ may have moved, the distances between them *may not* change. Thus, we can rearrange them in the order $i-1$, i, and $i+1$, but in different positions, or in the order $i+1$, i, $i-1$, in either the same or different positions. That limits our options. To count the number of possible symmetries, then, we start by counting the number of positions where we can move vertex 1: there are n such positions, one for each vertex. As we just observed, the vertex that follows vertex 1 *must* be vertex 2 or vertex n — if we are to preserve the distances between vertices, we have no other choice! (See Figure 5.5.) That gives us only two choices for the vertex that follows vertex 1! We can in fact create symmetries corresponding to these choices — simply count up or down, as appropriate. By the counting principle, D_n has $2n$ elements. But is it a group?

The associative property follows from the fact that permutations are functions, and composition of functions is associative. The identity symmetry, which moves the vertices onto themselves, corresponds to the identity element $\iota \in D_n$. The inverse property holds because (1) any permutation has an inverse permutation, and (2) Exercise 5.39 shows that this inverse permutation reverses the order of entries, so that the requirement that vertex $i-1$ precede or follow vertex i is preserved.

It remains to show closure. Let $\alpha, \beta \in D_n$, and let $i \in V$. Now, if $\beta(i) = j$, then the preservation of distance between vertices implies that $\beta(i+1)$ either precedes j or succeeds it; that is, $\beta(i+1) = j \pm 1$. If $\alpha(j) = k$, then the preservation of distance between vertices implies that $\alpha(j \pm 1)$ either precedes k or succeeds it; that is, $\alpha(j \pm 1) = k \pm 1$. By substitution,

$$(\alpha \circ \beta)(i) = \alpha(\beta(i)) = \alpha(j) = k$$

and

$$(\alpha \circ \beta)(i+1) = \alpha(\beta(i+1)) = \alpha(j \pm 1) = k \pm 1.$$

We see that $\alpha \circ \beta$ preserves the distance between the vertices, as vertex $i+1$ after the transformation either succeeds or precedes vertex i. Since i was arbitrary in V, this is true for all the vertices of the n-sided polygon. Thus, $\alpha \circ \beta \in D_n$, and D_n is closed.

We have shown that D_n has $2n$ elements, and that it satisfies the four properties of a group. □

The basic argument we followed above gives us the following result, as well.

Corollary 5.45. For any $n \geq 3$, D_n is isomorphic to a subgroup of S_n. If $n = 3$, then $D_3 \cong S_3$ itself.

Proof. You already proved that $D_3 \cong S_3$ in Exercise 5.35. □

What we have seen is that some problems, such as the symmetries of a regular polygon, fall naturally into a group-theoretical context if you can formulate the activity as a set of permutations. The next section shows that this is no accident.

Exercises.

Exercise 5.46. Write all eight elements of D_4 in cycle notation.

Exercise 5.47. Construct the Cayley table of D_4. Describe some similarities and differences between this result and that of Exercise 2.87.

Exercise 5.48. Show that the symmetries of any n-sided polygon can be described as a power of ρ and φ, where φ is a flip about the y-axis and ρ is a rotation of $2\pi/n$ radians.

Exercise 5.49. Show that D_n is solvable for all $n \geq 3$.

5.4: Cayley's Theorem

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.

Theorem 5.50 (Cayley's Theorem)**.** Every group of order n is isomorphic to a subgroup of S_n.

Before we give the proof, we give an example that illustrates how the proof of the theorem works.

Example 5.51. 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 will build the isomorphism by looking at the Cayley table for the Klein 4-group:

$\times$	e	a	b	ab
e	e	a	b	ab
a	a	e	ab	b
b	b	ab	e	a
ab	ab	b	a	e

To find a permutation appropriate to each element, we'll do the following. First, we label each element with a certain number:

$$\begin{aligned} e &\leftrightsquigarrow 1,\\ a &\leftrightsquigarrow 2,\\ b &\leftrightsquigarrow 3,\\ ab &\leftrightsquigarrow 4. \end{aligned}$$

We will use this along with tabular notation to determine the isomorphism. Define a map f from the Klein 4-group to S_4 by

$$f(x) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \ell(x \cdot e) & \ell(x \cdot a) & \ell(x \cdot b) & \ell(x \cdot ab) \end{pmatrix}, \tag{14}$$

where $\ell(y)$ is the label that corresponds to y.

This notation can make things hard to read. Why? Well, f maps an element g of the Klein 4-group to a permutation $f(x) = \sigma$ of S_4. Suppose $\sigma = (1\,2)(3\,4)$. 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 is hard to look at; is it saying "$f(x)$ times 2" or "$f(x)$ of 2"? In fact, it says the latter. 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 now to look at $f_x(2)$ and understand that we want $f_x(2) = 1$.

Let's compute f_a:

$$f_a = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \ell(a \cdot e) & \ell(a \cdot a) & \ell(a \cdot b) & \ell(a \cdot ab) \end{pmatrix}.$$

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 $\ell(a \cdot a) = \ell(a^2) = \ell(e) = 1$, 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 $\ell(a\cdot ab)=\ell(a^2b)=\ell(b)=3$, 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 (14) 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 (14) to the other elements of the Klein 4-group tells us that

$$\begin{aligned}f_e&=\begin{pmatrix}1&2&3&4\\1&2&3&4\end{pmatrix}=(1)\\f_b&=\begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix}=\begin{pmatrix}1&3\end{pmatrix}\begin{pmatrix}2&4\end{pmatrix}\\f_{ab}&=\begin{pmatrix}1&2&3&4\\4&3&2&1\end{pmatrix}=\begin{pmatrix}1&4\end{pmatrix}\begin{pmatrix}2&3\end{pmatrix}.\end{aligned}$$

The result is a subset of S_4; or, in cycle notation,

$$\begin{aligned}W&=\{f_e,f_a,f_b,f_{ab}\}\\&=\{(1),(1\,2)(3\,4),(1\,3)(2\,4),(1\,4)(2\,3)\}.\end{aligned}$$

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 element labeled by the number m. For instance, $\ell^{-1}(b)=3$.
- 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 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

$$f_x(m)=\ell\left(x\ell^{-1}(m)\right). \tag{15}$$

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 5.51 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 for any $x \in G$ denote $\ell(x) = i$ such that $x = g_i$. 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 Theorem 2.15 on page 79). The proof that f is a homomorphism is identical to the proof for Example 5.51: 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, *everything you need to know about finite groups can be learned from studying the groups of permutations!* A more flippant summary is that *the theory of finite groups is all about studying how to rearrange lists.*

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.

Exercises.

Exercise 5.52. In Example 5.51 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 Figures 28 on page 85 and 5.3.

Exercise 5.53. 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.

Exercise 5.54. The subgroup of S_4 that you identified in Exercise 5.53 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.

Exercise 5.55. 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.

5.5: Alternating groups

A special kind of group of permutations, with very important implications for later topics, are the *alternating groups*. To define them, we need to study permutations a little more closely, in particular the cycle notation.

Transpositions

Definition 5.56. 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 5.57. 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.

Remark 5.58. *Any transposition is its own inverse.* Why? Consider any transposition $(i\ j)$; it swaps the ith and jth elements of a list. Now consider the product $(i\ j)\,(i\ j)$. The rightmost $(i\ j)$ swaps these two, and the leftmost $(i\ j)$ swaps them back, restoring the list to its original arrangement. Hence $(i\ j)\,(i\ j) = (1)$.

Thanks to 1-cycles, any permutation can be written with many different numbers of cycles: for example,

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

A neat trick allows us to write every permutation as a composition of transpositions.

Example 5.59. Verify that

- $(1\,2\,3) = (1\,3)\,(1\,2)$;
- $(1\,4\,8\,2\,3) = (1\,3)\,(1\,2)\,(1\,8)\,(1\,4)$; and
- $(1) = (1\,2)\,(1\,2)$.

Did you see the relationship between the n-cycle and the corresponding transpositions?

Lemma 5.60. Any permutation can be written as a composition of transpositions.

Proof. You do it! See Exercise 5.71. □

Remark 5.61. Given an expression of σ as a product of transpositions, say $\sigma = \tau_1 \cdots \tau_n$, it is clear from Remark 5.58 that we can write $\sigma^{-1} = \tau_n \cdots \tau_1$, as an application of the associative property yields

$$\begin{aligned}(\tau_1 \cdots \tau_n)(\tau_n \cdots \tau_1) &= (\tau_1 \cdots \tau_{n-1})(\tau_n \tau_n)(\tau_{n-1} \cdots \tau_1) \\ &= (\tau_1 \cdots \tau_{n-1})(\ 1\ )(\tau_{n-1} \cdots \tau_1) \\ &\ \vdots \\ &= (1)\,.\end{aligned}$$

At this point it is worth looking at Example 5.59 and the discussion before it. Can we write $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$ with many different numbers of *transpositions*? Yes:

$$\begin{aligned}(1\,2\,3) &= (1\,3)(1\,2) \\ &= (1\,3)(1\,2)(2\,3)(2\,3) \\ &= (1\,3)(1\,2)(1\,3)(1\,3) \\ &= \cdots\,.\end{aligned}$$

Notice something special about the representation of $(1\,2\,3)$. No matter how you try, you only seem to be able to write it as an *even* number of transpositions. By contrast, consider

$$\begin{aligned}(2\,3) &= (2\,3)(2\,3)(2\,3) \\ &= (2\,3)(1\,2)(1\,3)(1\,3)(1\,2) = \cdots\,.\end{aligned}$$

No matter how you try, you only seem to be able to write it as an *odd* number of transpositions.

Is this always the case?

Even and odd permutations

Theorem 5.62. 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. Suppose that $\alpha \in S_n$. Consider the polynomials

$$g = \prod_{1 \le i < j \le n} \left(x_i - x_j\right) \qquad \text{and} \qquad g_\alpha = \prod_{1 \le i < j \le n} \left(x_{\alpha(i)} - x_{\alpha(j)}\right).$$

Since the value of g_α depends on the permutation α, and permutations are one-to-one functions, g_α is invariant with respect to the representation of α; that is, it won't change regardless of how we write α in terms of transpositions.

But what, precisely, *is* g_α? Sometimes $g = g_\alpha$; for example, if $\alpha = \left(\begin{array}{ccc} 1 & 3 & 2 \end{array}\right)$ then

$$g = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$$

and

$$g_\alpha = (x_3 - x_1)(x_3 - x_2)(x_1 - x_2) = [(-1)(x_1 - x_3)][(-1)(x_2 - x_3)](x_1 - x_2) = g. \tag{16}$$

Is it always the case that $g_\alpha = g$? Not necessarily: if $\alpha = \left(\begin{array}{cc} 1 & 2 \end{array}\right)$ then $g = x_1 - x_2$ and $g_\alpha = x_2 - x_1 \neq g$. In this case , $g_\alpha = -g$.

Since we cannot guarantee $g_\alpha = g$, can we write g_α in terms of g? Try the following. We know from Lemma 5.60 that α is a composition of transpositions, so let's think about what happens when we compute g_τ for any transposition $\tau = \left(\begin{array}{cc} i & j \end{array}\right)$. 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, $x_j - x_i$ is a factor of g_τ. This factor of g has changed in g_τ, since $x_j - x_i = -\left(x_i - x_j\right)$.
- If $i < j < k$, then $x_i - x_k$ and $x_j - x_k$ are factors of g. After applying τ, $x_i - x_k$ and $x_j - x_k$ are factors of g_τ. These factors of g have not changed in g_τ.
- 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.These factors of g have not changed in g_τ.
- 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. These factors of g have changed in g_τ, but the changes cancel each other out, since

$$\left(x_j - x_k\right)(x_k - x_i) = \left[-\left(x_k - x_j\right)\right][-(x_i - x_k)] = (x_i - x_k)\left(x_k - x_j\right).$$

To summarize: $x_i - x_j$ is the only factor that changes sign *and* does not pair with another factor that changes sign. Thus, $g_\tau = -g$.

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 5.60, 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,

$$g_\alpha = g_{\tau_1 \cdots \tau_m} = -g_{\tau_2 \cdots \tau_m} = (-1)^2 g_{\tau_3 \cdots \tau_m} = \cdots = (-1)^m g.$$

In short,

$$g_\alpha = (-1)^m g. \tag{17}$$

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 (17) 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 $(-1)^{2m} g = (-1)^{2k+1} g$; simplification yields $g = -g$, 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 5.60 tells us that any permutation can be written as a composition of transpositions, and Theorem 5.62 tells us that for any given permutation, this number is always either an even or odd number of transpositions. This relationship merits a definition.

Definition 5.63. If a permutation can be written with an even number of permutations, then we say that the permutation is **even**. Otherwise, we say that the permutation is **odd**.

Example 5.64. 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.

At this point, we are ready to define a new group.

The alternating group

Definition 5.65. Let $n \in \mathbb{N}^+$ and $n \geq 2$. Let $A_n = \{\alpha \in S_n : \alpha \text{ is even}\}$. We call A_n **the set of alternating permutations**.

Remark 5.66. Although A_3 is not the same as "A_3" in Example 3.68 on page 129, the two are isomorphic, because $D_3 \cong S_3$. For this reason, we need not worry about the difference in construction.

Theorem 5.67. For all $n \geq 2$, $A_n < S_n$.

Proof. Let $n \geq 2$, and let $x, y \in A_n$. By the definition of A_n, 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 5.61,

$$y^{-1} = \tau_{2n} \cdots \tau_1,$$

so

$$xy^{-1} = (\sigma_1 \cdots \sigma_{2m})(\tau_{2n} \cdots \tau_1).$$

Counting the transpositions, we find that xy^{-1} can be written as a product of $2m + 2n = 2(m+n)$ transpositions; in other words, $xy^{-1} \in A_n$. By the Subgroup Theorem, $A_n < S_n$. Thus, A_n is a group. □

How large is A_n, relative to S_n?

Theorem 5.68. For any $n \geq 2$, 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 from page 125.

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 3.37 on page 121 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 3.37, $\alpha A_n = \beta A_n$.

We have shown that any coset of A_n is either A_n itself or αA_n for some odd permutation α. Thus, there are only two cosets of A_n in S_n: A_n itself, and the coset of odd permutations. 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 5.69. For any $n \geq 2$, $A_n \triangleleft S_n$.

Proof. You do it! See Exercise 5.75. □

There are a number of *exciting* facts regarding A_n that have to wait until later; in particular, A_n has a pivotal effect on whether one can solve polynomial equations by radicals (such as the quadratic formula). In comparison, the facts presented here are relatively dull.

I say that only in comparison, though. The facts presented here are quite striking in their own right: A_n is half the size of S_n, and it is a normal subgroup of S_n. If I call these facts "rather dull", that tells you just how interesting this group can get!

Exercises.

Exercise 5.70. List the elements of A_2, A_3, and A_4 in cycle notation.

Exercise 5.71. Show that any permutation can be written as a product of transpositions.

Exercise 5.72. Show that the inverse of any transposition is a transposition.

Exercise 5.73. Recall the polynomials g and g_α defined in the proof of Theorem 5.62. 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?

Exercise 5.74. Recall the polynomials g and g_α defined in the proof of Theorem 5.62. The **sign function** $\operatorname{sgn}(\alpha)$ is defined to satisfy the property,

$$g = \operatorname{sgn}(\alpha) \cdot g_\alpha.$$

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,

$$(-1)^{\operatorname{sgn}(\alpha\beta)} = (-1)^{\operatorname{sgn}(\alpha)} (-1)^{\operatorname{sgn}(\beta)}.$$

Exercise 5.75. Show that for any $n \geq 2$, $A_n \triangleleft S_n$.

5.6: The 15-puzzle

The 15-puzzle is similar to a toy you probably played with as a child. It looks like 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 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	

The only permissible moves are those where one "slides" a square left, right, above, or below the empty square. Given the starting position above, the following first moves are permissible:

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	

.

We will use groups of permutations to show that that the challenge is 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 $x \notin 4\mathbb{Z}$; or
  - $y = x + 4$.

Example 5.76. 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 in the original configuration, 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 $y = 13 = 12 + 1 = x + 1$, $x = 12 \in 4\mathbb{Z}$.

How can we use this to show that it is impossible to solve 15-puzzle? We show this in two steps. The first shows that if there is a solution, it must belong to a particular group.

Lemma 5.77. If there is a solution to the 15-puzzle, it is a permutation $\sigma \in A_{16}$, where A_{16} is the alternating group.

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 $(x'\ y')$ where $[x'] = [x]$ in $\mathbb{Z}_4$ and $[y'] = [y]$ in $\mathbb{Z}_4$; 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 $(x'\ y')$ of the second type.

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 5.62 on page 186 and Definitions 5.63 and 5.65, $\sigma \in A_{16}$. □

We can now show that there is no solution to the 15-puzzle.

Theorem 5.78. The 15-puzzle has no solution.

Proof. By way of contradiction, assume that it has a solution σ. By Lemma 5.77, $\sigma \in A_{16}$. 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 $(14\ 15)$. Regardless of the transpositions used, the representation must simplify to $\sigma^{-1} = (14\ 15)$. This shows that $\sigma \notin A_{16}$, which contradicts the assumption that we have a contradiction. □

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.[22]

The textbook [Lau03] contains a more general discussions of solving puzzles of this sort using algebra.

Exercises

Exercise 5.79. 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	

.

[22]According to the website `www.measuringworth.com/ppowerus/result.php`.

Chapter 6:
Number theory

The theory of groups was originally developed to answer questions about the roots of polynomials. From such beginnings, it has grown to many applications that seem at first completely unrelated to this topic. Some of the most widely-used applications in recent decades are in number theory, the study of properties of the integers.

This chapter introduces several of these applications. Section 6.1 fills some background with two of the most important tools in computational algebra and number theory. The first is a fundamental definition; the second is a fundamental algorithm. Both recur throughout the chapter, and later in the notes. Section 6.2 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 the card trick described on page 2.

The rest of the chapter moves us toward Section 6.6, the RSA cryptographic scheme, a major component of internet communication and commerce. In Section 3.5 you learned of additive clockwork groups; in Section 6.4 you will learn of multiplicative clockwork groups. These allows us to describe in Section 6.5 the theoretical foundation of RSA, Euler's number and Euler's Theorem.

6.1: The Greatest Common Divisor

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 $ax = b$.

Definition 6.1. Let $m, n \in \mathbb{Z}$, not both zero. We say that $d \in \mathbb{Z}$ is a **common divisor of** m **and** n if $d \mid m$ and $d \mid n$. We say 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 6.2. Common divisors of 36 and −210 are 1, 2, 3, and 6. The greatest common divisor is 6.

In grade school, you learned how to compute the greatest common divisor of two integers. For example, given the integers 36 and 210, you can find their greatest common divisor, 6. Computing greatest common divisors—not only of integers, but of other objects as well — is an important problem in mathematics, with a large number of important applications. Arguably, it is one of the most important problems in mathematics, and it has an ancient pedigree.

But, do greatest common divisors always exist?

Theorem 6.3. Let $m, n \in \mathbb{Z}$, not both zero. There exists a unique greatest common divisor of m, n.

Algorithm 1. The Euclidean algorithm

```
inputs
  m, n ∈ ℤ
outputs
  gcd(m, n)
do
  Let s = max(m, n)
  Let t = min(m, n)
  repeat while t ≠ 0
    Let q, r ∈ ℤ be the result of dividing s by t
    Let s = t
    Let t = r
  return s
```

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 $D \neq \emptyset$. We also know that any $d \in D$ must satisfy $d \leq \min(m,n)$; 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) $d' \leq d$, and $d \neq d'$ implies $d' < d$.

□

How can we compute the greatest common divisor? One way is to make a list of all common divisors, and find the largest. That would require a list of all possible divisors of each integer. 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

Theorem 6.4 (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 $s = \max(m,n)$ and $t = \min(m,n)$.
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 $\gcd(m,n)$.

It is common to write algorithms in a form called *pseudocode*. You can see this done in Algorithm 1.

Before proving that the Euclidean algorithm gives us a correct answer, let's do an example.

Example 6.5. 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 6.2.

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,

$$\gcd(m,n) = \gcd(m,r_1) = \gcd(r_1,r_2) = \cdots = \gcd\left(r_{k-1},0\right),$$

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 6.6. 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 from page 17. Then $\gcd(s,t) = \gcd(t,r)$.

Example 6.7. We can verify Lemma 6.6 using the numbers from Example 6.5. We know that $\gcd(36,210) = 6$. The remainder from division of 36 by 210 is $r = 36$. The lemma claims that $\gcd(36,210) = \gcd(36,30)$; it should be clear to you that $\gcd(36,30) = 6$.

The example also shows that the lemma doesn't care whether $m < n$ or vice versa. We turn to the proof.

Proof of Lemma 6.6. Let $d = \gcd(s,t)$. First we show that d is a divisor of r. From Definition **??** on page ??, there exist $a, b \in \mathbb{Z}$ such that $s = ad$ and $t = bd$. By hypothesis, $s = qt + r$ and $0 \leq r < |t|$. 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 $d \leq d'$. Since d' is a common divisor of t and r, Definition **??** 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 $s = q\left(d'x\right) + d'y$; 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 $d' \leq d$. Earlier, we showed that $d \leq d'$. Hence $d \leq d' \leq d$, 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 have to guarantee that it terminates with the correct answer.
2. **Termination.** What if the algorithm doesn't terminate? If you look at the Euclidean algorithm, you see that one of its instructions 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 over zero from positive to negative values. That would mean that we never receive any answer from the algorithm, let alone a correct one.

We will identify both stages of the proof clearly. In addition, we will refer back to the the Division Theorem as well as the well-ordering property of the integers from Section 8; 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,

$$0 \leq r < |t|. \tag{18}$$

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 (18), 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

$$0 \leq r_2 < r_1.$$

Proceeding in this fashion, we generate a strictly decreasing sequence of elements,

$$r_1 > r_2 > r_3 > \cdots.$$

By Exercise 0.37, 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 6.6 repeatedly to the remainders to obtain the chain of equalities

$$\begin{aligned}
r_{k-1} = \gcd(0, r_{k-1}) &= \gcd(r_k, r_{k-1}) && \text{(definition of gcd, substitution)}\\
&= \gcd(r_{k-1}, r_{k-2}) && \text{(Lemma 6.6)}\\
&= \gcd(r_{k-2}, r_{k-3}) && \text{(Lemma 6.6)}\\
&\vdots\\
&= \gcd(r_2, r_1) && \text{(Lemma 6.6)}\\
&= \gcd(r_1, s) && \text{(substitution)}\\
&= \gcd(t, s) && \text{(substitution)}\\
&= \gcd(m, n). && \text{(substitution)}
\end{aligned}$$

The Euclidean Algorithm terminates with the correct answer. □

Bezout's identity

A fundamental fact of number theory is that the greatest common divisor of two integers can be expressed as a simple expression of those integers.

Theorem 6.8 (Bezout's Lemma, or, 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 reverse-substituting the chain of equations obtained by the repeated division in the Euclidean algorithm.

The expression, $am + bn = \gcd(m, n)$, is important enough to be known by the name, **Bezout's identity**. It can be used to prove a *lot* of properties of greatest common divisors.

Example 6.9. Recall from Example 6.5 the computation of $\gcd(210, 36)$. The divisions gave us a series of equations:

$$210 = 5 \cdot 36 + 30 \tag{19}$$

$$36 = 1 \cdot 30 + 6 \tag{20}$$

$$30 = 5 \cdot 6 + 0.$$

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 (20):

$$36 - 1 \cdot 30 = 6. \tag{21}$$

This looks a little like what we want, but we need 210 instead of 30. Equation (19) allows us to rewrite 30 in terms of 210 and 36:

$$30 = 210 - 5 \cdot 36. \tag{22}$$

Substituting this result into equation (21), 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 $b = 210$, $\gcd(a, b) = 6$.

The method we applied in Example (6.9) 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

$$\begin{aligned}
m &= q_1 n + r_1 \\
n &= q_2 r_1 + r_2 \\
r_1 &= q_3 r_2 + r_3 \\
&\vdots \\
r_{k-3} &= q_{k-1} r_{k-2} + r_{k-1} && (23) \\
r_{k-2} &= q_k r_{k-1} + r_k && (24) \\
r_{k-1} &= q_{k+1} r_k + 0 \\
\text{and } r_k &= \gcd(m,n).
\end{aligned}$$

Rewrite equation (24) as

$$r_{k-2} = q_k r_{k-1} + \gcd(m,n).$$

Solving for $\gcd(m,n)$, we have

$$r_{k-2} - q_k r_{k-1} = \gcd(m,n). \tag{25}$$

Solve for r_{k-1} in equation (23) to obtain

$$r_{k-3} - q_{k-1} r_{k-2} = r_{k-1}.$$

Substitute this into equation (25) to obtain

$$\begin{aligned}
r_{k-2} - q_k \left(r_{k-3} - q_{k-1} r_{k-2}\right) &= \gcd(m,n) \\
\left(q_{k-1} + 1\right) r_{k-2} - q_k r_{k-3} &= \gcd(m,n).
\end{aligned}$$

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. □

Pseudocode appears in Algorithm 2. One can also derive a method of computing both $\gcd(m,n)$ and the representation $am + bn = \gcd(m,n)$ simultaneously, which is to say, without having to reverse the process. We will not consider that here.

Exercises.

Exercise 6.10. Compute the greatest common divisor of 100 and 140 by (a) listing all divisors, then identifying the largest; and (b) the Euclidean Algorithm.

Exercise 6.11. 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.

Exercise 6.12. Show that any common divisor of any two integers divides the integers' greatest common divisor.

Algorithm 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 $n = 0$
    Let $d = m$, $a = 1$, $b = 0$
  else
    Let $r_0 = m$ and $r_1 = n$
    Let $k = 1$
    repeat while $r_k \neq 0$
      Increment k by 1
      Let q_k, r_k be the quotient and remainder from division of r_{k-2} by r_{k-1}
    Let $d = r_{k-1}$ and $p = r_{k-3} - q_{k-1} r_{k-2}$ (do not simplify p)
    Decrement k by 2
    repeat while $k \geq 2$
      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

Exercise 6.13. Recall the roots of unity $\Omega_n = \left\{1, \omega, \omega^2, \ldots, \omega^{n-1}\right\}$ where

$$\omega = \cos\left(\frac{2\pi}{n}\right) + i \sin\left(\frac{2\pi}{n}\right).$$

(If you do not recall them, see p. 103.) Definition 2.77 tells us that a root of unity is primitive if it generates Ω_n.

Exercise 6.14. In Lemma 6.6 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 $m, n, q \in \mathbb{Z}$ $\gcd(m,n) = \gcd(n, m - qn)$.

Exercise 6.15. Bezout's Identity (Theorem 6.8) 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 6.1 with the appropriate justification or statement.

6.2: The Chinese Remainder Theorem

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 property of $\mathbb{Z}$ 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 _____, $d \leq \gcd(m,n)$ and $\gcd(m,n) \leq d$.
4. By _____, $d = \gcd(m,n)$.

Figure 6.1. Material for Exercise 6.15

In this section we explain how the card trick on page 2 works. The result is based on an old Chinese observation.[23] Recall from Section 3.5 that for any $m \neq 0$ there exists a group $\mathbb{Z}_m$ of m elements, under the operation of adding, then taking remainder after division by m. Remember that we often write $[x]$ for the elements of $\mathbb{Z}_m$ if we want to emphasize that its elements are cosets.

The simple Chinese Remainder Theorem

Theorem 6.16 (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.

Example 6.17 (The card trick)**.** In the card trick, we took twelve cards and arranged them

- once in groups of three; and
- once in groups of four.

Each time, the player 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 in fact the remainders from division by three and by four. This corresponds to a system of linear congruences,

$$\begin{cases} [x] = [\alpha] \text{ in } \mathbb{Z}_3 \\ [x] = [\beta] \text{ in } \mathbb{Z}_4 \end{cases},$$

[23] 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. Unfortunately, this is not the same Sun Tzu who wrote *The Art of War*.

where x is the location of the mystery card. The simple version of the Chinese Remainder Theorem guarantees a solution for x, which 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." It turns out that Bezout's identity,

$$am + bn = \gcd(m, n),$$

is the key to unlocking the Chinese Remainder Theorem. Before doing so, we need an important lemma about numbers whose gcd is 1.

Lemma 6.18. 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 that 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 $[am]_n = [1]_n$. By Lemmas 3.95 and 3.98 on page 137, $[\alpha]_m = \alpha[1]_m = \alpha[bn]_m = [\alpha bn]_m$ 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.

Algorithm 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 $[\alpha bn + \beta am]_N$

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 3.101 on page 139, $m \mid (x - y)$. 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 6.18, $n \mid q$. 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 3.101 $[x]_N = [y]_N$, which means that the solution x is unique in $\mathbb{Z}_N$, as desired. □

Pseudocode to solve the Chinese Remainder Theorem appears as Algorithm 3.

Example 6.19. The algorithm of Corollary 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.

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 groups? You might like to arrange the cards into rows of 3, 4, 5, and 7. What about other groupings? What constraints do there have to be on the groupings, and how would we solve the new problem?

Theorem 6.20 (Chinese Remainder Theorem on $\mathbb{Z}$)**.** Let $m_1, m_2, \ldots, m_n \in \mathbb{Z}$ and assume that $\gcd\left(m_i, m_j\right) = 1$ for all $1 \leq i < j \leq n$. 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 6.21. Let $m_1, m_2, \ldots, m_n \in \mathbb{Z}$ such that $\gcd\left(m_i, m_j\right) = 1$ for all $1 \leq i < j \leq n$. 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

$$\begin{aligned} 1 &= \left(a' m_1 + b' m_2 m_3\right)(em_1 + fm_4) \\ &= \left(a' em_1 + a' fm_4 + b' em_2 m_e\right) m_1 + \left(b' f\right)(m_2 m_3 m_4). \end{aligned}$$

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 6.21, $\gcd(m_i, N_i) = 1$. 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 3.101). Hence

$$\begin{aligned} [x] &= [\alpha_1 b_1 N_1 + \alpha_2 b_2 N_2 + \cdots + \alpha_n b_n N_n] \\ &= [\alpha_1] + [0] + \cdots + [0] \end{aligned}$$

in $\mathbb{Z}_{m_1}$, as desired. 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 $i = 1, 2, \ldots, n$.

Since $m_1 \mid (x - y)$, the definition of divisibility implies that 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 6.18 implies that $m_2 \mid q_1$. The definition of divisibility implies that 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 6.21, $\gcd(m_1 m_2, m_3) = 1$, and Lemma 6.18 implies that $m_3 \mid q_2$. The definition of divisibility implies that 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, we show that $x - y = m_1 m_2 \cdots m_n q_n$ for some $q_n \in \mathbb{Z}$. By substitution, $x - y = N q_n$, 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.

Exercises

Exercise 6.22. 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 $0 \le x < 36$.

Exercise 6.23. Solve the system of linear congruences

$$\begin{cases} [x] = [2] \text{ in } \mathbb{Z}_5 \\ [x] = [3] \text{ in } \mathbb{Z}_6 \\ [x] = [4] \text{ in } \mathbb{Z}_7 \end{cases}.$$

Exercise 6.24. 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.

Exercise 6.25. 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?

Exercise 6.26. Is it possible to modify the card trick to work with only ten cards instead of 12? If so, how; if not, why not?

Exercise 6.27. Is it possible to modify the card trick to work with only eight cards instead of 12? If so, how; if not, why not?

6.3: The Fundamental Theorem of Arithmetic

In this section, we address a fundamental result of number theory with algebraic implications.

Definition 6.28. Let $n \in \mathbb{N}^+ \backslash \{1\}$. We say that n is **irreducible** if the only integers that divide n are ± 1 and $\pm n$.

You may read this and think, "Oh, he's talking about prime numbers." Yes and no. We'll say more about that in a moment.

Example 6.29. 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 6.30. Let $n \in \mathbb{Z} \backslash \{\pm 1\}$. 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 $n \in \mathbb{N}^+ \backslash \{1\}$. If n is not irreducible, then by definition $n = a_1 b_1$ 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 a, b are negative and we can replace them with their opposites). Observe further that $a_1 < n$ (this is a consequence of Exercise 0.32 on page 15). 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 $a_2 < a_1$.

Let $a_0 = n$. As long as a_i is not irreducible, we can find $a_{i+1}, b_{i+1} \in \mathbb{N}^+$ such that $a_i = a_{i+1} b_{i+1}$. By Exercise 0.32, $a_i > a_{i+1}$ for each i. Proceeding in this fashion, we generate a strictly decreasing sequence of elements,

$$a_0 > a_1 > a_2 > \cdots .$$

By Exercise 0.37, this sequence *must* be finite. Let a_m be the final element in the sequence. We claim that a_m is irreducible; after all, if it were not irreducible, then we could extend the sequence further, and we cannot. By substitution,

$$n = a_1 b_1 = a_2 (b_2 b_1) = \cdots = a_m \left(b_{m-1} \cdots b_1\right).$$

That is, a_m is an irreducible integer that divides n.

Case 3: Assume that $n \in \mathbb{Z} \setminus (\mathbb{N} \cup \{-1\})$. Let $m = -n$. Since $m \in \mathbb{N}^+ \setminus \{1\}$, 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 $p \mid n$. □

Let's turn now to the term you might have expected for the definition given above: a *prime* number. For reasons that you will learn later, we actually associate a different notion with this term.

Definition 6.31. Let $p \in \mathbb{N}^+ \setminus \{1\}$. We say that p is **prime** if for any two integers a, b

$$p \mid ab \quad \Longrightarrow \quad p \mid a \text{ or } p \mid b.$$

Example 6.32. 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 an infinite number of 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. Previously, you were probably taught that a *prime* number was what we have here called *irreducible*. I've now given a definition that seems different.

Could it be that the definitions are *distinctions without a difference?* Indeed, they are equivalent!

Theorem 6.33. An integer is prime if and only if it is irreducible.

Proof. This proof has two parts. You will show in Exercise 6.35 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 6.18, $n \mid b$.

We assumed that if n is irreducible and divides ab, then n must divide one of a or b. By definition, n is prime. □

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 later in Sections 8.4 and 10.1.

The following theorem is a cornerstone of Number Theory.

Theorem 6.34 (The Fundamental Theorem of Arithmetic)**.** Let $n \in \mathbb{N}^+ \setminus \{1\}$. 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 textboks. We use "irreducible" here to lay the groundwork for Definition 10.16 on page 319.

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, ..., $n-1$ have a factorization into irreducibles.

Inductive step: If n is irreducible, then we are finished. Otherwise, n is not irreducible. By Lemma 6.30, 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 $p_1 \neq 1$, Exercise 0.49 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 6.33). 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 $\alpha_1 \leq \beta_1$. 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 $p_2, p_3, \ldots, p_r$.

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 6.2. Material for Exercise 6.37

This contradicts the irreducibility of $p_2, p_3, \ldots, p_r$. 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. □

Exercises.

Exercise 6.35. Show that any prime integer p is irreducible.

Exercise 6.36. Show that there are infinitely many irreducible integers.

Exercise 6.37. Fill in each blank of Figure 6.2 with the justification.

Exercise 6.38. Let $n \in \mathbb{N}^+ \setminus \{1\}$. Modify the proof in Figure 6.2 to show that if p is irreducible, then $\sqrt[n]{p}$ is irrational.

Exercise 6.39. Let $n \in \mathbb{N}^+ \setminus \{1\}$. Modify the proof in Figure 6.2 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.

6.4: Multiplicative clockwork groups

Throughout this section, $n \in \mathbb{N}^+ \setminus \{1\}$, unless otherwise stated.

Multiplication in $\mathbb{Z}_n$

Recall that $\mathbb{Z}_n$ is an additive group, but not multiplicative. In this section we find a subset of $\mathbb{Z}_n$ that we can turn into a multiplicative group, where multiplication is "intuitive":

$$[2]_5 \cdot [3]_5 = [2 \cdot 3]_5 = [6]_5 = [1]_5 .$$

Remember, though: cosets can have various representations, and different representations may lead to different results. We have to ask ourselves, is this operation well-defined?

Lemma 6.40. The proposed multiplication of elements of $\mathbb{Z}_n$ as

$$[a]\,[b] = [ab]$$

is well-defined.

This lemma requires no special constraints on n, so it applies even if $n \in \mathbb{Z}$ is arbitrary.

Proof. Let $x, y \in \mathbb{Z}_n$. Choose $a, b, c, d \in \mathbb{Z}$ such that $x = [a] = [c]$ and $y = [b] = [d]$. By definition of the operation,

$$xy = [a]\,[b] = [ab] \quad \text{and} \quad xy = [c]\,[d] = [cd] .$$

We need to show that $[ab] = [cd]$. The best tool for this is Lemma 3.101 on page 139, which tells us that if we can show that $ab - cd \in n\mathbb{Z}$, then we're done.

How can we accomplish this? By assumption, $[a] = [c]$; this notation means that $a + n\mathbb{Z} = c + n\mathbb{Z}$. Lemma 3.101 tells us that $a - c \in n\mathbb{Z}$. By definition, $a - c = nt$ for some $t \in \mathbb{Z}$. Similarly, $b - d = nu$ for some $u \in \mathbb{Z}$. We can build ab using these differences by multiplying $b\,(a - c)$, but this actually equals $ac - bc$. We can cancel bc using these differences by adding $c\,(b - d)$, and that will give us precisely what we need:

$$\begin{aligned} ab - cd &= b\,(a - c) + c\,(b - d) \\ &= b\,(nt) + c\,(nu) \\ &= n\,(bt + cu), \end{aligned}$$

so $ab - cb \in n\mathbb{Z}$. Lemma 3.101 again tells us that $[ab] = [cb]$ as desired, so the proposed multiplication of elements in $\mathbb{Z}_n$ is well-defined. □

Example 6.41. Recall that $\mathbb{Z}_5 = \mathbb{Z}/\langle 5 \rangle$. The elements of $\mathbb{Z}_5$ are cosets; since $\mathbb{Z}$ is an additive group, we were able to define easily an addition on $\mathbb{Z}_5$ that turns it into an additive group in its own right.

Can we also turn it into a multiplicative group? We need to identify an identity, and inverses. Certainly $[0]$ won't have a multiplicative inverse, but what about $\mathbb{Z}_5 \backslash \{[0]\}$? This generates a multiplication table that satisfies the properties of an abelian (but non-additive) group:

×	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 the cyclic group of four elements. In $\mathbb{Z}_5^*$, however, the nominal operation is multiplication, whereas in $\mathbb{Z}_4$ the nominal operation is addition.

You might think that this trick of dropping zero and building a multiplication table always works, *but it doesn't*.

Example 6.42. Recall that $\mathbb{Z}_4 = \mathbb{Z}/\langle 4\rangle = \{[0],[1],[2],[3]\}$. Consider the set $\mathbb{Z}_4 \backslash \{[0]\} = \{[1],[2],[3]\}$. The multiplication table for this set *is not closed* because

$$[2]\cdot[2] = [4] = [0] \notin \mathbb{Z}_4 \backslash \{[0]\}.$$

If you are tempted to think that we made a mistake by excluding zero, think twice: zero has no inverse. So, we must exclude zero; our mistake seems to have been that we must also exclude 2. This finally works out:

×	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 $\mathbb{Z}_2 \backslash \{[0]\} = \{[1]\}$ and $\mathbb{Z}_3 \backslash \{[0]\} = \{[1],[2]\}$ 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 $\mathbb{Z}_6 \backslash \{[0],[2],[3],[4]\} = \{[1],[5]\}$.

Zero divisors

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.

That's strange: didn't we once learn that the product of two nonzero numbers is nonzero? Yet here we have non-zero elements whose product is zero! True, but this is a different set than the one where you learned the zero product property. Here is an instance where $\mathbb{Z}_n$ superficially behaves *very differently* from the integers.

In fact, you have seen this phenomenon before: look back at Exercises 3.47 and 0.87. This phenomenon is so important that it has a special name.

Definition 6.43. We say that nonzero elements $x, y \in \mathbb{Z}_n$ are **zero divisors** if $xy = [0]$.

In other words, zero divisors are non-zero elements of $\mathbb{Z}_n$ that violate the zero product property.

Can we find a criterion to detect this?

Lemma 6.44. Let $x \in \mathbb{Z}_n$ be nonzero. The following are equivalent:

(A) x is a zero divisor.

(B) For any representation $[a]$ of x, a and n have a common divisor besides ± 1.

Proof. *That (B) implies (A):* Let $[a]$ be any representation of x, and assume that a and n share a common divisor $d \neq 1$. Use the definition of divisibility to choose $t, q \in \mathbb{Z} \setminus \{0\}$ such that $n = qd$ and $a = td$. Let $y = [q]$. Substitution and Lemma 6.40 imply that

$$xy = [a][q] = [aq] = [(td)q] = t[qd] = t[n] = [0].$$

Since $d \neq 1$, $-n < q < n$, so $[0] \neq [q] = y$. By definition, x is a zero divisor.

That (A) implies (B): Assume that x is a zero divisor. By definition, we can find nonzero $y \in \mathbb{Z}_n$ such that $xy = [0]$. Choose $a, b \in \mathbb{Z}$ such that $x = [a]$ and $y = [b]$. Since $xy = [0]$, Lemma 3.101 implies that $n \mid (ab - 0)$, so we can find $k \in \mathbb{Z}$ such that $ab = kn$. Let p_0 be any irreducible number that divides n. Then p_0 also divides kn. Since $kn = ab$, we see that $p_0 \mid ab$. Since p_0 is irreducible, hence prime, it must divide one of a or b. If it divides a, then a and n have a common divisor p_0 that is not ± 1, and we are done; otherwise, it divides b. Use the definition of divisibility to find $n_1, b_1 \in \mathbb{Z}$ such that $n = n_1 p_0$ and $a = b_1 p_0$; it follows that $ab_1 = kn_1$. Again, let p_2 be any irreducible number that divides n_2; the same logic implies that p_2 divides ab_2; being prime, p_2 must divide a or b_2.

As long as we can find prime divisors of the n_i that divide b_i but not a, we repeat this process to find triplets $(n_2, b_2, p_2), (n_3, b_3, p_3), \ldots$ satisfying for all i the properties

- $ab_i = kn_i$;
- $b_{i-1} = p_i b_i$ and $n_{i-1} = p_i n_i$; and so, by Exercise 0.49,
- $|n_{i-1}| > |n_i|$.

The sequence $|n|, |n_1|, |n_2|, \ldots$ is a decreasing sequence of elements of $\mathbb{N}$; by Exercise (0.37), it is finite, and so has a least element, call it $|n_r|$. Observe that

$$b = p_1 b_1 = p_1(p_2 b_2) = \cdots = p_1(p_2(\cdots(p_r b_r))) \tag{26}$$

and

$$n = p_1 n_1 = p_1(p_2 n_2) = \cdots = p_1(p_2(\cdots(p_r n_r))).$$

Case 1. If $n_r = \pm 1$, then $n = p_1 p_2 \cdots p_r$. By substitution into equation 26, $b = nb_r$. By the definition of divisibility, $n \mid b$. By the definition of $\mathbb{Z}_n$, $y = [b] = [0]$. This contradicts the hypothesis.

Case 2. If $n_r \neq \{\pm 1\}$, then Theorem 6.30 tells us that n_r has an irreducible divisor p_{r+1}. Since $p_{r+1} \mid kn_r$, it must also divide ab_r. If $p_{r+1} \mid b_r$, then we can construct n_{r+1} and b_{r+1} that satisfy the properties above for $i = r+1$. As before, $|n_{r+1}| < |n_r|$, which contradicts the choice of n_r. Hence $p_{r+1} \nmid b_r$; since irreducible integers are prime, $p_{r+1} \mid a$.

Hence n and a share a common divisor that is not ± 1. □

Meet $\mathbb{Z}_n^*$

We can now make a *multiplicative* group out of the set of elements of $\mathbb{Z}_n$ that do not violate the zero product rule.

Definition 6.45. Define the set $\mathbb{Z}_n^*$ to be the set of nonzero elements of $\mathbb{Z}_n$ that are not zero divisors. In set builder notation,

$$\mathbb{Z}_n^* := \{X \in \mathbb{Z}_n \setminus \{0\} : \forall Y \in \mathbb{Z}_n \setminus \{0\}\ XY \neq 0\}.$$

By Lemma 6.44, we could also say that $\mathbb{Z}_n^*$ is the set of positive numbers less than n whose only common factors with n are ± 1. This is the usual definition of $\mathbb{Z}_n^*$ in number theory.

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 6.46. $\mathbb{Z}_n^*$ is an abelian group under its multiplication.

Proof. We showed in Lemma 6.40 that the operation is well-defined. We check each requirement of a group, slightly out of order. Let $X, Y, Z \in \mathbb{Z}_n^*$, and choose $a, b, c \in \mathbb{Z}$ such that $X = [a]$, $Y = [b]$, and $Z = [c]$.

(associative) By substitution and properties of $\mathbb{Z}_n^*$, $\mathbb{Z}_n$, and $\mathbb{Z}$,

$$X(YZ) = [a][bc] = [a(bc)] = [(ab)c] = [ab][c] = (XY)Z.$$

Notice that this applies for elements of $\mathbb{Z}_n$ as well as elements of $\mathbb{Z}_n^*$.

(closed) Since the operation is well-defined, $XY \in \mathbb{Z}_n$. How do we know that $XY \in \mathbb{Z}_n^*$? Assume to the contrary that it is not. That would mean that $XY = [0]$ or XY is a zero divisor; either way, $\gcd(ab, n) \neq 1$. By definition of $\mathbb{Z}_n^*$, neither X nor Y is a zero divisor, so $XY \neq [0]$, which forces us to conclude that XY is a zero divisor. By definition of zero divisor, there must be some $Z \in \mathbb{Z}_n$ such that $(XY)Z = [0]$. By the associative property, $X(YZ) = [0]$; that is, X is a zero divisor. This contradicts the choice of X! Thus, XY *cannot* be a zero divisor; the assumption that $XY \notin \mathbb{Z}_n^*$ must have been wrong.

(identity) We claim that $[1]$ is the identity. Since $\gcd(1, n) = 1$, Lemma 6.44 tells us that $[1] \in \mathbb{Z}_n^*$. By substitution and arithmetic in both $\mathbb{Z}_n^*$ and $\mathbb{Z}_n$,

$$X \cdot [1] = [a \cdot 1] = [a] = X.$$

A similar argument shows that $[1] \cdot X = X$.

(inverse) We need to find an inverse of X. From Lemma 6.44, a and n have no common divisors except ± 1; hence $\gcd(a, n) = 1$. Bezout's Identity tells us that there exist $b, m \in \mathbb{Z}$ such that $ab + mn = 1$. By arithmetic in both $\mathbb{Z}_n^*$ and $\mathbb{Z}$, as well as Lemma 3.101, we deduce that

$$\begin{aligned} ab - 1 &= n(-m) \\ \therefore ab - 1 &\in n\mathbb{Z} \\ \therefore [ab] &= [1] \\ \therefore [a][b] &= [1]. \end{aligned}$$

Let $Y = [b]$; by substitution, the last equation becomes

$$XY = [1].$$

But is $Y \in \mathbb{Z}_n^*$? In fact it is, and the justification is none other than the same Bezout Identity we used above! We had $ab + mn = 1$. I hope you agree that we can't find a positive integer smaller than 1. You will also agree that 1 is the *smallest* positive integer d for which we can find $w, z \in \mathbb{Z}$ such that $bw + nz = d$, if we can find such $w, z \in \mathbb{Z}$. In fact, we can: the Bezout Identity above provides a solution, where $d = 1$, $w = a$, and $z = m$. Guess what: Exercise 6.15(b) tells us that $\gcd(b, n) = 1$! By definition, then, $Y = [b] \in \mathbb{Z}_n^*$, and X has an inverse in $\mathbb{Z}_n^*$.

(commutative) Use the definition of multiplication in $\mathbb{Z}_n^*$ and the commutative property of integer multiplication to see

$$XY = [ab] = [ba] = YX.$$

□

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 6.47. Consider $\mathbb{Z}_{10}^*$. To find its elements, collect the elements of $\mathbb{Z}_{10}$ that are not zero divisors. Lemma 6.44 tells us that these are the elements whose representations $[a]$ satisfy $\gcd(a, n) \neq 1$. Thus

$$\mathbb{Z}_{10}^* = \{[1], [3], [7], [9]\}.$$

Theorem 6.46 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:

×	**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 $\mathbb{Z}_4$.

Exercises.

Exercise 6.48. 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.

Exercise 6.49. 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.

6.5: Euler's Theorem

In Section 6.4 we defined the group $\mathbb{Z}_n^*$ for all $n \in \mathbb{N}^+$ where $n > 1$. This group satisfies an important property called *Euler's Theorem*, a result about Euler's φ-function.

Euler's Theorem

Definition 6.50. Euler's φ-function is $\varphi(n) = |\mathbb{Z}_n^*|$.

In other words, Euler's φ-function counts the number of positive integers smaller than n that share no common factors with it.

Theorem 6.51 (Euler's Theorem)**.** 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 [Lau03].

On the other hand, a proof of Euler's Theorem using group theory is short and straightforward.

Proof. Let $x \in \mathbb{Z}_n^*$. By Exercise 3.59, $x^{|\mathbb{Z}_n^*|} = 1$. By substitution, $x^{\varphi(n)} = 1$. □

Corollary 6.52. For all $x \in \mathbb{Z}_n^*$, $x^{-1} = x^{\varphi(n)-1}$.

Proof. You do it! See Exercise 6.61. □

Corollary 6.52 says that we can compute x^{-1} for any $x \in |\mathbb{Z}_n^*|$ "relatively easily;" all we need to know is $\varphi(n)$.

Computing $\varphi(n)$

The natural followup question is, what is $\varphi(n)$? 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 reducible integers, 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 quite hard. It is a major research topic in number theory, and its difficulty makes the RSA algorithm secure (see Section 6.6). One approach, of course, is to factor n and compute all the positive 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. However, this requires us to know first that 2 and 7 are factors of 28, and no one knows a very *efficient* way to do this.

Another way would be to compute $\varphi(m)$ for each factor m of n, then recombine them. But, how? Lemma 6.53 gives us a first step.

Lemma 6.53. Let $a, b, n \in \mathbb{N}^+$. If $n = ab$ and $\gcd(a, b) = 1$, then $\varphi(n) = \varphi(a)\varphi(b)$.

Example 6.54. 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. Assume $n = ab$. Recall from Exercise 2.32 on page 82 that $\mathbb{Z}_a^* \times \mathbb{Z}_b^*$ is a group; the size of this group is $|\mathbb{Z}_a^*| \times \left|\mathbb{Z}_b^*\right| = \varphi(a)\varphi(b)$. We claim that $\mathbb{Z}_n^* \cong \mathbb{Z}_a^* \times \mathbb{Z}_b^*$. If true, this would prove the lemma, since

$$\varphi(n) = |\mathbb{Z}_n^*| = \left|\mathbb{Z}_a^* \times \mathbb{Z}_b^*\right| = |\mathbb{Z}_a^*| \times \left|\mathbb{Z}_b^*\right| = \varphi(a)\varphi(b).$$

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 $[x] \in \mathbb{Z}_n^*$ indeed, let d be any common divisor of x and n. By way of contradiction, assume $d \neq \pm 1$; by Theorem 6.30, we can find an irreducible divisor r of d; by Exercise 0.50 on page 23, $r \mid n$ 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 3.101 on page 139 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 6.44, $\gcd(y, a) = 1$. But the definition of greatest common divisor means that

$$\gcd(y,a) \geq r > 1 = \gcd(y,a),$$

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, $x \in \mathbb{Z}_n^*$. □

Fast exponentiation

Corollary 6.52 gives us an "easy" way to compute the inverse of any $x \in \mathbb{Z}_n^*$. However, it can take a long time to compute $x^{\varphi(n)}$, so let's take a moment to explain how we can compute canonical forms of exponents in this group more quickly. We will take two steps towards a fast exponentiation in $\mathbb{Z}_n^*$.

Lemma 6.55. For any $n \in \mathbb{N}^+$, $[x^a] = [x]^a$ in $\mathbb{Z}_n^*$.

Proof. You do it! See Exercise 6.63 on the next page. □

Example 6.56. In $\mathbb{Z}_{15}^*$ we can determine easily that $\left[4^{20}\right] = [4]^{20} = \left([4]^2\right)^{10} = [16]^{10} = [1]^{10} = [1]$. Notice that this is a *lot* faster than computing $4^{20} = 1099511627776$ and 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?

Theorem 6.57 (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 $2^b \leq a$.
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 $q_i \neq 0$ then replace y with the product of y and z.
   (b) Replace z with z^2.
   (c) Replace i with $i + 1$.

This ends with $x^a = y$.

Theorem 6.57 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 6.55 gives an added bonus in $\mathbb{Z}_n^*$, which is what we care about most.

Example 6.58. Since $10 = 2^3 + 2^1$, we can compute $\left[4^{10}\right]_7$ following the algorithm of Theorem 6.57:

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 $\left[4^{10}\right]_7 = [4]_7$. 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 $2^b \leq a$; if we need one more bit, then the definition of binary representation means that $2^{b+1} \leq x^a$, 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. □

Exercises.

Exercise 6.59. Compute 3^{28} in $\mathbb{Z}$ using fast exponentiation. Show each step.

Exercise 6.60. Compute 24^{28} in $\mathbb{Z}_7^*$ using fast exponentiation. Show each step.

Exercise 6.61. Prove that for all $x \in \mathbb{Z}_n^*$, $x^{\varphi(n)-1} = x^{-1}$.

Exercise 6.62. Prove that for all $x \in \mathbb{N}^+$, if x and n have no common divisors, then $n \mid \left(x^{\varphi(n)} - 1\right)$.

Exercise 6.63. Prove that for any $n \in \mathbb{N}^+$, $[x^a] = [x]^a$ in $\mathbb{Z}_n^*$.

6.6: RSA Encryption

From the viewpoint of practical applications, some of the most important results of group theory and number theory enable security in internet commerce. We described this problem on page 2: when you buy something online, you submit some private information, at least a credit card or bank account number, and usually more. There is no guarantee that, as this information passes through the internet, it 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.

Given the inherent insecurity of the internet, the solution is to disguise private information so that snoopers cannot understand it. A common method in use today is the RSA encryption algorithm.[24] First we describe the algorithms for encryption and decryption; afterwards we explain the ideas behind each stage, illustrating with an example; finally we prove that it succesfully encrypts and decrypts messages.

Description and example

Theorem 6.64 (RSA algorithm)**.** Let M be a list of positive integers. Let p,q be two irreducible integers such that:

- $\gcd(p,q) = 1$; and
- $(p-1)(q-1) > \max\{m : m \in M\}$.

Let $N = pq$, and let $e \in \mathbb{Z}^*_{\varphi(N)}$, where φ is the Euler phi-function. If we apply the following algorithm to M:

1. Let $e \in \mathbb{Z}^*_{\varphi(N)}$.
2. 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 6.65. 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 6.45 and Lemma 6.44 show that, yes, $e \in \mathbb{Z}^*_{\varphi(n)}$.

[24] RSA stands for Rivest (of MIT), Shamir (of the Weizmann Institute in Israel), and Adleman (of USC).

Encrypt by computing m^e for each $m \in M$:

$$C = \left(1^3, 12^3, 7^3, 5^3, 2^3, 18^3, 1^3, 0^3, 18^3, 21^3, 12^3, 26^3\right)$$
$$= (1, 23, 13, 15, 8, 2, 1, 0, 2, 21, 23, 31).$$

A snooper who intercepts C and tries to read it as a plain message would have several problems trying to read it. 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 6.52, but it's not hard to see that $3 \times 27 = 81$ and $[81]_{40} = [1]_{40}$.) Decrypt by computing c^d for each $c \in C$:

$$D = \left(1^{27}, 23^{27}, 13^{27}, 15^{27}, 8^{27}, 2^{27}, 1^{27}, 0^{27}, 2^{27}, 21^{27}, 23^{27}, 31^{27}\right)$$
$$= (1, 12, 7, 5, 2, 18, 1, 0, 18, 21, 12, 26).$$

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

$$\text{ALGE} \quad \rightarrow \quad 1 \times 26^3 + 12 \times 26^2 + 7 \times 26 + 5 = 25{,}785.$$

In order to encrypt this, we would 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 6.52 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 6.53 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!* To make things worse, there is a careful science to choosing p and q in such a way that 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. At the end of this section, after the exercises, we list programs that will 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);
    ```

Now, *why* does the RSA algorithm work?

Theory

Before reading the proof, let's reexamine the theorem.

Theorem (RSA algorithm)**.** Let M be a list of positive integers. Let p, q be two irreducible integers such that:

- $\gcd(p,q) = 1$; and
  - $(p-1)(q-1) > \max\{m : m \in M\}$.

Theorem. Let $N = pq$, and let $e \in \mathbb{Z}^*_{\varphi(N)}$, where φ is the Euler phi-function. If we apply the following algorithm to M:

1. Let $e \in \mathbb{Z}^*_{\varphi(N)}$.
   (a) Let C be a list of positive integers found by computing the canonical representation of $[m^e]_N$ for each $m \in M$.

Theorem. and subsequently apply the following algorithm to C:

1. Let $d = e^{-1} \in \mathbb{Z}^*_{\varphi(N)}$.
   (a) Let D be a list of positive integers found by computing the canonical representation of $\left[c^d\right]_N$ for each $c \in C$.

Theorem. then $D = M$.

Proof of the RSA algorithm. Let $i \in \{1,2,\ldots,|C|\}$. 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 3.101, $\varphi(N) \mid (1-de)$, 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 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 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 Exercise 6.68 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. □

Exercises.

Exercise 6.66. 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.

Exercise 6.67. 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.

Exercise 6.68. 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:
```

Part II

Rings

Chapter 7: Rings

While monoids are defined by one operation, groups are arguably defined by two: addition and subtraction, for example, or multiplication and division. The second operation is so closely tied to the first that we consider groups to have only one operation, for which (unlike monoids) every element has an inverse.

Of course, a set can be closed under more than one operation; for example, $\mathbb{Z}$ is closed under both addition and multiplication. As with subtraction, it is possible to define the multiplication of integers in terms of addition, just as we did with groups. However, this is not possible for all sets where an addition and a multiplication are both defined. Think of the multiplication of polynomials; how would you define $(x+1)(x-1)$ as repeated addition of $x-1$, a total of $x+1$ times? Does that even make sense? This motivates the study of a structure that incorporates common properties of two operations, which are related as loosely as possible.

Section 7.1 of this chapter introduces us to this structure, called a *ring*. A ring has two operations, "addition" and "multiplication". As you should expect from your experience with groups, what we call "addition" and "multiplication" may look nothing at all like the usual addition and multiplication of numbers. In fact, while the multiplication of integers has a natural definition from addition, multiplication in a ring may have absolutely nothing to do with addition, with one exception: the distributive property must still hold.

The rest of the chapter examines special kinds of rings. In Section 7.2 we introduce special kinds of rings that model useful properties of $\mathbb{Z}$ and $\mathbb{Q}$. In Section 7.3 we introduce rings of polynomials. The Euclidean algorithm, which proved so important in chapter 6, serves as the model for a special kind of ring described in Section 7.4.

7.1: A structure for addition and multiplication

What sort of properties do we associate with both addition and multiplication? We typically associate the properties of addition with an abelian group, and the properties of multiplication with a monoid, although it really depends on the set. The most basic properties of multiplication are encapsulated by the notion of a **semigroup**, which we defined way back in Definition 1.71 on page 70.

Definition 7.1. Let R be a set *with at least one element*, and $+$ and $\times$ two binary operations on that set. We say that $(R,+,\times)$ is a **ring** if it satisfies the following properties:

(R1) $(R,+)$ is an abelian group.
(R2) $(R,\times)$ is a semigroup.
(R4) R satisfies the distributive property of addition over multiplication: that is,
for all $a,b,c\in R$, $a(b+c)=ab+ac$ and $(a+b)c=ac+bc$.

Notation 7.2. As with monoids and groups, we usually refer simply to the ring as R, rather than $(R,+,\times)$. Since $(R,+)$ is an abelian group, the ring has an additive identity, 0. We sometimes write 0_R to emphasize that it is the additive identity of R.

Notice the following:

- While addition is commutative on account of (R1), multiplication need not be.
- There is no requirement that a multiplicative identity exists.
- There is no requirement that multiplicative inverses exist.
- There is no guarantee (yet) that the additive identity interacts with multiplication according to properties you have seen before. In particular, there is *no guarantee* that
  - the zero-product rule holds; or even that
  - $0_R \cdot a = 0_R$ for any $a \in R$.

Example 7.3. Let $R = \mathbb{R}^{m\times m}$ for some positive integer m. It turns out that R is a ring under the usual addition and multiplication of matrices. After all, Example 1.8 shows that the matrices satisfy the properties of a monoid under multiplication, and Example 2.6 shows that they are a group under addition, though most of the work was done in Section 0.3. The only part missing is distribution, and while that isn't hard, it is somewhat tedious, so we defer to your background in linear algebra.

However, we do want to point out something that should make you at least a *little* uncomfortable. Let

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Routine computation shows that

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

or in other words, $AB = 0$. This is true even though $A, B \neq 0$! Hence

Not every ring R satisfies the **zero product property**

$$\forall a, b \in R \qquad ab = 0 \quad \Longrightarrow \quad a = 0 \text{ or } b = 0.$$

Example 7.3 shouldn't surprise you that much; first, you've seen it in linear algebra, and second, you met zero divisors in Section 6.4. In fact, we will shortly generalize that idea into zero divisors for rings.

Likewise, the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, with which you are long familiar, are also rings. We omit the details, but you should think about them a little bit, and ask your instructor if some part of it isn't clear. You will study other example rings in the exercises. For now, we prove a familiar property of the additive identity.

Proposition 7.4. For all $r \in R$,

$$r \cdot 0_R = 0_R \cdot r = 0_R.$$

If you see that and ask, "Isn't that obvious?" then you *really* need to read the proof. While you read it, ask yourself, "What properties of a ring make this statement true?" The answer to that question will indicate your hidden assumptions. Try to prove the proposition without those properties, and you will see why it is *not* in fact obvious.

Proof. Let $r \in R$. Since $(R, +)$ is an abelian group, we know that $0_R + 0_R = 0_R$. By substitution, $r(0_R + 0_R) = r \cdot 0_R$. By distribution, $r \cdot 0_R + r \cdot 0_R = r \cdot 0_R$. Since $(R, +)$ is an abelian group,

$r \cdot 0_R$ has an additive inverse; call it s. Applying the properties of a ring, we have

$$\begin{aligned} s + (r \cdot 0_R + r \cdot 0_R) &= s + r \cdot 0_R && \text{(substitution)} \\ (s + r \cdot 0_R) + r \cdot 0_R &= s + r \cdot 0_R && \text{(associative)} \\ 0_R + r \cdot 0_R &= 0_R && \text{(additive inverse)} \\ r \cdot 0_R &= 0_R. && \text{(additive identity)} \end{aligned}$$

A similar argument shows that $0_R \cdot r = 0_R$. □

We now turn our attention to two properties that, while pleasant, are not necessary for a ring.

Definition 7.5. Let R be a ring. If R has a multiplicative identity 1_R such that

$$r \cdot 1_R \quad = \quad 1_R \cdot r \quad = \quad r \qquad \forall r \in R,$$

we say that R is a **ring with unity**. (Another name for the multiplicative identity is **unity**.)

If R is a ring and the multiplicative operation is commutative, so that

$$rs = sr \quad \forall r \in R,$$

then we say that R is a **commutative ring**.

A ring with unity is

- an abelian group under multiplication, and
- a (possibly commutative) monoid under addition.

Example 7.6. The set of matrices $\mathbb{R}^{m\times m}$ is a ring with unity, where I_m is the multiplicative identity. However, it is not a commutative ring.

You will show in Exercise 7.12 that $2\mathbb{Z}$ is a ring. It is a commutative ring, but not a ring with unity.

For a commutative ring with unity, consider $\mathbb{Z}$.

Remark 7.7. While non-commutative rings are interesting,

Unless we state otherwise,
all rings in these notes are commutative.

As with groups, we can characterize all rings with only two elements.

Example 7.8. Let R be a ring with only two elements. There are two possible structures for R.

Why? Since $(R, +)$ is an abelian group, by Example 2.11 on page 78 the addition table of R has the form

$+$	0_R	a
0_R	0_R	a
a	a	0_R

.

By Proposition 7.4, we know that the multiplication table *must* have the form

$\times$	0_R	a
0_R	0_R	0_R
a	0_R	?

where $a \cdot a$ is undetermined. Nothing in the properties of a ring tell us whether $a \cdot a = 0_R$ or $a \cdot a = a$; in fact, rings exist with both properties:

- if $R = \mathbb{Z}_2$ (see Exercise 7.13 to see that this is a ring), then $a = [1]$ and $a \cdot a = a$; but
- if
$$R = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right\} \subsetneq (\mathbb{Z}_2)^{2\times 2},$$
then $a \cdot a = 0 \neq a$.

Just as groups have subgroups, rings have subrings:

Definition 7.9. Let R be a ring, and S a nonempty subset of R. If S is also a ring under the same operations as R, then S is a subring of R.

Example 7.10. Recall from Exercise 7.12 that $2\mathbb{Z}$ is a ring; since $2\mathbb{Z} \subsetneq \mathbb{Z}$, it is a subring of $\mathbb{Z}$.

To show that a subset of a ring is a subring, do we have to show all four ring properties? No: as with subgroups, we can simplify the characterization to two properties:

Theorem 7.11 (The Subring Theorem)**.** Let R be a ring and S be a nonempty subset of R. The following are equivalent:
(A) S is a subring of R.
(B) S is closed under subtraction and multiplication: for all $a, b \in S$
  (S1) $a - b \in S$, and
  (S2) $ab \in S$.

Proof. That (A) implies (B) is clear, so assume (B). From (B) we know that for any $a, b \in S$ we have (S1) and (S2). As (S1) is essentially the Subgroup Theorem, S is an additive subgroup of the additive group R. On the other hand, (S2) only tells us that S satisfies property $(R2)$ of a ring, but any elements of S are elements of R, so the associative and distributive properties follow from inheritance. Thus S is a ring in its own right, which makes it a subring of R. □

Exercises

Exercise 7.12.
(a) Show that $2\mathbb{Z}$ is a ring under the usual addition and multiplication of integers.
(b) Show that for any $n \in \mathbb{Z}$, $n\mathbb{Z}$ is a ring under the usual addition and multiplication of integers.

Exercise 7.13. Recall the definition of multiplication for $\mathbb{Z}_n$ from Section 6.4: for $[a], [b] \in \mathbb{Z}_n$, $[a][b] = [ab]$.
(a) Show that $\mathbb{Z}_2$ is a ring under the addition and multiplication of cosets defined in Section 3.5.
(b) Show that for any $n \in \mathbb{N}^+$ where $n > 1$, $\mathbb{Z}_n$ is a ring under the addition and multiplication of cosets defined in Section 3.5.
(c) Show that there exist a, b, n such that $[a]_n [b_n] = [0]_n$ but $[a]_n, [b]_n \neq [0]_n$.

Exercise 7.14. Let R be a ring.

(a) Show that for all $r, s \in R$, $(-r)s = r(-s) = -(rs)$.
(b) Suppose that R has unity. Show that $-r = -1_R \cdot r$ for all $r \in R$.

Exercise 7.15. Let R be a ring with unity. Show that $1_R = 0_R$ if and only if R has only one element.

Exercise 7.16. Consider the two possible ring structures from Example 7.8. Show that if a ring R has only two elements, one of which is unity, then it can have only one of the structures.

Exercise 7.17. Let $R = \{T, F\}$ with the additive operation $\oplus$ (Boolean xor) and a multiplicative operation $\wedge$ (Boolean and where

$$\begin{aligned} F \oplus F &= F & \qquad F \wedge F &= F \\ F \oplus T &= T & F \wedge T &= F \\ T \oplus F &= T & T \wedge F &= F \\ T \oplus T &= F & T \wedge T &= T. \end{aligned}$$

(See also Exercises 2.28 and 2.29 on page 82.) Is $(R, \oplus, \wedge)$ a ring? If it is a ring, then
(a) what is the zero element?
(b) does it have a unity element? if so, what is it?
(c) is it commutative?

Exercise 7.18. Let R and S be rings, with $R \subseteq S$ and $\alpha \in S$. The **extension of** R **by** α is

$$R[\alpha] = \{r_n \alpha^n + \cdots + r_1 \alpha + r_0 : n \in \mathbb{N},\ r_0, r_1, \ldots, r_n \in R\}.$$

(a) Show that $R[\alpha]$ is also a ring.
(b) Suppose $R = \mathbb{Z}$, $S = \mathbb{C}$, and $\alpha = \sqrt{-5}$.
  (i) Explain why every element of $R[\alpha]$ can be written in the form $a + b\alpha$.
  (ii) Show that 6 can be factored two distinct ways in $R[\alpha]$: one is the ordinary factorization in $R = \mathbb{Z}$, while the other exploits the difference of squares with $\alpha = \sqrt{-5}$.

Exercise 7.19. In Exercise 7.13, you showed that $\mathbb{Z}_n$ is a ring. A nonzero element r of a ring R is **nilpotent** if we can find $n \in \mathbb{N}^+$ such that $r^n = 0_R$.
(a) Identify the nilpotent elements, *if any*, of $\mathbb{Z}_n$ for $n = 2, 3, 4, 5, 6$. If not, state that.
(b) Do you think there is a relationship between n and the nilpotents of $\mathbb{Z}_n$? If so, state it.

7.2: Integral Domains and Fields

In this section, R is always a commutative ring with unity.

Example 7.3 illustrates an important point: not all rings satisfy properties that we might like to take for granted. Not only does it show that not all rings possess the zero product property, it also demonstrates that multiplicative inverses do not necessarily exist in all rings. Both multiplicative inverses and the zero product property are very useful; we use them routinely to solve equations! Rings with these properties deserve special attention.

Two convenient kinds of rings

We first classify rings that satisfy the zero product property.

> **Definition 7.20.** If the elements of R satisfy the zero product property, then we call R an **integral domain**.

We use the word "integral" here because R is like the ring of "integ"ers, $\mathbb{Z}$. We do *not* mean that you can compute the integrals of calculus.

Whenever R is not an integral domain, we can find two elements of R that *do not* satisfy the zero product property; that is, we can find nonzero $a, b \in R$ such that $ab = 0_R$. Recall that we used a special term for this phenomenon in the group $\mathbb{Z}_n^*$, **zero divisors** (Section 6.4). The ideas are identical, so the term is appropriate, and we will call a and b **zero divisors** in a ring, as well.

Example 7.21. As you might expect, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are integral domains.

In Exercise 7.13, you showed that $\mathbb{Z}_n$ was a ring under clockwork addition and multiplication. However, it need not be an integral domain. For example, in $\mathbb{Z}_6$ we have $[2] \cdot [3] = [6] = [0]$, making $[2]$ and $[3]$ zero divisors. On the other hand, it isn't hard to see that $\mathbb{Z}_2$, $\mathbb{Z}_3$, and $\mathbb{Z}_5$ are integral domains, if only via an exhaustive check. What about $\mathbb{Z}_4$? We leave that, and all of $\mathbb{Z}_n$ to the exercises.

Next, we turn to multiplicative inverses.

> **Definition 7.22.** If every non-zero element of R has a multiplicative inverse, then we call R a **field**.

Example 7.23. The rings $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are fields, while $\mathbb{Z}$ is not.

What about $\mathbb{Z}_n$ and $\mathbb{Z}_n^*$? Again, we leave those to the exercises. For now, we need to notice an important relationship between fields and integral domains.

The examples show that some integral domains are not fields, but all the fields we've listed are also integral domains. It would be great if this turned out to be true in general: that is, if every field is an integral domain. Determining the relationships between different classes of rings, and remembering which class you're working with, is a crucial point of ring theory.

> **Theorem 7.24.** Every field is an integral domain.

Proof. Let $\mathbb{F}$ be a field. We claim that $\mathbb{F}$ is an integral domain: that is, the elements of $\mathbb{F}$ satisfy the zero product property. Let $a, b \in \mathbb{F}$ and assume that $ab = 0$. We need to show that $a = 0$ or $b = 0$. If $a = 0$, we're done, so assume that $a \neq 0$. Since $\mathbb{F}$ is a field, a has a multiplicative inverse. Apply Proposition 7.4 to obtain

$$b = 1 \cdot b = \left(a^{-1}a\right) b = a^{-1}(ab) = a^{-1} \cdot 0 = 0.$$

Hence $b = 0$.

We had assumed that $ab = 0$ and $a \neq 0$. By concluding that $b = 0$, the fact that a and b are arbitrary show that $\mathbb{F}$ is an integral domain. Since $\mathbb{F}$ is an arbitrary field, every field is an integral domain. □

Not every integral domain is a field, however. The most straightforward example is $\mathbb{Z}$.

The field of fractions

Speaking of $\mathbb{Q}$, it happens to be the smallest field that contains $\mathbb{Z}$, an integral domain. So there's another interesting question: can we form a field from any ring R, simply by adding fractions?

No, of course not — we just saw that a field must be an integral domain, and some rings are not integral domains. Even if you add fractions, the zero divisors remain, so you cannot have a field. So, then, can we form a field from any integral domain in the same way that we form $\mathbb{Q}$ from $\mathbb{Z}$? We need some precision in this discussion, which requires a definition.

Definition 7.25. Let R be an arbitrary ring. **The set of fractions over** a ring R is

$$\mathrm{Frac}(R) := \left\{\frac{p}{q} : p, q \in R \text{ and } q \neq 0\right\},$$

with addition and multiplication defined in the usual way for "fractions", and equality defined by

$$\frac{a}{b} = \frac{p}{q} \iff aq = bp.$$

The answer to our question turns out to be yes!

Theorem 7.26. If R is an integral domain, then $\mathrm{Frac}(R)$ is a ring.

To prove Theorem 7.26, we need two useful properties of fractions that you should be able to prove yourself.

Proposition 7.27. Let R be a ring, $a, b, r \in R$. If $br \neq 0$, then in $\mathrm{Frac}(R)$

- $\dfrac{a}{b} = \dfrac{ar}{br}$, and
- $\dfrac{0_R}{a} = \dfrac{0_R}{b}$.

Proof. You do it! See Exercise 7.33. □

Watch for these properties in what follows.

Proof of Theorem 7.26. Assume that R is an integral domain. First we show that $\mathrm{Frac}(R)$ is an additive group. Let $f, g, h \in R$; choose $a, b, p, q, r, s \in \mathrm{Frac}(R)$ such that $f = a/b$, $g = p/q$, and $h = r/s$. First we show that $\mathrm{Frac}(R)$ is an abelian group.

closure: This is fairly routine, using common denominators. Since R is a domain and $b, q \neq 0$, we know that $bq \neq 0$. Thus,

$$\begin{aligned} f + g &= \frac{a}{b} + \frac{p}{q} && \text{(substitution)} \\ &= \frac{aq}{bq} + \frac{bp}{bq} && \text{(Proposition 7.27)} \\ &= \frac{aq + bp}{bq} && \text{(definition of addition in Frac}(R)\text{)} \\ &\in \operatorname{Frac}(R). \end{aligned}$$

Why did we need R do be an integral domain? If not, then it is possible that $bq = 0$, and if so, $f + g \notin \operatorname{Frac}(R)$!

associative: This is the hardest one; watch for Proposition 7.27 to show up in many places. As before, since R is a domain and $b, q, s \neq 0$, we know that bq, $(bq)s$, $b(qs)$, and qs are all non-zero. Thus,

$$\begin{aligned} (f + g) + h &= \frac{aq + bp}{bq} + \frac{r}{s} \\ &= \frac{(aq + bp)s}{(bq)s} + \frac{(bq)r}{(bq)s} \\ &= \frac{((aq)s + (bp)s) + (bq)r}{(bq)s} \\ &= \frac{a(qs) + (b(ps) + b(qr))}{b(qs)} \\ &= \frac{a(qs)}{b(qs)} + \frac{b(ps) + b(qr)}{b(qs)} \\ &= \frac{a}{b} + \frac{ps + qr}{qs} \\ &= \frac{a}{b} + \left(\frac{p}{q} + \frac{r}{s}\right) \\ &= f + (g + h.) \end{aligned}$$

identity: We claim that the additive identity of $\operatorname{Frac}(R)$ is $0_R/1_R$. This is easy to see, since

$$f + \frac{0_R}{1_R} = \frac{a}{b} + \frac{0_R \cdot b}{1_R \cdot b} = \frac{a}{b} + \frac{0_R}{b} = \frac{a}{b} = f.$$

additive inverse: For each $f = p/q$, we claim that $(-p)/q$ is the additive inverse. This is easy to see, but a little tedious. It is straightforward enough that,

$$f + \frac{-p}{q} = \frac{p}{q} + \frac{-p}{q} = \frac{(p + (-p))}{q} = \frac{0_R}{q}.$$

Don't conclude too quickly that we are done! We have to show that $f + (-p)/q =$

$0_{\text{Frac}(R)}$, which is $0_R/1_R$. By Proposition 7.27, $0_R/1_R = 0_R/q_R$, so we did in fact compute the identity.

commutative: Using the fact that R is commutative, we have

$$\begin{aligned} f+g &= \frac{a}{b}+\frac{c}{d} = \frac{ad}{bd}+\frac{bc}{bd} \\ &= \frac{ad+bc}{bd} = \frac{cb+da}{db} \\ &= \frac{cb}{db}+\frac{da}{db} = \frac{c}{d}+\frac{a}{b} \\ &= g+f. \end{aligned}$$

Next we have to show that Frac (R) satisfies the requirements of a ring.

closure: Using closure in R and the fact that R is an integral domain, this is straightforward: $fg = (ap)/(bq) \in \text{Frac}(R)$.

associative: Using the associative property of R, this is straightforward:

$$\begin{aligned} (fg)h &= \left(\frac{ap}{bq}\right)\frac{r}{s} = \frac{(ap)r}{(bq)s} = \frac{a(pr)}{b(qs)} \\ &= \frac{a}{b}\frac{(pr)}{qs} = f(gh). \end{aligned}$$

distributive: We rely on the distributive property of R:

$$\begin{aligned} f(g+h) &= \frac{a}{b}\left(\frac{p}{q}+\frac{r}{s}\right) = \frac{a}{b}\left(\frac{ps+qr}{qs}\right) \\ &= \frac{a(ps+qr)}{b(qs)} = \frac{a(ps)+a(qr)}{b(qs)} \\ &= \frac{a(ps)}{b(qs)}+\frac{a(qr)}{b(qs)} = \frac{ap}{bq}+\frac{ar}{bs} \\ &= fg+fh. \end{aligned}$$

Finally, we show that Frac (R) is a field. We have to show that it is commutative, that it has a multiplicative identity, and that every non-zero element has a multiplicative inverse.

commutative: We claim that the multiplication of Frac (R) is commutative. This follows from the fact that R, as an integral domain, has a commutative multiplication, so

$$fg = \frac{a}{b}\cdot\frac{p}{q} = \frac{ap}{bq} = \frac{pa}{qb} = \frac{p}{q}\cdot\frac{a}{b} = gf.$$

multiplicative identity: We claim that $\frac{1_R}{1_R}$ is a multiplicative identity for $\text{Frac}(R)$. In fact,

$$f\cdot\frac{1_R}{1_R} = \frac{a}{b}\cdot\frac{1_R}{1_R} = \frac{a\cdot 1_R}{b\cdot 1_R} = \frac{a}{b} = f.$$

multiplicative inverse: Let $f \in \text{Frac}(R)$ be a non-zero element. Let $a, b \in R$ such that $f = a/b$ and $a \neq 0$. Let $g = b/a$; then

$$fg = \frac{a}{b}\cdot\frac{b}{a} = \frac{ab}{ab}.$$

By Proposition 7.27

$$\frac{ab}{ab} = \frac{1_R}{1_R},$$

which we just showed to be the identity of $\text{Frac}(R)$.

□

Definition 7.28. For any integral domain R, we call $\text{Frac}(R)$ the **field of fractions of R.**

Exercises.

Exercise 7.29. Explain why $n\mathbb{Z}$ is not always an integral domain. For what values of n is it an integral domain?

Exercise 7.30. Is the boolean ring of Exercise 7.17 an integral domain?

Exercise 7.31. Show that $\mathbb{Z}_n$ is an integral domain if and only if n is irreducible. Is it also a field in these cases?

Exercise 7.32. You might think from Exercise 7.31 that we can turn $\mathbb{Z}_n$ into a field, or at least an integral domain, in the same way that we turned $\mathbb{Z}_n$ into a multiplicative group: that is, working with $\mathbb{Z}_n^*$. Explain that this doesn't work in general, because $\mathbb{Z}_n^*$ isn't even a ring.

Exercise 7.33. Show that if R is an integral domain, then the set of fractions has the following properties for any nonzero $a, b, c \in R$:

$$\frac{ac}{bc} = \frac{ca}{cb} = \frac{a}{b}, \qquad \frac{0_R}{a} = \frac{0_R}{1} = 0_{\text{Frac}(R)},$$
$$\text{and} \qquad \frac{a}{a} = \frac{1_R}{1_R} = 1_{\text{Frac}(R)}.$$

Exercise 7.34. To see concretely why Frac(R) is not a field if R is not a domain, consider $R = \mathbb{Z}_4$. Find nonzero $b, q \in R$ such that $bq = 0$, using them to find $f, g \in \text{Frac}(R)$ such that $fg \notin \text{Frac}(R)$.

7.3: Polynomial rings

When the average man on the street thinks of "algebra", he typically thinks not of "monoids", "groups", or "rings", but of "polynomials". Polynomials are certainly the focus of high school algebra, and they are also a major focus of higher algebra. The last few chapters of these notes are dedicated to the classical applications of the structural theory to important problems about polynomials.

While one can talk of a monoid or group of polynomials under addition, it is more natural to talk about a ring of polynomials under addition and multiplication. Polynomials helped motivate the distinction between the "two operations" of groups, which we decided was really two sides of one coin, and the "two operations" of rings, which really can be quite different operations. Polynomials provide great examples for the remaining topics. It is time to give them a good, hard look.

Some of the following may seem pedantic and needlessly detailed, and there's some truth to that, but it is important to fix these terms now to avoid confusion later. The difference between a "monomial" and a "term" is of special note; some authors reverse the notions. Similarly, pay attention to the notion of the support $\mathcal{T}_f$ of a polynomial f.

As usual, R is a ring.

Fundamental notions

Definition 7.35. An **indeterminate over** R is a symbol that represents an unknown value of R. A **constant of** R is a symbol that represents a fixed value of R. An **variable over** R is an indeterminate whose value is *not* fixed.

Notice that a constant can be indeterminate, as in the usual use of letters like a, b, and c, or quite explicitly determined, as in 1_R, 0_R, and so forth. Variables are always indeterminate. The main difference is that a constant is *fixed,* while a variable is not.

Definition 7.36. A **monomial over** R is a finite product of variables over R.

The use of "monomial" here is meant to be both consistent with its definition in Section 1.1, and with our needs for future work. Typically, though, we refer simply to "a monomial" rather than "a monomial over R".

By referring to "variables", the definition of a monomial explicitly excludes constants. Even though a^2 looks like a monomial, if a is a constant, we do not consider it a monomial; from our point of view, it is a constant.

Definition 7.37. The **total degree** of a monomial is the number of factors in the product. We say that two monomials are **like monomials** if they have the same variables, and corresponding variables have the same exponents.

A **term of** R is a constant, *or* the product of a monomial over R and a constant of R. The constant in a term is called the **coefficient** of the term. Two terms are **like terms** if their monomials are like monomials.

Now we define *polynomials*.

Definition 7.38. A **polynomial over** R is a finite sum of terms of R. We can write a generic polynomial f as $f = a_1t_1 + a_2t_2 + \cdots + a_mt_m$ where each $a_i \in R$ and each t_i is a monomial.

We call the set of monomials of f with non-zero coefficient its **support**. If we denote the support of f by $\mathcal{T}_f$, then we can write f as

$$f = \sum_{i=1,\dots,\#\mathcal{T}_f} a_i t_i = \sum_{t\in\mathcal{T}_f} a_t t.$$

We call R the **ground ring** of each polynomial.

We say that two polynomials f and g are equal if $\mathcal{T}_f = \mathcal{T}_g$ and the coefficients of corresponding monomials are equal.

Notation 7.39. We adopt a convention that $\mathcal{T}_f$ is the support of a polynomial f.

Definition 7.40. $R[x]$ is the set of **univariate** polynomials in the variable x over R. That is, $f \in R[x]$ if and only if there exist $m \in \mathbb{N}$ and $a_m, a_{m-1}, \dots, a_1 \in R$ such that

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0.$$

The set $R[x,y]$ is the set of **bivariate** polynomials in the variables x and y whose coefficients are in R.

For $n \geq 2$, the set $R[x_1, x_2, \dots, x_n]$ is the set of **multivariate** polynomials in the variables $x_1, x_2, \dots, x_n$ whose coefficients are in R.

The **degree** of a univariate polynomial f, written $\deg f$, is the largest of the total degrees of the monomials of f. We write $\mathrm{lm}(f)$ for the monomial of f with that degree, and $\mathrm{lc}(f)$ for its coefficient. Unless we say otherwise, the degree of a multivariate polynomial is undefined.

Example 7.41. Definition 7.40 tells us that $\mathbb{Z}_6[x,y]$ is the set of bivariate polynomials in x and y whose coefficients are in $\mathbb{Z}_6$. For example,

$$f(x,y) = 5x^3 + 2x \in \mathbb{Z}_6[x,y]$$

and

$$g(x,y) = x^2y^2 - 2x^3 + 4 \in \mathbb{Z}_6[x,y].$$

The ground ring for both f and g is $\mathbb{Z}_6$. Observe that f can be considered a univariate polynomial, in which case $\deg f = 3$.

We also consider constants to be polynomials of degree 0; thus $4 \in \mathbb{Z}_6[x,y]$ and even $0 \in \mathbb{Z}_6[x,y]$.

It is natural to think of a constant as a polynomial. This leads to some unexpected, but interesting and important consequences.

Definition 7.42. Let $f \in R[x_1,\ldots,x_n]$.

We say that f is a **constant polynomial** if $\mathcal{T}_f = \{1\}$ or $\mathcal{T}_f = \emptyset$; in other words, all the non-constant terms have coefficient zero.

We say that f is a **vanishing polynomial** if for all $r_1,\ldots,r_n \in R$, $f(r_1,\ldots,r_n) = 0$. We will see that this can happen even if $f \neq 0_R$.

The definition of vanishing and constant polynomials implies that 0_R satisfies both. However, the definition of equality means that vanishing polynomials need not be zero polynomials!

Example 7.43. Let $f(x) = x^2 + x \in \mathbb{Z}_2[x]$. Since $\mathcal{T}_f \neq \emptyset$, $f \neq 0_R$. However,

$$\begin{aligned} f(0) &= 0^2 + 0 && \text{and} \\ f(1) &= 1^2 + 1 = 0 && (\text{in } \mathbb{Z}_2!). \end{aligned}$$

Here f is a vanishing polynomial *even though it is not zero.*

Properties of polynomials

We can now turn our attention to the properties of $R[x]$ and $R[x_1,\ldots,x_n]$. First up is a question raised by Example 7.43: when must a vanishing polynomial be the constant polynomial 0?

Proposition 7.44. If R is a non-zero integral domain, then the following are equivalent.

(A) 0 is the only vanishing polynomial in $R[x_1,\ldots,x_n]$.

(B) R has infinitely many elements.

As is often the case, we can't answer that question immediately. Before proving Proposition 7.44, we need the following, extraordinary theorem.

Theorem 7.45 (The Factor Theorem)**.** If R is a non-zero integral domain, $f \in R[x]$, and $a \in R$, then $f(a) = 0$ if and only if $x - a$ divides $f(x)$.

To prove Theorem 7.45, we need to make precise our notions of addition and multiplication of polynomials.

Definition 7.46. To **add** two polynomials $f, g \in R[x_1, \ldots, x_n]$, let $\mathcal{T} = \mathcal{T}_f \cup \mathcal{T}_g$. Choose $a_t, b_t \in R$ such that

$$f = \sum_{t \in \mathcal{T}} a_t t \quad \text{and} \quad g = \sum_{t \in \mathcal{T}} b_t t.$$

We add the polynomials by adding like terms; that is,

$$f + g = \sum_{t \in \mathcal{T}} (a_t + b_t) t.$$

To **multiply** f and g, compute the sum of all products of terms in the first polynomial with terms in the second; that is,

$$fg = \sum_{t \in \mathcal{T}} \sum_{u \in \mathcal{T}} (a_t b_u)(tu).$$

We use u in the second summand to distinguishes the terms of g from those of f. Notice that fg is really the distribution of g to the terms of f, followed by the distribution of each term of f to the terms of g.

Proof of the Factor Theorem. If $x - a$ divides $f(x)$, then there exists $q \in R[x]$ such that $f(x) = (x - a) \cdot q(x)$. By substitution, $f(a) = (a - a) \cdot q(a) = 0_R \cdot q(a) = 0_R$.

Conversely, assume $f(a) = 0$. You will show in Exercise 7.49 that we can write $f(x) = q(x) \cdot (x - a) + r$ for some $r \in R$. Thus

$$0 = f(a) = q(a) \cdot (a - a) + r = r,$$

and substitution yields $f(x) = q(x) \cdot (x - a)$. In other words, $x - a$ divides $f(x)$, as claimed. □

We now turn our attention to proving Proposition 7.44.

Proof of Lemma 7.44. Assume that R is a non-zero integral domain.

(A) ⇒ (B): We proceed by the contrapositive. Assume that R has finitely many elements. We can enumerate them all as $r_1, r_2, \ldots, r_m$. Let

$$f(x_1, \ldots, x_n) = (x_1 - r_1)(x_1 - r_2) \cdots (x_1 - r_m).$$

Let $b_1, \ldots, b_n \in R$. By assumption, R is finite, so $b_1 = r_i$ for some $i \in \{1, 2, \ldots, m\}$. Notice that f is not only multivariate, it is also univariate: $f \in R[x_i]$. By the Factor Theorem, $f = 0$. We have shown that ¬(B) implies ¬(A); thus, (A) implies (B).

(A) ⇐ (B): Assume that R has infinitely many elements. Let f be any vanishing polynomial. We proceed by induction on n, the number of variables in $R[x_1, \ldots, x_n]$.

Inductive base: Suppose $n = 1$. By the Factor Theorem, $x - a$ divides f for every $a \in R$. By definition of polynomial multiplication, each distinct factor of f adds 1 to the degree of f; for example, if $f = (x - 0)(x - 1)$, then $\deg f = 2$. However, the definition of a polynomial implies that f has finite degree. Hence, if $f \neq 0$, then it can be factored as only finitely many polynomials of the form $x-a$. If so, then choose $a_1, a_2, \ldots, a_n$ such that

$$f = (x-a_1)(x-a_2)\cdots(x-a_n).$$

Since R has infinitely many elements, we can find $b \in R$ such that $b \neq a_1, \ldots, a_n$. That means $b-a_i \neq 0$ for each $i = 1, \ldots, n$. As R is an integral domain,

$$f(b) = (b-a_1)(b-a_2)\cdots(b-a_n) \neq 0.$$

This contradicts the choice of f as a vanishing polynomial. Hence, $f = 0$.

Inductive hypothesis: Assume for all i satisfying $1 \leq i < n$, if $f \in R[x_1, \ldots, x_i]$ is a zero polynomial, then f is the constant polynomial 0.

Inductive step: Let $n > 1$, and $f \in R[x_1, \ldots, x_n]$ be a vanishing polynomial. Let $a_n \in R$, and substitute $x_n = a_n$ into f. Denote the resulting polynomial as g. The substitution means that $x_n \notin \mathcal{T}_g$. Hence, $g \in R[x_1, \ldots, x_{n-1}]$.

It turns out that g is also a vanishing polynomial in $R[x_1, \ldots, x_{n-1}]$. *Why?* By way of contradiction, assume that it is not. Then there exist $a_1, \ldots, a_{n-1} \in R$ such that $f(a_1, \ldots, a_{n-1}) \neq 0$. However, the definition of g implies that

$$f(a_1, \ldots, a_n) = g(a_1, \ldots, a_{n-1}) \neq 0.$$

This contradicts the choice of f as a vanishing polynomial. The assumption was wrong; g must be a vanishing polynomial in $R[x_1, \ldots, x_{n-1}]$, after all. We can now apply the inductive hypothesis, and infer that g is the constant polynomial 0.

We chose a_n arbitrarily, so this argument holds for any $a_n \in R$. Thus, any of the terms of f containing any of the variables $x_1, \ldots, x_{n-1}$ has a coefficient of zero. The only non-zero terms are those whose only variables are x_n, so $f \in R[x_n]$. This time, the inductive base implies that f is zero. □

We come to the main purpose of this section.

Theorem 7.47. The univariate and multivariate polynomial rings over a ring R are themselves rings.

Proof. Let $n \in \mathbb{N}^+$ and R a ring. We claim that $R[x_1, \ldots, x_n]$ is a ring. To consider the requirements of a ring, let $f, g, h \in R[x_1, \ldots, x_n]$, and let $\mathcal{T} = \mathcal{T}_f \cup \mathcal{T}_g \cup \mathcal{T}_h$. For each $t \in \mathcal{T}$, choose $a_t, b_t, c_t \in R$ such that

$$f = \sum_{t \in \mathcal{T}} a_t t, \quad g = \sum_{t \in \mathcal{T}} b_t t, \quad h = \sum_{t \in \mathcal{T}} c_t t.$$

(Naturally, if $t \in \mathcal{T} \backslash \mathcal{T}_f$, then $a_t = 0$; if $t \in \mathcal{T} \backslash \mathcal{T}_g$, then $b_t = 0$, and if $t \in \mathcal{T} \backslash \mathcal{T}_h$, then $c_t = 0$.) Although we do not write it, all the sums below are indexed over $t \in \mathcal{T}$.

(R1) First we show that $R[x_1, \ldots, x_n]$ is an abelian group.

(closure) By the definition of polynomial addition,

$$(f+g)(x) = \sum (a_t + b_t) t.$$

Since R is closed under addition, we conclude that $f+g \in R[x_1,\ldots,x_n]$.

(associative) We rely on the associativity of R:

$$\begin{aligned} f+(g+h) &= \sum a_t t + \left(\sum b_t t + \sum c_t t\right) \\ &= \sum a_t t + \sum (b_t + c_t)\, t \\ &= \sum [a_t + (b_t + c_t)]\, t \\ &= \sum [(a_t + b_t) + c_t]\, t \\ &= \sum (a_t + b_t)\, t + \sum_{t\in T} c_t t \\ &= \left(\sum a_t t + \sum b_t t\right) + \sum c_t t \\ &= (f+g)+h. \end{aligned}$$

(identity) We claim that the constant polynomial 0 is the identity. Recall that 0 is a polynomial whose coefficients are all 0. We have

$$\begin{aligned} f+0 &= \sum a_t t + 0 \\ &= \sum a_t t + \sum 0\cdot t \\ &= \sum (a_t + 0)\, t \\ &= f. \end{aligned}$$

(inverse) Let $p = \sum(-a_t)\, t$. We claim that p is the additive inverse of f. In fact,

$$\begin{aligned} p+f &= \sum (-a_t)\, t + \sum a_t t \\ &= \sum (-a_t + a_t)\, t \\ &= \sum 0\cdot t \\ &= 0. \end{aligned}$$

(commutative) By the definition of polynomial addition, $g+f = \sum (b_t + a_t)\, t$. Since R is commutative under addition, addition of coefficients is commutative, so

$$\begin{aligned} f+g &= \sum a_t t + \sum b_t t \\ &= \sum (a_t + b_t)\, t \\ &= \sum (b_t + a_t)\, t \\ &= \sum b_t t + \sum a_t t \\ &= g+f. \end{aligned}$$

(R2) Next, we show that $R[x_1,\ldots,x_n]$ is a semigroup.

(closed) Applying the definition of polynomial multiplication, we have

$$fg = \sum_{t\in T}\left[\sum_{u\in T}(a_t b_u)(tu)\right].$$

Since R is closed under multiplication, each $(a_t b_u)(tu)$ is a term. Thus fg is a sum of sums of terms, or a sum of terms. In other words, $fg \in R[x_1, \ldots, x_n]$.

(associative) We start by applying the product fg, then multiplying the result to h:

$$\begin{aligned}(fg)h &= \left[\sum_{t\in T}\left[\sum_{u\in T}(a_t b_u)(tu)\right]\right]\cdot\sum_{v\in T}c_v v\\ &= \sum_{t\in T}\left[\sum_{u\in T}\left[\sum_{v\in T}[(a_t b_u)c_v][(tu)v]\right]\right].\end{aligned}$$

Now apply the associative property of multiplication in R:

$$(fg)h = \sum_{t\in T}\left[\sum_{u\in T}\left[\sum_{v\in T}[a_t(b_u c_v)][t(uv)]\right]\right].$$

(Notice the associative property of R applies to terms over R, as well, inasmuch as those terms represent undetermined elements of R.) Now unapply the product:

$$\begin{aligned}(fg)h &= \sum_{t\in T}\left[\sum_{u\in T}\left[\sum_{v\in T}[a_t(b_u c_v)][t(uv)]\right]\right]\\ &= \sum_{t\in T}a_t t\cdot\left[\sum_{u\in T}\left[\sum_{v\in T}(b_u c_v)(uv)\right]\right]\\ &= f(gh).\end{aligned}$$

(R3) To show the distributive property, first apply addition, then multiplication:

$$\begin{aligned}f(g+h) &= \sum_{t\in T}a_t t\cdot\left(\sum_{u\in T}b_u u + \sum_{u\in T}c_u u\right)\\ &= \sum_{t\in T}a_t t\cdot\sum_{u\in T}(b_u + c_u)u\\ &= \sum_{t\in T}\left[\sum_{u\in T}[a_t(b_u + c_u)](tu)\right].\end{aligned}$$

Now apply the distributive property in the ring, and unapply the addition and multiplication:

$$\begin{aligned}
f(g+h) &= \sum_{t\in T}\left[\sum_{u\in T}(a_t b_u + a_t c_u)(tu)\right]\\
&= \sum_{t\in T}\left[\sum_{u\in T}[(a_t b_u)(tu) + (a_t c_u)(tu)]\right]\\
&= \sum_{t\in T}\left[\sum_{u\in T}(a_t b_u)(tu) + \sum_{u\in T}(a_t c_u)(tu)\right]\\
&= \sum_{t\in T}\left[\sum_{u\in T}(a_t b_u)(tu)\right] + \sum_{t\in T}\left[\sum_{u\in T}(a_t c_u)(tu)\right]\\
&= fg + fh.
\end{aligned}$$

(commutative) Since we are working in commutative rings, we must also show that that $R[x_1,\ldots,x_n]$ is commutative. This follows from the commutativity of R:

$$\begin{aligned}
fg &= \left(\sum_{t\in T} a_t t\right)\left(\sum_{u\in T} b_u u\right)\\
&= \sum_{t\in T}\sum_{u\in T}(a_t b_u)(tu)\\
&= \sum_{u\in T}\sum_{t\in T}(b_u a_t)(ut)\\
&= gf.
\end{aligned}$$

(We can swap the sums because of the commutative and associative properties of addition.)

□

Exercises.

Exercise 7.48. Let $f(x) = x$ and $g(x) = x+1$ in $\mathbb{Z}_2[x]$.
(a) Show that f and g are not vanishing polynomials.
(b) Compute the polynomial $p = fg$.
(c) Show that $p(x)$ is a vanishing polynomial.
(d) Explain why this does *not* contradict Proposition 7.44.

Exercise 7.49. Fill in each blank of Figure 7.1 with the justification.

Exercise 7.50. Pick at random a degree 5 polynomial f in $\mathbb{Z}[x]$. Then pick at random some $a \in \mathbb{Z}$.
(a) Find $q \in \mathbb{Z}[x]$ and $r \in \mathbb{Z}$ such that $f(x) = q(x)\cdot(x-a)+r$.
(b) Explain why you *cannot* pick a nonzero integer b at random and expect willy-nilly to find $q \in \mathbb{Z}[x]$ and $r \in \mathbb{Z}$ such that $f(x) = q(x)\cdot(bx-a)+r$.
(c) Explain why you *can* pick a nonzero integer b at random and expect willy-nilly to find $q \in \mathbb{Z}[x]$ and $r,s \in \mathbb{Z}$ such that $s\cdot f(x) = q(x)\cdot(bx-a)+r$. (Neat, huh?)

Let R be an integral domain, $f \in R[x]$, and $a \in R$.
Claim: There exist $q \in R[x]$ and $r \in R$ such that $f(x) = q(x) \cdot (x - a) + r$.
Proof:

1. Without loss of generality, we may assume that $\deg f = n$.
2. By _____, choose $a_1, \ldots, a_n$ such that $f = \sum_{k=1}^{n} a_k x^k$. We proceed by induction on n.
3. For the *inductive base*, assume that $n = 0$. Then $q(x) =$_____ and $r =$_____.
4. For the *inductive hypothesis*, assume that for all $i \in \mathbb{N}$ satisfying $0 \leq i < n$, there exist $q \in R[x]$ and $r \in R$ such that $f(x) = q(x) \cdot (x - a) + r$.
5. For the *inductive step*,
   (a) Let $p(x) = a_n x^{n-1}$, and $g(x) = f(x) - p(x) \cdot (x - a)$.
   (b) Notice that $\deg g <$_____.
   (c) By _____, there exist $p' \in R[x]$ and $r \in R$ such that $g(x) = p'(x) \cdot (x - a) + r$.
   (d) Let $q = p + p'$. By _____, $q \in R[x]$.
   (e) By _____ and _____, $f(x) = q(x) \cdot (x - a) + r$.
6. We have shown that, for arbitrary n, we can find $q \in R[x]$ and $r \in R$ such that $f(x) = q(x) \cdot (x - a) + r$. The claim holds.

Figure 7.1. Material for Exercise 7.49

(d) If the requirements of (b) were changed to finding $q \in \mathbb{Q}[x]$ and $r \in \mathbb{Q}$, would you then be able to carry out (b)? Why or why not?

Exercise 7.51. Let $R = \mathbb{Z}_3[x]$ and $f(x) = x^3 + 2x + 1 \in R$.
(a) Explain how we can infer that f does not factor in R without performing a brute force search of factorizations.
(b) If we divide $g \in R$ by f, how many possible remainders can we obtain?

Exercise 7.52. Show that $x^4 + x^2 + 1$ factors in $\mathbb{Z}_2$, even though it has no roots. Explain how the Factor Theorem can apply to the polynomial of Exercise 7.51, but not to this one.

Exercise 7.53. Let R be an integral domain.
(a) Show that $R[x]$ is also an integral domain.
(b) How does this not contradict Exercise 7.48? After all, $\mathbb{Z}_2$ is a field, and thus an integral domain!

Exercise 7.54. Let R be a ring, and $f, g \in R[x]$. Show that $\deg(f + g) \leq \max(\deg f, \deg g)$.

Exercise 7.55. Let R be a ring and define

$$R(x) = \operatorname{Frac}(R[x]);$$

for example,

$$\mathbb{Z}(x) = \operatorname{Frac}(\mathbb{Z}[x]) = \left\{ \frac{p}{q} : p, q \in \mathbb{Z}[x] \right\}.$$

Is $R(x)$ a ring? is it a field?

Exercise 7.56. Let $R = \mathbb{Q}\left[\sqrt{2}\right]$, an extension of $\mathbb{Q}$ by $\sqrt{2}$. (See Exercise 7.18.)

(a) Find $g \in \mathbb{Q}[x]$ such that g factors with coefficients in R, but not with coefficients in $\mathbb{Q}$.
(b) Let $S = \mathbb{Q}\left[\sqrt{2}+\sqrt{3}\right]$ and $T = R\left[\sqrt{3}\right]$. Show that $S = T$.
(c) Is $\mathbb{Z}\left[\sqrt{2}+\sqrt{3}\right] = \mathbb{Z}\left[\sqrt{2}\right]\left[\sqrt{3}\right]$?

Exercise 7.57. Let $p \in \mathbb{Z}$ be irreducible, and $R = \mathbb{Z}_p[x]$. Show that $\varphi : R \to R$ by $\varphi(f) = f^p$ is a group automorphism. This is called the **Frobenius automorphism**.

7.4: Euclidean domains

In this section we consider an important similarity between the ring of integers and the ring of polynomials. This similarity will motivate us to define a new kind of ring. We will then show that all rings of this type allow us to perform important operations that we find both useful and necessary. What is the similarity? The ability to *divide with remainder*.

Division of polynomials

We start with polynomials, but we will take this a step higher in a moment.

Theorem 7.58 (The Division Theorem for polynomials)**.** Let $\mathbb{F}$ be a field, and consider the polynomial ring $\mathbb{F}[x]$. Let $f, g \in \mathbb{F}[x]$ with $f \neq 0$. There exist unique $q, r \in \mathbb{F}[x]$ satisfying (D1) and (D2) where
(D1) $g = qf + r$;
(D2) $r = 0$ or $\deg r < \deg f$.
We call g the **dividend**, f the **divisor**, q the **quotient**, and r the **remainder**.

Proof. The proof is essentially the procedure of long division of polynomials.

If $g = 0$, let $r = q = 0$. Then $g = qf + r$ and $r = 0$.

Now assume $g \neq 0$. If $\deg g < \deg f$, let $r = g$ and $q = 0$. Then $g = qf + r$ and $\deg r < \deg f$.

Otherwise, $\deg g \geq \deg f$. Let $m = \deg f$ and $n = \deg g - \deg f$. We proceed by induction on n.

For the *inductive base* $n = 0$, we have $\deg g = \deg f = m$. Let $a_m, \ldots, a_1, b_m, \ldots, b_1 \in R$ such that

$$g = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0$$
$$f = b_m x_m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0.$$

Let $q = \frac{a_m}{b_m}$ and $r = g - qf$. Since $\mathbb{F}$ is a field and $b_m \neq 0$, we can safely conclude that q is a constant polynomial. Arithmetic shows that $g = qf + r$, but can we guarantee that $r = 0$ or $\deg r < \deg f$? Apply substitution, distribution, and polynomial addition to obtain

$$\begin{aligned}
r &= g - qf \\
&= \left(a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0\right) \\
&\quad - \frac{a_m}{b_m}\left(b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0\right) \\
&= \left(a_m - \frac{a_m}{b_m}\cdot b_m\right)x^m + \left(a_{m-1} - \frac{a_m}{b_m}\cdot b_{m-1}\right)x^{m-1} + \cdots + \left(a_0 - \frac{a_m}{b_m}\cdot b_0\right) \\
&= 0x^m + \left(a_{m-1} - \frac{a_m}{b_m}\cdot b_{m-1}\right)x^{m-1} + \cdots + \left(a_0 - \frac{a_m}{b_m}\cdot b_0\right).
\end{aligned}$$

Since the coefficient of x^m is zero, we see that if $r \neq 0$, then $\deg r < \deg f$.

For the *inductive hypothesis*, assume that for all $i < n$ there exist $q, r \in R[x]$ such that $g = qf + r$ and $r = 0$ or $\deg r < \deg f$.

For the *inductive step*, let $\ell = \deg g$. Let $a_m, \ldots, a_0$, $b_\ell, \ldots, b_0 \in R$ such that

$$\begin{aligned}
f &= a_m x^m + \cdots + a_0 \\
g &= b_\ell x^\ell + \cdots + b_0.
\end{aligned}$$

Let $p = \frac{b_\ell}{a_m} \cdot x^n$ and $r = g - pf$. Once again, since $\mathbb{F}$ is a field and $a_m \neq 0$, we can safely conclude that $p \in \mathbb{F}[x]$. Apply substitution and distribution to obtain

$$\begin{aligned}
g' &= g - pf \\
&= g - \frac{b_\ell}{a_m}\cdot x^n\left(a_m x^m + \cdots + a_0\right) \\
&= g - \left(b_\ell x^{m+n} + \frac{b_\ell a_{m-1}}{a_m}\cdot x^{m-1+n} + \cdots + \frac{b_\ell a_0}{a_m}\cdot x^n\right).
\end{aligned}$$

Recall that $n = \deg g - \deg f = \ell - m$, so $\ell = m + n$. Apply substitution and polynomial addition to obtain

$$\begin{aligned}
g' = g - pf &= \left(b_\ell x^\ell + \cdots + b_0\right) \\
&\quad - \left(b_\ell x^\ell + \frac{b_\ell a_{m-1}}{a_m}\cdot x^{\ell-1} + \cdots + \frac{b_\ell a_0}{a_m}\cdot x^n\right) \\
&= 0x^\ell + \left(b_{\ell-1} - \frac{b_\ell a_{m-1}}{a_m}\right)x^{\ell-1} \\
&\quad + \cdots + \left(b_n - \frac{b_\ell a_0}{a_m}\right)x^n + b_{n-1}x^{n-1} \cdots + b_0.
\end{aligned}$$

Since $\mathbb{F}$ is a field and $a_m \neq 0$, we can safely conclude that $g' \in \mathbb{F}[x]$. Observe that $\deg g' < \ell = \deg g$, so $\deg g' - \deg f < n$. Apply the inductive hypothesis to find $p', r \in R[x]$ such that

$g' = p'f + r$ and $r = 0$ or $\deg r < \deg f$. Then

$$\begin{aligned} g = pf + g' &= pf + (p'f + r) \\ &= (p + p')f + r. \end{aligned}$$

Let $q = p + p'$. By closure, $q \in R[x]$, and we have shown the existence of a quotient and remainder.

For uniqueness, assume that there exist $q_1, q_2, r_1, r_2 \in R[x]$ such that $g = q_1 f + r_1 = q_2 f + r_2$ and $\deg r_1, \deg r_2 < \deg f$. Then

$$\begin{aligned} q_1 f + r_1 &= q_2 f + r_2 \\ 0 &= (q_2 - q_1) f + (r_2 - r_1). \end{aligned} \tag{27}$$

If $q_2 - q_1 \neq 0$, then no term of $(q_2 - q_1)\operatorname{lm}(f)$ has degree smaller than $\deg f$. Since every term of $r_2 - r_1$ has degree smaller than $\deg f$, there are no like terms between the two. Thus, there can be no cancellation between $(q_2 - q_1)\operatorname{lm}(f)$ and $r_2 - r_1$, and for similar reasons there can be no cancellation between $(q_2 - q_1)\operatorname{lm}(f)$ and lower-degree terms of $(q_2 - q_1) f$. However, the left hand side of equation 27 is the zero polynomials, so coefficients of $(q_2 - q_1)\operatorname{lm}(f)$ are all 0 on the left hand side. They must likewise be all zero on the right hand side. That implies $(q_2 - q_1)\operatorname{lm}(f)$ is equal to the constant polynomial 0. We are working in an integral domain (Exercise 7.53), and $\operatorname{lm}(f) \neq 0$, so it mus tbe that $q_2 - q_1 = 0$. In other words, $q_1 = q_2$.

Once we have $q_2 - q_1 = 0$, substitution into (27) implies that $0 = r_2 - r_1$. Immediately we have $r_1 = r_2$. We have shown that q and r are unique. □

Notice that the theorem does *not* apply if $R = \mathbb{Z}$, and Exercise 7.50 explains why. That's a shame.

Euclidean domains

Recall from Section 6.1 that the Euclidean algorithm for integers is basically repeated division. You can infer, more or less correctly, that a similar algorithm works for polynomials.

Why stop there? We have a notion of divisibility in rings, and we just found that the Division Theorem for integers can be generalized to any polynomial ring whose ground ring is a field. Can we generalize the Division Theorem beyond a ring of polynomials over a field? We can, but it requires us to generalize the notion of a remainder, as well.

Definition 7.59. Let R be an integral domain and v a function mapping the nonzero elements of R to $\mathbb{N}^+$. We say that R is a **Euclidean Domain** with respect to the **valuation function** v if it satisfies (E1) and (E2) where

(E1) $v(r) \leq v(rs)$ for all nonzero $r, s \in R$.

(E2) For all nonzero $f \in R$ and for all $g \in R$, there exist $q, r \in R$ such that
- $g = qf + r$, and
- $r = 0$ or $v(r) < v(f)$.

Example 7.60. By the Division Theorem, $\mathbb{Z}$ is a Euclidean domain with the valuation function $v(r) = |r|$.

Theorem 7.61. Let $\mathbb{F}$ be a field. Then $\mathbb{F}[x]$ is a Euclidean domain with the valuation function $v(r) = \deg r$.

Proof. You do it! See Exercise 7.71. □

Example 7.62. On the other hand, $\mathbb{Z}[x]$ is *not* a Euclidean domain if the valuation function is $v(r) = \deg r$. 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 $\deg x > \deg 2$.

If you think back to the Euclidean algorithm, you might remember that it requires only *the ability to perform a division with a unique remainder that was smaller than the divisor*. This means that we can perform the Euclidean algorithm in a Euclidean ring! — But will the result have the same properties as when we perform it in the ring of integers?

Yes and no. We *do* get an object whose properties resemble those of the greatest common divisor of two integers. However, the result *might not be unique!* On the other hand, if we relax our expectation of uniqueness, we can get a greatest common divisor that is… *sort of* unique.

Definition 7.63. Let R be a ring. If $a, b, r \in R$ satisfy $ar = b$ or $ra = b$, then a **divides** b, a is a **divisor of** b, and b is **divisible** by a.

Now suppose that R is a Euclidean domain with respect to v, and let $a, b \in R$. If there exists $d \in R$ such that $d \mid a$ and $d \mid b$, then we call d a **common divisor** of a and b. If in addition all other common divisors d' of a and b divide d, then d is a **greatest common divisor** of a and b.

Two subtle differences with the definition for the integers have profound consequences.

- The definition refers to "a" greatest common divisor, not "the" greatest common divisor. *There can be many great"est" common divisors!*
- Euclidean domains measure "greatness" using divisibility (or multiplication) rather than order (or subtraction). As a consequence, the Euclidean domain R need not have a well ordering, or even a linear ordering — it needs only a valuation function! This is *why* there can be many great"est" common divisors.

Example 7.64. Consider $x^2 - 1, x^2 + 2x + 1 \in \mathbb{Q}[x]$. By Theorem 7.61, $\mathbb{Q}[x]$ is a Euclidean domain with respect to the valuation function $v(p) = \deg p$. Both of the given polynomials factor:

$$x^2 - 1 = (x+1)(x-1) \quad \text{and} \quad x^2 + 2x + 1 = (x+1)^2,$$

so we see that $x + 1$ is a divisor of both. In fact, it is a greatest common divisor, since no polynomial of degree two divides both $x^2 - 1$ and $x^2 + 2x + 1$.

However, $x + 1$ is not the *only* greatest common divisor. Another greatest common divisor is $2x + 2$. It may not be obvious that $2x + 2$ divides both $x^2 - 1$ and $x^2 + 2x + 1$, but it does:

$$x^2 - 1 = (2x+2)\left(\frac{x}{2} - \frac{1}{2}\right)$$

and

$$x^2 + 2x + 1 = (2x+2)\left(\frac{x}{2} + \frac{1}{2}\right).$$

Notice that $2x+2$ divides $x+1$ and vice-versa; also that $\deg(2x+2)=\deg(x+1)$.

Likewise, $\frac{x+1}{3}$ is also a greatest common divisor of x^2-1 and x^2+2x+1.

This new definition will allow more than one greatest common divisor even in $\mathbb{Z}$! For example, for $a=8$ and $b=12$, both 4 *and* -4 are greatest common divisors! This happens because each divides the other, emphasizing that in a generic Euclidean domain, the notion of a "greatest" common divisor is relative to divisibility, not to other orderings. However, when speaking of greatest common divisors in the integers, we typically use the ordering, not divisibility.

That said, all greatest common divisors have something in common.

Proposition 7.65. Let R be a Euclidean domain with respect to v, and $a,b\in R$. Suppose that d is a greatest common divisor of a and b. If d' is a common divisor of a and b, then $v(d')\leq v(d)$. If d' is another greatest common divisor of a and b, then $v(d)=v(d')$.

Proof. Since d is a greatest common divisor of a and b, and d' is a common divisor, the definition of a greatest common divisor tells us that d divides d'. Thus there exists $q\in R$ such that $qd'=d$. From property (E1) of a Euclidean domain,

$$v(d')\leq v(qd')=v(d).$$

On the other hand, if d' is also a greatest common divisor of a and b, an argument similar to the one above shows that

$$v(d)\leq v(d')\leq v(d).$$

Hence $v(d)=v(d')$. □

Finally we come to the point of a Euclidean domain: we can use the Euclidean algorithm to compute a gcd of any two elements! Essentially we transcribe the Euclidean Algorithm for integers (Theorem 6.4 on page 194 of Section 6.1).

Theorem 7.66 (The Euclidean Algorithm for Euclidean domains)**.** Let R be a Euclidean domain with valuation v and $m,n\in R\backslash\{0\}$. One can compute a 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 a greatest common divisor of m and n.

Proof. You do it! See Exercise 7.72. □

Just as we could adapt the Euclidean Algorithm for integers to the Extended Euclidean Algorithm in order to compute $a,b\in\mathbb{Z}$ such that Bezout's Identity holds,

$$am+bn=\gcd(m,n),$$

we can do the same in Euclidean domains. You will need this for Exercise 7.72.

Exercises.

Exercise 7.67. Try to devise a division algorithm for $\mathbb{Z}_n$? Does the value of n matter?

Exercise 7.68. Let $f = 2x^2 + 1$ and $g = x^3 - 1$.
(a) Show that 1 is a greatest common divisor of f and g in $\mathbb{Q}[x]$, and find $a, b \in \mathbb{Q}[x]$ such that $1 = af + bg$.
(b) Recall that $\mathbb{Z}_5$ is a field. Show that 1 is a greatest common divisor of f and g in $\mathbb{Z}_5[x]$, and find $a, b \in \mathbb{Z}_5[x]$ such that $1 = af + bg$.
(c) Recall that $\mathbb{Z}[x]$ is not a Euclidean domain. Explain why the result of part (a) cannot be used to show that 1 is a greatest common divisor of f and g in $\mathbb{Z}[x]$. What would you get if you used the Euclidean algorithm on f and g in $\mathbb{Z}[x]$?

Exercise 7.69. Let $f = x^4 + 9x^3 + 27x^2 + 31x + 12$ and $g = x^4 + 13x^3 + 62x^2 + 128x + 96$.
(a) Compute a greatest common divisor of f and g in $\mathbb{Q}[x]$.
(b) Recall that $\mathbb{Z}_{31}$ is a field. Compute a greatest common divisor of f and g in $\mathbb{Z}_{31}[x]$.
(c) Recall that $\mathbb{Z}_3$ is a field. Compute a greatest common divisor of f and g in $\mathbb{Z}_3[x]$.
(d) Even though $\mathbb{Z}[x]$ is not a Euclidean domain, it still has greatest common divisors. What's more, we can compute the greatest common divisors using the Euclidean algorithm! How?
(e) You can even compute the greatest common divisors *without* using the Euclidean algorithm, but by examining the answers to parts (b) and (c) slowly. How?

Exercise 7.70. Show that every field is a Euclidean domain.

Exercise 7.71. Prove Theorem 7.61.

Exercise 7.72. Prove Theorem 7.66, the Euclidean Algorithm for Euclidean domains.

Exercise 7.73. A famous Euclidean domain is the ring of *Gaussian integers*

$$\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$$

where $i^2 = -1$. Let $v : \mathbb{Z}[i] \to \mathbb{Z}$ by

$$v(a + bi) = a^2 + b^2.$$

(a) Show that $a + bi$ is "orthogonal" to $i(a + bi)$, in the sense that the slope of the line segment connecting 0 and $a + bi$ in the complex plane is orthogonal to the slope of the line segment connecting 0 and $i(a + bi)$.
(b) Assuming the facts given about v, divide:
  (i) 11 by 3;
  (ii) 11 by $3i$;
  (iii) $2 + 3i$ by $1 + 2i$.
(c) Show that v is, in fact, a valuation function suitable for a Euclidean domain.
(d) Describe a method for dividing Gaussian integers. (Again, it helps to think of them as vectors in the plane. See Exercise **??** on page ??.)

Chapter 8:
Ideals

This chapter fills two roles. Some sections describe ring analogs to structures that we introduced in group theory:

- Section 8.1 introduces the *ideal*, an analog to a normal subgroup;
- Section 8.3 provides an analog of quotient groups; and
- Section 8.5 decries ring homomorphisms.

The remaining sections use these ring structures to introduce new kinds of ring structures:

- Section 8.2 describes an important class of rings; and
- Section 8.4 highlights an important class of ideals.

8.1: Ideals

Given that normal subgroups were so important to group theory, it will not surprise you that a special kind of subring plays an crucial role in ring theory. But, what sort of properties should it have? Rather than take the structural approach that we took last time, and find a criterion on a subring that guarantees we can create a "quotient" that gives us a new ring, let's look at the mathematical applications of rings that interest us.

An application which may strike the reader as more concrete is the question of the roots of polynomials. Start with a ring R, an element $a \in R$, and three univariate polynomials f, g, and p over R. How do the roots of f and/or g behave with respect to ring operations? If a is a root of both f and g, then a is also a root of their sum $h = f + g$, since

$$h(a) = (f + g)(a) = f(a) + g(a) = 0.$$

Also, if a is a root *only of f*, then it is a root of *any* multiple of f, such as $h = fp$. After all,

$$h(a) = (fp)(a) = f(a)\,p(a) = 0 \cdot p(a) = 0.$$

Something subtle is going on here, and you may have missed it, so let's look more carefully. Let S be the subring of R that contains *all* polynomials that have a as a root. By definition, f, g, and h are all in S, but p is not! Compare this to group theory: the product of an element of a subgroup and a element outside the subgroup is *never* in the subgroup. What we are seeing is a property we had studied way back when we looked at monoids: S is an *absorbing subset* of R.

Notice how absorption creates an important difference from group theory. With groups, multiplying an element of a subgroup with an element outside the subgroup *always* gave us another element outside the subgroup! This allowed us to create cosets, and partition the group. We obviously cannot rely on this property to the same thing in rings, because some subrings absorb multiplication from outside the subgroup! You might argue that it still holds for addition, and that is true – in fact, we will use that fact later to create cosets that partition a ring.

Recall our definition of S as the subring of R that contains all polynomials that have a as a root. This definition is quite simple, and clearly important. The fact that S "absorbs" any polynomial that does *not* have a as a root indicates that the absorption property is important. This property likewise occurs with other subrings that have straightforward and obvious definitions;

for example, the subring A of $\mathbb{Z}$ that contains all multiples of 4 and 6: $3 \notin A$, but $4 \cdot 3 \in A$ and $6 \cdot 3 \in A$. When a property appears in many contexts that are very different but important, it merits investigation.

Definition and examples

Definition 8.1. Let A be a subring of R that satisfies the **absorption property**:

$$\forall r \in R \quad \forall a \in A \qquad ra \in A.$$

Then A is an **ideal subring** of R, or simply, an **ideal**, and we write $A \triangleleft R$. An ideal A is **proper** if $\{0\} \neq A \neq R$.

Recall that our rings are assumed to be commutative, so if $ra \in A$ then $ar \in A$, also.

Example 8.2. Recall the subring $2\mathbb{Z}$ of the ring $\mathbb{Z}$. We claim that $2\mathbb{Z} \triangleleft \mathbb{Z}$. Why? Let $r \in \mathbb{Z}$, and $a \in 2\mathbb{Z}$. By definition of $2\mathbb{Z}$, there exists $d \in \mathbb{Z}$ such that $a = 2d$. Substitution gives us

$$ra = r \cdot 2d = 2(rd) \in 2\mathbb{Z},$$

so $2\mathbb{Z}$ "absorbs" multiplication by $\mathbb{Z}$. This makes $2\mathbb{Z}$ an ideal of $\mathbb{Z}$.

Naturally, we can generalize this proof to arbitrary $n \in \mathbb{Z}$: see Exercises 8.14 and 8.16.

Ideals in the ring of integers have a nice property that we will use in future examples.

Example 8.3. Certainly $3 \mid 6$ since $3 \cdot 2 = 6$. Look at the ideals generated by 3 and 6:

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

Inspection suggests that $6\mathbb{Z} \subseteq 3\mathbb{Z}$. Is it? Let $x \in 6\mathbb{Z}$. By definition, $x = 6q$ for some $q \in \mathbb{Z}$. By substitution, $x = (3 \cdot 2)q = 3(2 \cdot q) \in 3\mathbb{Z}$. Since x was arbitrary in $6\mathbb{Z}$, we have $6\mathbb{Z} \subseteq 3\mathbb{Z}$.

Lemma 8.4. Let $a, b \in \mathbb{Z}$. The following are equivalent:
(A) $a \mid b$;
(B) $b\mathbb{Z} \subseteq a\mathbb{Z}$.

Proof. You do it! See Exercise 8.17. □

Earlier in the section, we looked at roots of univariate polynomials. The same properties hold when we move to 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 8.5. You showed in Exercise 7.3 that $\mathbb{C}[x, y]$ is a ring. Let $f = x^2 + y^2 - 4$, $g = xy - 1$, and $A = \{hf + kg : h, k \in \mathbb{C}[x, y]\}$. From a geometric perspective what's interesting about A is that *the common roots of f and g are roots of any element of A.* To see this, let (α, β) be a

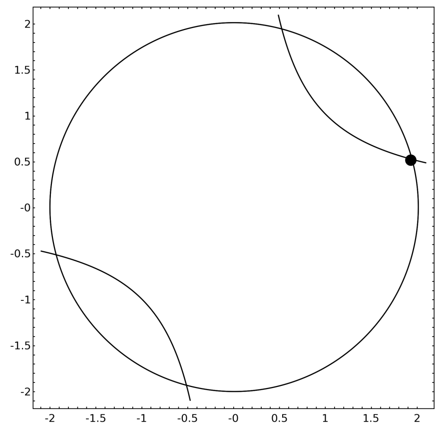


Figure 8.1. A common root of x^2+y^2-4 and $xy-1$

common root of f and g; that is, $f(\alpha,\beta)=g(\alpha,\beta)=0$. Let $p\in A$; by definition, we can write $p=hf+kg$ for some $h,k\in\mathbb{C}[x,y]$. By substitution,

$$\begin{aligned}p(\alpha,\beta)&=(hf+kg)(\alpha,\beta)\\&=h(\alpha,\beta)\cdot f(\alpha,\beta)+k(\alpha,\beta)\cdot g(\alpha,\beta)\\&=h(\alpha,\beta)\cdot 0+k(\alpha,\beta)\cdot 0\\&=0;\end{aligned}$$

that is, (α,β) is a root of p. Figure 8.1 depicts the root

$$(\alpha,\beta)=\left(\sqrt{2+\sqrt{3}},2\sqrt{2+\sqrt{3}}-\sqrt{6+3\sqrt{3}}\right).$$

The remarkable thing is that A is an ideal. To show this, we must show that A is a subring of $\mathbb{C}[x,y]$ that absorbs multiplication.

- Is A a subring? Let $a,b\in A$. By definition, we can find $h_a,h_b,k_a,k_b\in\mathbb{C}[x,y]$ such that $a=h_af+k_ag$ and $b=h_bf+k_bg$. A little arithmetic gives us

$$\begin{aligned}a-b&=(h_af+k_ag)-(h_bf+k_bg)\\&=(h_a-h_b)f+(k_a-k_b)g\in A.\end{aligned}$$

  To show that $ab\in A$, we will distribute over *one* of the two polynomials:

$$\begin{aligned}ab&=a(h_bf+k_bg)\\&=a(h_bf)+a(k_bg)\\&=(ah_b)f+(ak_b)g.\end{aligned}$$

  Let

$$h'=ah_b\quad\text{and}\quad k'=ak_b;$$

then $ab = h'f + k'g$, and by closure, $h', k' \in \mathbb{C}[x, y]$. By definition, $ab \in A$, as well. By the Subring Theorem, A is a subring of $\mathbb{C}[x, y]$.

- Does A absorb multiplication? Let $a \in A$, and $r \in \mathbb{C}[x, y]$. By definition, we can write $a = h_a f + k_a g$, as above. A little arithmetic gives us

$$\begin{aligned} ra &= r(h_a f + k_a g) = r(h_a f) + r(k_a g) \\ &= (rh_a)f + (rk_a)g \in A. \end{aligned}$$

Let

$$h' = rh_a \quad \text{and} \quad k' = rk_a;$$

then $ra = h'f + k'g$, and by closure, $h', k' \in \mathbb{C}[x, y]$. By definition, $ra \in A$, as well. By definition, A satisfies the absorption property.

We have shown that A satisfies the subring and absorption properties; thus, $A \triangleleft \mathbb{C}[x, y]$.

You will show in Exercise 8.24 that the ideal of Example 8.5 can be generalized to other rings and larger numbers of variables.

Remark 8.6. Recall from linear algebra that *vector spaces* are an important tool for the study of systems of linear equations. If we find a *triangular basis* of a system of linear polynomials, we can analyze the subspace of solutions of the system.

Example 8.5 illustrates that ideals are an important analog for non-linear polynomial equations. If we can find a "triangular basis" of an ideal, then we can analyze the solutions of the system in a method very similar to methods for linear systems. We take up this task in Chapter 11.

Properties and elementary theory

Since ideals are fundamental, we would like an analog of the Subring Theorem to decide whether a subset of a ring is an ideal. You might have noticed from the example above that absorption actually implies closure under multiplication. After all, if $rb \in A$ for every $r \in R$, then since $a \in A$ implies $a \in R$, we really have $ab \in A$, too. The Ideal Theorem uses this fact to simplify the criteria for an ideal.

Theorem 8.7 (The Ideal Theorem)**.** Let R be a ring and $A \subseteq R$ with A nonempty. The following are equivalent:
(A) A is an ideal subring 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 $ar, ra \in A$.

Proof. You do it! See Exercise 8.19. □

We conclude by defining a special kind of ideal, with a notation similar to that of cyclic subgroups, but with a different meaning.

Notation 8.8. Let R be a ring with unity, $m \in \mathbb{N}^+$, and $r_1, r_2, \ldots, r_m \in R$. Define the set $\langle r_1, r_2, \ldots, r_m \rangle$ as the intersection of all the ideals of R that contain all of $r_1, r_2, \ldots, r_m$.

Proposition 8.9. For all $r_1, \dots, r_m \in R$, $\langle r_1, \dots, r_m \rangle$ is an ideal.

We will not prove this proposition, as it is a direct consequence of the next:

Proposition 8.10. For every set $\mathcal{I}$ of ideals of a ring R, $\bigcap_{I \in \mathcal{I}} I$ is also an ideal.

Proof. Denote $J = \bigcap_{I \in \mathcal{I}} I$. Observe that $J \neq \emptyset$ because 0_R is an element of every ideal. Let $a, b \in J$ and $r \in R$. Let $I \in \mathcal{I}$. Since J contains only those elements that appear in every element of $\mathcal{I}$, 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 $\mathcal{I}$, *every* element of $\mathcal{I}$ 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 $\langle r_1, \dots, r_m \rangle$ is defined as the intersection of ideals containing $r_1, \dots, r_m$, Proposition 8.10 implies that $\langle r_1, \dots, r_m \rangle$ is an ideal. It is important enough to identify by a special name.

Definition 8.11. We call $\langle r_1, r_2, \dots, r_m \rangle$ the **ideal generated by** $r_1, r_2, \dots, r_m$, and $\{r_1, r_2, \dots, r_m\}$ a **basis** of $\langle r_1, r_2, \dots, r_m \rangle$.

This ideal is closely related to the ideal we used in Example 8.5.

Proposition 8.12. If R has unity, then $\langle r_1, r_2, \dots, r_m \rangle$ is precisely the set

$$I = \{h_1 r_1 + h_2 r_2 + \cdots + h_m r_m : h_i \in R\}.$$

Proof. First, we show that $I \subseteq \langle r_1, \dots, r_m \rangle$. Let $p \in I$; by definition, there exist $h_1, \dots, h_m \in R$ such that $p = \sum_{i=1}^m h_i r_i$. Let J be any ideal that contains all of $r_1, \dots, r_m$. By absorption, $h_i r_i \in J$ for each i. By closure, $p = \sum_{i=1}^m h_i r_i \in J$. Since J was an arbitrary ideal containing all of $r_1, \dots, r_m$, we infer that all the ideals containing all of $r_1, \dots, r_m$ contain p. Since p is an arbitrary element of I, I is a subset of all the ideals containing all of $r_1, \dots, r_m$. By definition, $I \subseteq \langle r_1, \dots, r_m \rangle$.

Now we show that $I \supseteq \langle r_1, \dots, r_m \rangle$. We claim that I is an ideal that contains each of $r_1, \dots, r_m$. If true, the definition of $\langle r_1, \dots, r_m \rangle$ does the rest, as it consists of elements common to *every* ideal that contains all of $r_1, \dots, r_m$.

But why is I an ideal? We first consider the absorption property. Let $f \in I$. By definition, there exist $h_1, \dots, h_m \in R$ such that

$$f = h_1 r_1 + \cdots + h_m r_m.$$

Let $p \in R$; we have

$$pf = (ph_1) r_1 + \cdots + (ph_m) r_m.$$

By closure, $ph_i \in R$ for each $i = 1, \dots, m$. We have written pf in a form that satisfies the definition of I, so $pf \in I$. As for the closure of subtraction, let $f, g \in I$; then choose $p_i, q_i \in R$ such that

$$f = p_1 r_1 + \cdots + p_m r_m \text{ and}$$
$$g = q_1 r_1 + \cdots + q_m r_m.$$

Using the associative property, the commutative property of addition, the commutative property of multiplication, distribution, and the closure of subtraction in R, we see that

$$\begin{aligned} f-g &= (p_1r_1+\cdots+p_mr_m)-(q_1r_1+\cdots+q_mr_m) \\ &= (p_1r_1-q_1r_1)+\cdots+(p_mr_m-q_mr_m) \\ &= (p_1-q_1)\,r_1+\cdots+(p_m-q_m)\,r_m. \end{aligned}$$

By closure, $p_i-q_i\in R$ for each $i=1,\ldots,m$. We have written $f-g$ in a form that satisfies the definition of I, so $f-g\in I$. By the Ideal Theorem, I is an ideal.

But, is $r_i\in I$ for each $i=1,2,\ldots,m$? Well,

$$r_i = 1_R\cdot r_i + \sum_{j\neq i} 0\cdot r_j \in I.$$

Since R has unity, this expression of r_i satisfies the definition of I, so $r_i\in I$.

Hence I is an ideal containing all of $r_1,r_2,\ldots,r_m$. By definition of $\langle r_1,\ldots,r_m\rangle$, $I\supseteq\langle r_1,\ldots,r_m\rangle$.

We have shown that $I\subseteq\langle r_1,\ldots,r_m\rangle\subseteq I$. Hence $I=\langle r_1,\ldots,r_m\rangle$ as claimed. □

As with vector spaces, the basis of an ideal *is not unique*.

Example 8.13. Consider the ring $\mathbb{Z}$, and let $I=\langle 4,6\rangle$. Proposition 8.12 claims that

$$I=\{4m+6n:\ m,n\in\mathbb{Z}\}.$$

Choosing concrete values of m and n, we see that

$$\begin{aligned} 4 &= 4\cdot 1+6\cdot 0\in I \\ 0 &= 4\cdot 0+6\cdot 0\in I \\ -12 &= 4\cdot(-3)+6\cdot 0\in I \\ -12 &= 4\cdot 0+6\cdot(-2)\in I. \end{aligned}$$

Notice that for some elements of I, we can provide representations in terms of 4 and 6 in more than one way.

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 $2=4\cdot(-1)+6\cdot 1$, so $2\in I$. (Even if it wasn't that easy, though, Bezout's Identity would do the trick: $\gcd(4,6)=4m+6n$ for some $m,n\in\mathbb{Z}$.) 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,

$$x=2q=[4\cdot(-1)+6\cdot 1]\cdot q=4\cdot(-q)+6\cdot q\in I.$$

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

$$x=4m+6n=2\,(2m+3n)\in 2\mathbb{Z}.$$

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 $I=\langle 4,6\rangle=\langle 2\rangle=2\mathbb{Z}$. 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 Section 8.2, the answer is, "Not always." However, the statement is true for the ring $\mathbb{Z}$ (and a number of other rings as well). You will explore this in Exercise 8.18, and Section 8.2.

Exercises.

Exercise 8.14. Show that for any $n\in\mathbb{N}$, $n\mathbb{Z}$ is an ideal of $\mathbb{Z}$.

Exercise 8.15. Suppose A is an ideal of R and B is an ideal of S. Is $A\times B$ an ideal of $R\times S$?

Exercise 8.16. Show that every ideal of $\mathbb{Z}$ has the form $n\mathbb{Z}$, for some $n\in\mathbb{N}$.

Exercise 8.17.
(a) Prove Lemma 8.4.
(b) More generally, prove that in *any* ring, $a\mid b$ if and only if $\langle b\rangle\subseteq\langle a\rangle$.

Exercise 8.18. In this exercise, we explore how $\langle r_1,r_2,\ldots,r_m\rangle$ behaves in $\mathbb{Z}$. Keep in mind that the results do not necessarily generalize to other rings.
(a) For the following values of $a,b\in\mathbb{Z}$, verify that $\langle a,b\rangle=\langle c\rangle$ for a certain $c\in\mathbb{Z}$.
  (i) $a=3$, $b=5$
  (ii) $a=3$, $b=6$
  (iii) $a=4$, $b=6$
(b) What is the relationship between a, b, and c in part (a)?
(c) Prove the conjecture you formed in part (b).

Exercise 8.19. Prove Theorem 8.7 (the Ideal Theorem).

Exercise 8.20. (a) Suppose R is a ring with unity, and A an ideal of R. Show that if $1_R\in A$, then $A=R$.
(b) Let q be an element of a ring with unity. Show that q has a multiplicative inverse if and only if $\langle q\rangle=\langle 1\rangle$.

Exercise 8.21. Show that in any field $\mathbb{F}$, the only two distinct ideals are the zero ideal and $\mathbb{F}$ itself.

Exercise 8.22. Let R be a ring and A and I two ideals of R. Decide whether the following subsets of R are also ideals, and explain your reasoning:
(a) $A\cap I$
(b) $A\cup I$
(c) $A+I=\{x+y:x\in A,y\in I\}$
(d) $A\cdot I=\{xy:x\in A,y\in I\}$
(e) $AI=\left\{\sum_{i=1}^{n}x_iy_i:n\in\mathbb{N},x_i\in A,y_i\in I\right\}$

Exercise 8.23. Let A,B be two ideals of a ring R. The definition of AB appears in Exercise 8.22.
(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 $AB\neq A\cap B$.

Exercise 8.24. Let R be a ring with unity. Recall the polynomial ring $P = R[x_1, x_2, \ldots, x_n]$, whose ground ring is R (Section 7.3). Let

$$\langle f_1, \ldots, f_m \rangle = \{h_1 f_1 + \cdots + h_m f_m : h_1, h_2, \ldots, h_m \in P\}.$$

Example 8.5 showed that the set $A = \langle x^2 + y^2 - 4, xy - 1 \rangle$ was an ideal; Proposition 8.12 generalizes this to show that $\langle f_1, \ldots, f_m \rangle$ is an ideal of P. Show that the common roots of $f_1, f_2, \ldots, f_m$ are common roots of all polynomials in the ideal I.

Exercise 8.25. Let A be an ideal of a ring R. Define its **radical** to be

$$\sqrt{A} = \{r \in R : r^n \in A \; \exists r \in \mathbb{N}^+\}.$$

(a) Suppose $R = \mathbb{Z}$. Compute $\sqrt{A}$ for
   (i) $A = 2\mathbb{Z}$
   (ii) $A = 9\mathbb{Z}$
   (iii) $A = 12\mathbb{Z}$
(b) Suppose $R = \mathbb{Q}[x]$. Compute $\sqrt{A}$ for
   (i) $A = \langle x + 1 \rangle$
   (ii) $A = \langle x^2 + 2x + 1 \rangle$
   (iii) $A = \langle x^2 + 1 \rangle$
(c) Show that $\sqrt{A}$ is an ideal.

8.2: Principal Ideal Domains

In the previous section, we described ideals for commutative rings with identity that are generated by a finite set of elements, denoting them by $\langle r_1, \ldots, r_m \rangle$. An important subclass of these ideals consists of ideals generated by only one element.

Principal ideal domains

Definition 8.26. Let A be an ideal of a ring R. If $A = \langle a \rangle$ for some $a \in R$, then A is a **principal ideal**.

Notice that, by Proposition 8.12, we have $\langle a \rangle = \{ra : r \in R\}$.

Many ideals can be rewritten as principal ideals. For example, the zero ideal $\{0\} = \langle 0 \rangle$. If R has unity, we can write $R = \langle 1 \rangle$. On the other hand, not all ideals are principal; we will show that if $A = \langle x, y \rangle$ in the ring $\mathbb{C}[x, y]$, there is no $f \in \mathbb{C}[x, y]$ such that $A = \langle f \rangle$.

The following property of principal ideals is extremely useful.

Lemma 8.27. Let R be a ring with unity, and $a, b \in R$. There exists $q \in R$ such that $qa = b$ if and only if $\langle b \rangle \subseteq \langle a \rangle$. In addition, if R is an integral domain and $a, b \neq 0$, then the same q has a multiplicative inverse if and only if $\langle b \rangle = \langle a \rangle$.

Proof. The first assertion is just Exercise 8.17(b).

For the second, assume first that R is an integral domain, $a, b \neq 0$, and $qa = b$. We first show that if q has a multiplicative inverse, then $\langle b \rangle = \langle a \rangle$. So, assume that q has a multiplicative inverse. The first assertion gives us $\langle b \rangle \subseteq \langle a \rangle$. By definition, q has a multiplicative inverse r iff $rq = 1_R$. By substitution, $rb = r(qa) = a$. By absorption, $a \in \langle b \rangle$. Hence $\langle b \rangle \supseteq \langle a \rangle$. We already had $\langle b \rangle \subseteq \langle a \rangle$, so we conclude that $\langle b \rangle = \langle a \rangle$.

We have shown that if q has a multiplicative inverse, then $\langle b \rangle = \langle a \rangle$. It remains to show the converse; namely, that if $\langle b \rangle = \langle a \rangle$, then q has a multiplicative inverse. So, assume that $\langle b \rangle = \langle a \rangle$. By definition, there exist $r, q \in R$ such that $a = rb$ and $b = qa$. By substitution, $a = r(qa) = (rq)a$, so $a(1 - rq) = 0$. Since R is an integral domain and $a \neq 0$, $1 - rq = 0$. Rewritten as $rq = 1$, it shows that q does have a multiplicative inverse, r. □

Outside an integral domain, a could divide b with an element that has no multiplicative inverse, yet $\langle b \rangle = \langle a \rangle$. For example, in $\mathbb{Z}_6$, we have $[2] \cdot [2] = [4]$, but $\langle [2] \rangle = \{[0], [2], [4]\} = \langle [4] \rangle$.

There are rings in which all ideals are principal.

Definition 8.28. A **principal ideal domain** is an integral domain where every ideal can be written as a principal ideal.

Example 8.29. We claim that $\mathbb{Z}$ is a principal ideal domain, and we can prove this using a careful application of Exercise 8.18. Let A be any ideal of $\mathbb{Z}$. The zero ideal is $\langle 0 \rangle$, so assume that $A \neq \{0\}$. In this case, A contains at least one non-zero element; call it a_1. Without loss of generality, we may assume that $a_1 \in \mathbb{N}^+$ (if not, we could take $-a_1$ instead, since the definition of an ideal requires $-a_1 \in A$ as well).

Is $A = \langle a_1 \rangle$? If not, we can choose $b_1 \in A \setminus \langle a_1 \rangle$. Let $q_1, r_1 \in \mathbb{Z}$ be the quotient and remainder from division of b_1 by a_1; notice that $r_1 = b_1 - q_1 a_1 \in A$. Let $a_2 = \gcd(a_1, r_1)$. By the Extended Euclidean Algorithm, we can find $x, y \in \mathbb{Z}$ such that $xa_1 + yr_1 = a_2$. Since $a_1, r_1 \in A$, absorption and closure imply that $a_2 \in A$. In addition, $b_1 \notin \langle a_1 \rangle$, so Lemma 8.27 implies that $a_1 \nmid b_1$, so $r_1 \neq 0$, so $a_2 = \gcd(a_1, r_1) \neq 0$. We have $0 < a_2 \leq r_1 < a_1$. In fact, since $a_2 \mid a_1$, Exercise 8.18 tells us that $\langle a_1, a_2 \rangle = \langle a_2 \rangle$.

Is $A = \langle a_2 \rangle$? If not, we can repeat the previous process to find $b_2 \in A \setminus \langle a_2 \rangle$, divide b_2 by a_2 to obtain a nonzero remainder $r_2 \in A$, and compute $a_3 = \gcd(a_2, r_2)$. Reasoning similar to that above implies that $0 < a_3 < a_2 < a_1$ and $\langle a_1, a_2, a_3 \rangle = \langle a_3 \rangle$.

Continuing in this fashion, we see that as long as $A \neq \langle a_i \rangle$, we can find nonzero $b_i \in A \setminus \langle a_i \rangle$, a nonzero remainder $r_i \in A$ from division of b_i by a_i, and nonzero $a_{i+1} = \gcd(a_i, r_i)$, so that $0 < a_{i+1} < a_i$ and $\langle a_1, \ldots, a_{i-1} \rangle = \langle a_i \rangle$. This gives us a strictly decreasing chain of integers $a_1 > a_2 > \cdots > a_i > 0$. By Exercise 0.37, this cannot continue indefinitely. Let d be the final a_i computed; since we cannot compute anymore, $A = \langle a_1, \ldots, d \rangle$. As the greatest common divisor of the previously computed a_i, however, we have $a_1, a_2, \ldots \in \langle d \rangle$. Thus, $A = \langle d \rangle$.

Before moving on, let's take a moment to look at how the ideals are related, as well. Let $B_1 = \langle a_1 \rangle$, and $B_2 = \langle a_1, a_2 \rangle$. Lemma 8.27 implies that $B_1 \subsetneq B_2$. Likewise, if we set $B_3 = \langle a_1, a_2, a_3 \rangle = \langle a_3 \rangle$, then $B_2 \subsetneq B_3$. In fact, as long as $A \neq \langle a_i \rangle$, we generate an ascending sequence of ideals $B_1 \subsetneq B_2 \subsetneq \cdots$. In other words, another way of looking at this proof is that it *expands the principal ideal B_i until $B_i = A$,* by adding elements not in B_i. Rather amazingly, the argument above implies that this ascending chain of ideals must stabilize, at least in $\mathbb{Z}$. This property that an ascending chain of ideals must stabilize is one that some rings satisfy, but not all; we return to it in a moment.

We can extend the argument of Example 8.29 to more general rings.

Theorem 8.30. 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 $v(r_i) < v(a_i)$;
- compute a gcd a_{i+1} of a_i and r_i;
  - notice $v(a_{i+1}) \leq v(r_i) < v(a_i)$;
- 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 $v(a_i) \leq v(a_{i+1}) < v(a_i)$
- hence, $\langle a_i \rangle \subsetneq \langle a_{i+1} \rangle$ and $v(a_{i+1}) < v(a_i)$.

By Exercise 0.37, 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 $A = \langle a_i \rangle$. □

Not all integral domains are principal ideal domains; you will show in the exercises that for any field $\mathbb{F}$ and its polynomial ring $\mathbb{F}[x,y]$, the ideal $\langle x,y \rangle$ is not principal.

Noetherian rings and the Ascending Chain Condition

For now, though, we will turn to a phenomenon that appeared in Example 8.29 and Theorem 8.30. In each case, we built a chain of ideals

$$\langle a_1 \rangle \subsetneq \langle a_2 \rangle \subsetneq \langle a_3 \rangle \subsetneq \cdots$$

and were able to show that the procedure we used to find the a_i must eventually terminate.

This property is very useful for a ring. In both Example 8.29 and Theorem 8.30, we relied on the well-ordering of $\mathbb{N}$, but that is not always available to us. So the property might be useful in other settings, even in cases where ideals aren't guaranteed to be principal. For example, eventually we will show that $\mathbb{F}[x,y]$ satisfies this property.

Definition 8.31. Let R be a ring. If for every **ascending chain of ideals** $A_1 \subseteq A_2 \subseteq \cdots$ we can find an integer k such that $A_k = A_{k+1} = \cdots$, then R satisfies the **Ascending Chain Condition**.

Remark 8.32. Another name for a ring that satisfies the Ascending Chain Condition is a **Noetherian ring**, after the German mathematician Emmy Noether.

Theorem 8.33. Each of the following holds.

(A) Every principal ideal domain satisfies the Ascending Chain Condition.

(B) Any field $\mathbb{F}$ satisfies the Ascending Chain Condition.

(C) If a ring R satisfies the Ascending Chain Condition, so does $R[x]$.

(D) If a ring R satisfies the Ascending Chain Condition, so does $R[x_1, x_2, \ldots, x_n]$.

Proof. (A) Let R be a principal ideal domain, and let $A_1 \subseteq A_2 \subseteq \cdots$ be an ascending chain of ideals in R. Let $B = \bigcup_{i=1}^{\infty} A_i$. By Exercise 8.39, B is an ideal. Since R is a principal ideal domain, $B = \langle b \rangle$ for some $b \in R$. By definition of a union, $b \in A_i$ for some $i \in \mathbb{N}$. The definition of an ideal now implies that $rb \in A_i$ for all $r \in R$; since $\langle b \rangle = \{rb : r \in R\}$, we infer that $\langle b \rangle \subseteq A_i$. By substitution, $B \subseteq A_i$. By definition of union, we also have $A_i \subseteq B$. Hence $A_i = B$, and a similar argument shows that $A_j = B$ for all $j \geq i$. In other words, the chain of ideals stabilizes at A_i. Since the chain was arbitrary, every ascending chain of ideals in R stabilizes, so R satisfies the ascending chain condition.

(B) By Exercise 7.70, any field $\mathbb{F}$ is a Euclidean domain, so this follows from (A) and Theorem 8.30. However, it's instructive to look at it from the point of view of a field as well. Recall from Exercise 8.21 that a field has only two distinct ideals: the zero ideal, and the field itself. Hence, any ascending chain of ideals stabilizes either at the zero ideal or at $\mathbb{F}$ itself.

(C) Assume that R satisfies the Ascending Chain Condition. The argument is based on two claims.

Claim 1: Every ideal of $R[x]$ is finitely generated. Let A be any ideal of $R[x]$, and choose $f_1, f_2, \ldots \in A$ in the following way:

- Let $B_0 = \{0\}$, and $k = 0$.
- While $A \neq \langle B_k \rangle$:
  - Let $S_k = \{\deg f : f \in A \backslash \langle B_k \rangle\}$. Since $S_k \subseteq \mathbb{N}$, it has a least element; call it d_k.
  - Let $f_k \in A \backslash \langle B_k \rangle$ be any polynomial of degree d_k. Notice that $f_k \in A \backslash \langle B_k \rangle$ implies that $\langle B_k \rangle \subsetneq \langle B_k \cup \{f_k\} \rangle$.
  - Let $B_{k+1} = B_k \cup \{f_k\}$, and add 1 to k.

Does this process terminate? We built $\langle f_1 \rangle \subsetneq \langle f_1, f_2 \rangle \subsetneq \cdots$ as an ascending chain of ideals. Denote the leading coefficient of f_k by a_k and let $C_k = \langle a_1, a_2, \ldots, a_k \rangle$. Since R satisfies the Ascending Chain Condition, the ascending chain of ideals $C_1 \subseteq C_2 \subseteq \cdots$ stabilizes for some $m \in \mathbb{N}$.

We claim that the chain $\langle B_0 \rangle \subsetneq \langle B_1 \rangle \subsetneq \langle B_2 \rangle \subsetneq \cdots$ has also stabilized at m; that is, we cannot find $f_{m+1} \in A \backslash \langle B_m \rangle$. By way of contradiction, suppose we can find f_{m+1} of minimal degree in $A \backslash \langle B_{m+1} \rangle$. By hypothesis, the chain of ideals C_k has stabilized, so $C_m = C_{m+1}$. Thus, $a_{m+1} \in C_{m+1} = C_m$. That means we can write $a_{m+1} = b_1 a_1 + \cdots + b_m a_m$ for some $b_1, \ldots, b_m \in R$. Write $d_i = \deg_x f_i$, and let

$$p = b_1 f_1 x^{d_{m+1} - d_1} + \cdots + b_m f_m x^{d_{m+1} - d_m}.$$

We chose each f_i to be of minimal degree, so for each i, we have $d_i \leq d_{m+1}$. Thus, $d_{m+1} - d_i \in \mathbb{N}$, and $p \in R[x]$. Moreover, we have set up the sum and products so that $\operatorname{lt}\left(b_i f_i x^{d_{m+1} - d_i}\right) =$

$b_i\left(a_i x^{d_i}\right) x^{d_{m+1}-d_i} = b_i a_i x^{d_{m+1}}$. This implies that the leading term of p is

$$\left(b_1 a_1 + \cdots + b_m a_m\right) x^{d_{m+1}} = a_{m+1} x^{d_{m+1}}.$$

Let $r = f_{m+1} - p$. Since $\operatorname{lt}\left(f_{m+1}\right) = \operatorname{lt}(p)$, the leading terms cancel, and $\deg r < \deg f_{m+1}$. By construction, $p \in B_{m+1}$. If $r \in \left\langle B_{m+1}\right\rangle$, we could rewrite $r = f_{m+1} - p$ as $f_{m+1} = r + p$, which would imply that $f_{m+1} \in \left\langle B_{m+1}\right\rangle$. This contradicts the choice of $f_{m+1} \in A \backslash \left\langle B_{m+1}\right\rangle$. Thus, $r \notin \left\langle B_{m+1}\right\rangle$. Since f_{m+1} and p are both in A, we have $r \in A \backslash \left\langle B_{m+1}\right\rangle$. However, $\deg r < \deg f_{m+1}$; this contradicts the choice of f_{m+1} as a polynomial with minimal degree in $A \backslash \left\langle B_{m+1}\right\rangle$.

The only unfounded assumption was that we could find $f_{m+1} \in A \backslash \langle B_m \rangle$. Apparently, we cannot do so, and the process of choosing elements of $A \backslash \langle B_i \rangle$ must terminate at $i = m$. Since it does not terminate unless $A = \langle B_m \rangle$, we conclude that $A = \langle B_m \rangle = \langle f_1, \ldots, f_m \rangle$. In other words, A is finitely generated.

Claim 2: Every ascending chain of ideals in R eventually stabilizes. Let $I_1 \subseteq I_2 \subseteq \cdots$ be an ascending chain. By Exercise 8.39, the set $I = \cup_{i=1}^{\infty} I_i$ is also an ideal. By Claim 1, I is finitely generated; let $f_1, \ldots, f_m \in I$ such that $I = \langle f_1, \ldots, f_m \rangle$. By definition of union, there exist $k_1, \ldots, k_m \in \mathbb{N}^+$ such that $f_j \in I_{k_j}$. By definition of subset, $f_j \in I_\ell$ for all $\ell > k_j$. Let $\ell \geq \max\{k_1, \ldots, k_m\}$; then $f_1, \ldots, f_m \in I_\ell$, so

$$I = \langle f_1, \ldots, f_m \rangle \subseteq I_\ell \subseteq I,$$

which implies equality. Thus, the chain stabilizes at $\max\{k_1, \ldots, k_m\}$.

(D) follows from (C) by induction on the number of variables n: use R to show $R[x_1]$ satisfies the Ascending Chain Condition; use $R[x_1]$ to show that $R[x_1, x_2] = (R[x_1])[x_2]$ satisfies the Ascending Chain Condition; etc. □

Corollary 8.34 (Hilbert Basis Theorem)**.** For any Noetherian ring R, $R[x_1, x_2, \ldots, x_n]$ satisfies the Ascending Chain Condition. In particular, this is true when R is a field $\mathbb{F}$. Thus, for any ideal I of $\mathbb{F}[x_1, \ldots, x_n]$, we can find $f_1, \ldots, f_m \in I$ such that $I = \langle f_1, \ldots, f_m \rangle$.

Proof. Apply (B) and (D) of Theorem 8.33. □

Exercises

Exercise 8.35. Let $d \in \mathbb{Z}$. Explain why:

(a) $d\mathbb{Z}$ is not a principal ideal domain, but

(b) every ideal is still principal.

Exercise 8.36. Is $\mathbb{F}[x]$ a principal ideal domain for every field $\mathbb{F}$? What about $R[x]$ for every ring R?

Exercise 8.37. Is the ring $\mathbb{Z}[i]$ of Gaussian integers a principal ideal domain? Why or why not?

Exercise 8.38. Let $\mathbb{F}$ be any field, and consider the polynomial ring $\mathbb{F}[x, y]$. Explain why $\langle x, y \rangle$ cannot be principal.

Exercise 8.39. Let R be a ring and $I_1 \subseteq I_2 \subseteq \cdots$ an ascending chain of ideals. Show that $\mathcal{I} = \bigcup_{i=1}^{\infty} I_i$ is itself an ideal.

Exercise 8.40. Show that $\mathbb{Z}$ satisfies the Ascending Chain Condition.

Exercise 8.41. Let R be a ring and $a, b \in R$.

(a) Show that if R has unity, $\langle a \rangle \langle b \rangle = \langle ab \rangle$.
(b) Show that if R does not have unity, it can happen that $\langle a \rangle \langle b \rangle \neq \langle ab \rangle$.

Exercise 8.42. Suppose R and S are Noetherian rings. Is $R \times S$ also a Noetherian ring?

8.3: Cosets and Quotient Rings

Recall that in group theory, we could use cosets of a subgroup to create equivalence classes in a group. We want to do the same thing for rings. Since a ring has two operations, we need to decide which one we ought to use to do this. The decision isn't very hard; as we saw in Section 8.1, some subrings absorb multiplication — in particular, *ideals* absorb multiplication — so we cannot expect to create cosets using that operation. We will have to try with addition alone.

Definition 8.43. Let R be a ring and S a subring of R. For every $r \in R$, denote
$$r + S := \{r + s : s \in S\},$$
called a **coset**. Then define
$$R/S := \{r + S : r \in R\}.$$

Since a sub*ring* is always a sub*group* under addition — in fact, it is a *normal* subgroup — and subgroups partition a group, we can immediately identify three properties of the cosets of a subring:

- they partition the ring as an additive group;
- they create a set of equivalence classes of the additive group;
- they create a quotient *group* under *addition;* and
- coset equality in rings follows the rules of coset equality in groups, listed in Lemma 3.37 on page 121.

Do they also create a quotient *ring*? In fact, they might not!

The necessity of ideals

The absorption property plays a critical role in guaranteeing that multiplication is well-defined.

Example 8.44. Let $R = \mathbb{Z}[x]$, and S the smallest subring of R that contains $x^2 - 1$. It is not hard to see that $S = \left\{\sum_{i=1}^{n} a_i \left(x^2-1\right)^{p_1} : n, p_i \in \mathbb{N}^+, a_i \in \mathbb{Z}\right\}$.

Let $X = x + S$ and $Y = 1 + S$. Notice that we can write $Y = x^2 + S$ as well, because $x^2 - 1 \in S$. However, the value of XY is not equal for both representations of Y! The first gives us
$$XY = (x + S)(1 + S) = x + S,$$

while the second gives us

$$XY = (x+S)\left(x^2+S\right) = x^3+S,$$

and $x^3 - x$ does not have the form necessary for members of S. Thus, R/S is not a ring.

We will see that the absorption property of ideals *does* guarantee that the multiplication of cosets is well-defined, which opens the door to creating quotient rings.

Lemma 8.45. The "natural" addition and multiplication of cosets is well-defined whenever a subring is an ideal.

Proof. First we show that the operations are well-defined. Let $X, Y \in R/A$ and $w, x, y, z \in R$. such that $w + A = x + A = X$ and $y + A = z + A = Y$.

Is addition well-defined? The definition of the operation tells us both $X + Y = (x + y) + A$ and $X + Y = (w + z) + A$. By the hypothesis that $x + A = w + A$ and $y + A = z + A$, Lemma 3.37 implies that $x - w \in A$ and $y - z \in A$. By closure, $(x - w) + (y - z) \in A$. Using the properties of a ring,

$$(x+y) - (w+z) = (x-w) + (y-z) \in A.$$

Again from Lemma 3.37, $(x + y) + A = (w + z) + A$, so, by definition,

$$\begin{aligned}(x+A)+(y+A) &= (x+y)+A \\ &= (w+z)+A = (w+A)+(z+A).\end{aligned}$$

It does not matter, therefore, which representations we use for X and Y; the sum $X + Y$ has the same value, so addition in R/A is well-defined.

Is multiplication well-defined? Observe that $XY = (x + A)(y + A) = xy + A$. As explained above, $x - w \in A$ and $y - z \in A$. Let $a, \widehat{a} \in A$ such that $x - w = a$ and $y - z = \widehat{a}$; **from the absorption property of an ideal**, $ay \in A$, so

$$\begin{aligned}xy - wz &= (xy - xz) + (xz - wz) \\ &= x(y-z) + (x-w)z \\ &= x\widehat{a} + az \in A.\end{aligned}$$

Again from Lemma 3.37, $xy + A = wz + A$, and by definition

$$(x+A)(y+A) = xy + A = wz + A = (w+A)(z+A).$$

It does not matter, therefore, what representations we use for X and Y; the product XY has the same value, so multiplication in R/A is well-defined. □

Using an ideal to create a new ring

We now generalize the notion of *quotient groups* to rings, and prove some interesting properties of certain quotient groups that help explain various phenomena we observed in both group theory and ring theory.

Theorem 8.46. Let R be a ring, and A an ideal. Define addition and multiplication for R/A in the "natural" way: for all $X, Y \in R/A$ denoted as $x + A, y + A$ for some $x, y \in R$,

$$X + Y = (x + y) + A$$
$$XY = (xy) + A.$$

The set R/A is a ring under these operations, called the **quotient ring**.

Example 8.47. Recall that $\mathbb{Z}$ is a ring, and $d\mathbb{Z}$ is an ideal for any $d \in \mathbb{Z}$. Thus, $\mathbb{Z}/d\mathbb{Z}$ is a quotient ring, and $3 + d\mathbb{Z}$ is a coset.

Example 8.48. Recall that $\mathbb{Z}_2[x]$ is a ring. Let $A = \langle x^2 + 1 \rangle$. We construct the addition and multiplication tables for $\mathbb{Z}_2[x]/A$.

First, recall that $\mathbb{Z}_2[x]$ is a Euclidean domain, so we can perform division, so any polynomial can be written as $p = q(x^2+1) + r$, where $\deg r < \deg(x^2+1) = 2$. By absorption, $p - r = q(x^2+1) \in A$, so coset equality implies $[p] = [r]$. No remainder has degree more than 1, so every element of $\mathbb{Z}_2[x]$ has the form $[ax+b]$=$(ax+b) + \mathbb{Z}_2[x]$. That means there are only four elements of the quotient ring:

$$[0], [1], [x], [x+1].$$

Superficially, then, we get the following tables.

$+$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

$\times$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	x^2	x^2+x
$x+1$	0	$x+1$	x^2+x	x^2+1

(Notice that $2x = 0$ in $\mathbb{Z}_2$, which is why $(x+1)^2 = x^2+1$.)

While the multiplication table is accurate, it is unsatisfactory, because every element of the table can be written as a *linear* polynomial. Applying division again, we get

$$x^2 = 1 \cdot (x^2+1) + 1, \qquad x^2 + 1 = 1 \cdot (x^2+1) + 0, \qquad x^2 + x = 1 \cdot (x^2+1) + (x+1).$$

Thus, the multiplication table can be written *in canonical form* as follows.

$\times$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	1	$x+1$
$x+1$	0	$x+1$	$x+1$	0

Notation 8.49. When we consider elements of $X \in R/A$, we refer to the "usual representation" of X as $x + A$ for appropriate $x \in R$; that is, "big" X is represented by "little" x. Likewise, if $X = 3 + d\mathbb{Z}$, we often write $x = [3]$ or even $x = 3$.

You may remember that, when working in quotient rings, we made heavy use of Lemma 3.37 on page 121. You will see that here, too.

Proof of Theorem 8.46. We have already shown that addition and multiplication are well-defined in R/A, so we turn to showing that R/A is a ring. First we show the properties of a group under addition:

closure: Let $X, Y \in R/A$, with the usual representation. By substitution, $X + Y = (x + y) + A$. Since R, a ring, is closed under addition, $x + y \in R$. Thus $X + Y \in R/A$.

associative: Let $X, Y, Z \in R/A$, with the usual representation. Applying substitution and the associative property of R, we have

$$\begin{aligned}(X+Y)+Z &= ((x+y)+A)+(z+A)\\ &= ((x+y)+z)+A\\ &= (x+(y+z))+A\\ &= (x+A)+((y+z)+A)\\ &= X+(Y+Z).\end{aligned}$$

identity: We claim that $A = 0 + A$ is itself the identity of R/A; that is, $A = 0_{R/A}$. Let $X \in R/A$ with the usual representation. Indeed, substitution and the additive identity of R demonstrate this:

$$\begin{aligned}X+A &= (x+A)+(0+A)\\ &= (x+0)+A\\ &= x+A\\ &= X.\end{aligned}$$

inverse: Let $X \in R/A$ with the usual representation. We claim that $-x + A$ is the additive inverse of X. Indeed,

$$\begin{aligned}X+(-x+A) &= (x+(-x))+A\\ &= 0+A\\ &= A\\ &= 0_{R/A}.\end{aligned}$$

Hence $-x + A$ is the additive inverse of X.

Now we show that R/A satisfies the ring properties. Each property falls back on the corresponding property of R.

closure: Let $X, Y \in R/A$ with the usual representation. By definition and closure in R,

$$\begin{aligned}XY &= (x+A)(y+A)\\ &= (xy)+A\\ &\in R/A.\end{aligned}$$

associative: Let $X, Y, Z \in R/A$ with the usual representation. By definition and the associative property in R,

$$\begin{aligned}(XY)Z &= ((xy)+A)(z+A)\\ &= ((xy)z)+A\\ &= (x(yz))+A\\ &= (x+A)((yz)+A)\\ &= X(YZ).\end{aligned}$$

distributive: Let $X, Y, Z \in R/A$ with the usual representation. By definition and the distributive property in R,

$$\begin{aligned}X(Y+Z) &= (x+A)((y+z)+A)\\ &= (x(y+z))+A\\ &= (xy+xz)+A\\ &= ((xy)+A)+((xz)+A)\\ &= XY+XZ.\end{aligned}$$

Hence R/A is a ring. □

We conclude with an obvious property of quotient rings.

Proposition 8.50. If R is a ring with unity, then R/A is also a ring with unity, which is $1_R + A$.

Proof. You do it! See Exercise 8.54. □

In Section 3.5 we showed that one could define a group using the quotient group $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$. Since $\mathbb{Z}$ is a ring and $n\mathbb{Z}$ is an ideal of $\mathbb{Z}$ by Exercise 8.14, it follows that $\mathbb{Z}_n$ is also a ring. Of course, you had already argued this in Exercise 7.13.

Exercises.

Exercise 8.51. Compute addition and multiplication tables for
(a) $\mathbb{Z}_2[x]/\langle x\rangle$;
(b) $\mathbb{Z}_2[x]/\langle x^2+x\rangle$;
(c) $\mathbb{Z}_2[x,y]/\langle x^2, y^2\rangle$.

Exercise 8.52. The example at the beginning of the section came from the ring $\mathbb{Z}[x]$. Show that in the ring of integers, *any* subring creates a quotient ring.

Exercise 8.53. Let $R = \mathbb{Z}_5[x]$ and $I = \langle x^2+2x+2\rangle$.
(a) Explain why $(x^2+x+3)+I = (4x+1)+I$.
(b) Find a factorization of x^2+2x+2 in R.
(c) Find two non-zero elements of R/I whose product is the zero element of R/I.
(d) Explain why R/I is, therefore, not an integral domain, and, therefore, not a field.

Exercise 8.54. Prove Proposition 8.50.

Exercise 8.55. Suppose R is a ring, and A, B ideals of R. Let $Q = R/A$, and $C = \{b + A : b \in B\}$. Is C an ideal of Q?

8.4: When is a quotient ring an integral domain or a field?

You found in Exercise 7.31 that $\mathbb{Z}_n$ is not, in general, an integral domain, let alone a field. The curious thing is that we started with an integral domain $\mathbb{Z}$, computed a quotient ring by an ideal $n\mathbb{Z}$ that satisfies the zero product property, yet we *still* didn't end up with an integral domain! Why did this happen? We found that it occurred when n was not irreducible, which also means it was not prime.

We can view this as a relationship not just of divisibility, but of ideals. From the definition of an irreducible integer, we know that n is irreducible if its only divisors are ± 1 and $\pm n$. Lemma 8.27 translates this into the language of ideals as this remarkable statement:

The only ideals "larger" than $\langle n \rangle$ are $\mathbb{Z}$ (of course) and $\langle n \rangle$ itself.

In other words, the ideal generated by an irreducible number is the "largest" sort of proper ideal in $\mathbb{Z}$. We ought to generalize that.

On the other hand, while the notions of "prime" and "irreducible" are equivalent in the integers, they may not mean the same thing in all rings. For example, in $\mathbb{Z}_6$,

- 2 is *not* irreducible, since $2 = 20 = 4 \cdot 5$, but
- 2 seems to be "prime", since from the 36 products possible in $\mathbb{Z}_6$, the only ones where 2 does not divide one of the factors are

$$1 \times 1,\ 1 \times 3,\ 1 \times 5,\ 3 \times 3,\ 3 \times 5,\ 5 \times 5,$$

  and 2 divides none of those products, either.

This observation raises a number of questions, but looking at them carefully would lead us astray for now, so we delay them until Chapter 10; for now, however, we prefer to focus on ideals. Recall that in algebra, n is prime if any time $n \mid ab$, then $n \mid a$ or $n \mid b$. Lemma 8.27 translates this into the language of ideals as this equally remarkable statement:

If $\langle n \rangle$ contains $\langle ab \rangle$, then it must contain $\langle a \rangle$ or $\langle b \rangle$.

Maximal and prime ideals

Let R be a ring.

Definition 8.56. A proper ideal A of R is a **maximal ideal** if no other proper ideal of R contains A.

Another way of expressing that A is maximal is the following: for any other ideal I of R, $A \subseteq I$ implies that $A = I$ or $I = R$.

Example 8.57. In Exercise 8.16 you showed that all ideals of $\mathbb{Z}$ have the form $n\mathbb{Z}$ for some $n \in \mathbb{Z}$. Are any of these (or all of them) maximal ideals?

Let $n \in \mathbb{Z}$ and suppose that $n\mathbb{Z}$ is maximal. Certainly $n \neq 0$, since $2\mathbb{Z} \not\subseteq \{0\}$. We claim that $|n|$ is irreducible; in other words, n is divisible only by $\pm 1, \pm n$. To see this, recall Lemma 8.4:

$m \in \mathbb{Z}$ is a divisor of n iff $n\mathbb{Z} \subseteq m\mathbb{Z}$. Since $n\mathbb{Z}$ is maximal, either $m\mathbb{Z} = \mathbb{Z}$ or $m\mathbb{Z} = n\mathbb{Z}$. In the first case, $m = \pm 1$; in the second case, $m = \pm n$. Hence $|n|$ is irreducible.

For prime ideals, you need to recall from Exercise 8.22 that for any two ideals A, B of R, AB is also an ideal.

Definition 8.58. A proper ideal P of R is a **prime ideal** if for every two ideals A, B of R we know that

$$\text{if } AB \subseteq P \text{ then } A \subseteq P \text{ or } B \subseteq P.$$

Definition 8.58 might remind you of our definition of prime integers from page 6.31. Indeed, the two are connected.

Example 8.59. Let $n \in \mathbb{Z}$ be a prime integer. Let $a, b \in \mathbb{Z}$ such that $p \mid ab$. Hence $p \mid a$ or $p \mid b$. Suppose that $p \mid a$.

Let's turn our attention to the corresponding ideals. Since $p \mid ab$, Lemma 8.4 tells us that $(ab)\mathbb{Z} \subseteq p\mathbb{Z}$. It is routine to show that $(ab)\mathbb{Z} = (a\mathbb{Z})(b\mathbb{Z})$, but in case you think otherwise, it's also Exercise 8.41. Put $A = a\mathbb{Z}$, $B = b\mathbb{Z}$, and $P = p\mathbb{Z}$; thus $AB \subseteq P$.

Recall that $p \mid a$; applying Lemma 8.4 again, we have $A = a\mathbb{Z} \subseteq p\mathbb{Z} = P$.

Conversely, if n is not prime, $n\mathbb{Z}$ is not a prime ideal: for example, $6\mathbb{Z}$ is not a prime ideal because $(2\mathbb{Z})(3\mathbb{Z}) \subseteq 6\mathbb{Z}$ but by Lemma 8.17 neither $2\mathbb{Z} \subseteq 6\mathbb{Z}$ nor $3\mathbb{Z} \subseteq 6\mathbb{Z}$. This can be generalized easily to all integers that are not prime: see Exercise 8.67.

Let's summarize our examples. We found in Example 8.57 that an ideal in $\mathbb{Z}$ is maximal iff it is generated by a prime integer, and in Example 8.59 we argued that an ideal is prime iff it is generated by a prime integer. We learned in Theorem 6.33 that an integer is prime if and only if it is irreducible. Thus, an ideal is maximal if and only if it is prime — in the ring of integers, anyway.

What about other rings? Showing a maximal ideal is prime doesn't require too many additional constraints.

Theorem 8.60. Let R be a ring. If R has unity, then every maximal ideal is prime.

Proof. Assume that R has unity. We want to show that every maximal ideal is prime, so let M be a maximal ideal of R. Let A, B be any two ideals of R such that $AB \subseteq M$. We claim that $A \subseteq M$ or $B \subseteq M$.

If $A \subseteq M$, then we are done, so assume that $A \not\subseteq M$. Recall from Exercise 8.22 that $A + M$ is also an ideal. In addition, it should be clear that $M \subseteq A + M$. (If it isn't clear, try it. It really isn't hard.) Since M is maximal, $A + M = M$ or $A + M = R$. Which is it?

We claim that $A + M = R$. To see why, observe that if $A + M = M$, then for any $a \in A$ and any $m \in M$ we could find $m' \in M$ such that $a + m = m'$. We can rewrite this as $a = m' - m$; closure tells us $m' - m \in M$, and substitution gives us $a \in M$. But a was arbitrary, implying that $A \subseteq M$, contradicting the hypothesis that $A \subsetneq M$. Thus, $A + M \neq M$, which means $A + M = R$.

Since R has unity, $1_R \in R = A + M$, so there exist $a \in A$, $m \in M$ such that

$$1_R = a + m. \tag{28}$$

Let $b \in B$. Multiply both sides of (28) by b; we have

$$\begin{aligned} 1_R \cdot b &= (a+m)\,b \\ b &= ab + mb. \end{aligned}$$

Recall that $AB \subseteq M$; since $ab \in AB$, $ab \in M$. Likewise, absorption implies that $mb \in M$. Closure of addition implies that $ab + mb \in M$. Substitution implies that $b \in M$. Since b was arbitrary in B, $B \subseteq M$.

We assumed that $AB \subseteq M$, and found that $A \subseteq M$ or $B \subseteq M$. Thus, M is prime. □

Is the requirement that the ring have unity that important? Yes, even in simple rings.

Theorem 8.61. If R is a ring without unity, then maximal ideals might not be prime.

Proof. The proof is by counterexample: we use $2\mathbb{Z}$, a ring without unity. We claim that $4\mathbb{Z}$ is a maximal ideal of $R = 2\mathbb{Z}$ that is not prime:

closed under subtraction? Let $x, y \in 4\mathbb{Z}$. By definition of $4\mathbb{Z}$, $x = 4a$ and $y = 4b$ for some $a, b \in \mathbb{Z}$. Using the distributive property and substitution, we have $x - y = 4a - 4b = 4(a-b) \in 4\mathbb{Z}$.

absorbs multiplication? Let $x \in 4\mathbb{Z}$ and $r \in 2\mathbb{Z}$. By definition of $4\mathbb{Z}$, $x = 4q$ for some $q \in \mathbb{Z}$. By substitution, the associative property, and the commutative property of integer multiplication, $rx = 4(rq) \in 4\mathbb{Z}$.

maximal? Let A be any ideal of $2\mathbb{Z}$ such that $4\mathbb{Z} \subseteq A$. Choose $a \in \mathbb{Z}$ such that $A = a\mathbb{Z}$. (We can do this thanks to Exercise 8.35.) By Lemma 8.4, $a \mid 4$, so the only possible values of a are ± 1, ± 2, and ± 4. Certainly $a \neq \pm 1$; after all, $\pm 1 \notin R$. If $a = \pm 4$, then $A = 4\mathbb{Z}$. If $a = \pm 2$, then $A = R$. We took an arbitrary ideal A such that $4\mathbb{Z} \subseteq A$, and found that $A = 4\mathbb{Z}$ or $A = 2\mathbb{Z}$, the entire ring. Hence, $4\mathbb{Z}$ is maximal.

prime? An easy counterexample does the trick: $(2\mathbb{Z})(2\mathbb{Z}) = 4\mathbb{Z}$, but $2\mathbb{Z} \not\subseteq 4\mathbb{Z}$.

□

The situation with prime ideals is less… well, to be cute about it, "less than ideal".

Theorem 8.62. A prime ideal is not necessarily maximal, even in a ring with unity.

Proof. Recall that $R = \mathbb{C}[x, y]$ is a ring with unity, and that $I = \langle x \rangle$ is an ideal of R.

We claim that I is a prime ideal of R. Let A, B be ideals of R such that $AB \subseteq I$. If $A \subseteq I$, then we are done, so suppose that $A \not\subseteq I$. We need to show that $B \subseteq I$. Let $a \in A \backslash I$. For any $b \in B$, $ab \in AB \subseteq I = \langle x \rangle$, so $ab \in \langle x \rangle$. This implies that $x \mid ab$; let $q \in R$ such that $qx = ab$. Write $a = f \cdot x + a'$ and $b = g \cdot x + b'$ where $a', b' \in R \backslash I$; that is, a' and b' are polynomials with *no* terms that are multiples of x. By substitution,

$$\begin{aligned} ab &= (f\cdot x + a')(g\cdot x + b') \\ qx &= (f\cdot x)\cdot(g\cdot x) + a'\cdot(g\cdot x) \\ &\quad + b'\cdot(f\cdot x) + a'\cdot b' \\ (q - fg - a'g - b'f)x &= a'b'. \end{aligned}$$

Hence $a'b' \in \langle x\rangle$. However, no term of a' or b' is a multiple of x, so no term of $a'b'$ is a multiple of x. The only element of $\langle x\rangle$ that satisfies this property is 0. Hence $a'b' = 0$, which by the zero product property of complex numbers implies that $a' = 0$ or $b' = 0$.

Which is it? If $a' = 0$, then $a = f\cdot x + 0 \in \langle x\rangle = I$, which contradicts the assumption that $a \in A\backslash I$. Hence $a' \neq 0$, implying that $b' = 0$, so $b = gx + 0 \in \langle x\rangle = I$. Since b is arbitrary, this holds for all $b \in B$; that is, $B \subseteq I$.

We took two arbitrary ideals such that $AB \subseteq I$ and showed that $A \subseteq I$ or $B \subseteq I$; hence $I = \langle x\rangle$ is prime. However, I is not maximal, since

- $y \notin \langle x\rangle$, implying that $\langle x\rangle \subsetneq \langle x,y\rangle$; and
- $1 \notin \langle x,y\rangle$, implying that $\langle x,y\rangle \notin \mathbb{C}[x,y]$.

□

So prime and maximal ideals need not be equivalent. In Chapter 10, we will find conditions on a ring that ensure that prime and maximal ideals are equivalent.

A criterion that determines when a quotient ring is an integral domain or a field

We can now answer the question that opened this section.

Theorem 8.63. If R is a ring with unity and M is a maximal ideal of R, then R/M is a field. The converse is also true.

Proof. ($\Rightarrow$) Assume that R is a ring with unity and M is a maximal ideal of R. Let $X \in R/M$ and assume that $X \neq M$; that is, X is non-zero. Since $X \neq M$, $X = x + M$ for some $x \notin M$. By Exercise 8.22, $\langle x\rangle + M$ is also an ideal. Since $x \notin M$, we know that $M \subsetneq \langle x\rangle + M$. Since M is a maximal ideal, $M \subsetneq \langle x\rangle + M = R$. Since R is a ring with unity, $1 \in R$ by definition. Substitution implies that $1 \in \langle x\rangle + M$, so there exist $h \in R$, $m \in M$ such that $1 = hx + m$. Rewrite this as $1 - hx = m \in M$; by Lemma 3.37,

$$1 + M = hx + M = (h + M)(x + M).$$

In other words, $h + M$ is a multiplicative inverse of $X = x + M$ in R/M. Since X was an arbitrary non-zero element of R/M, every element of R/M has a multiplicative inverse, and R/M is a field.

($\Leftarrow$) For the converse, assume that R/M is a field. We want to show that M is maximal, so let N be any ideal of R such that $M \subseteq N \subseteq R$. If $M = N$, then we are done, so assume that $M \neq N$. We want to show that $N = R$. Let $x \in N\backslash M$; then $x + M \neq M$, and since R/M is a field, $x + M$ has a multiplicative inverse; call it $Y = y + M$. That is,

$$1 + M = (x + M)(y + M) = (xy) + M,$$

which by Lemma 3.37 implies that $xy-1 \in M$. Let $m \in M$ such that $xy-1=m$; then $1 = xy-m$. Now, $xy \in N$ by absorption, and $m \in N$ by inclusion. (After all, $x \in N$ and $m \in M \subsetneq N$.) Closure of the subring N implies that $1 = xy-m \in N$, and Exercise 8.20 implies that $N = R$. Since N was an arbitrary ideal that contained M properly, M is maximal. □

A similar property holds true for prime ideals.

Theorem 8.64. If R is a ring with unity and P is a prime ideal of R, then R/P is an integral domain. The converse is also true.

Proof. (⇒) Assume that R is a ring with unity and P is a prime ideal of R. Let $X, Y \in R/P$ with the usual representation, and assume that $XY = 0_{R/P} = P$. By definition of the operation, $XY = (xy) + P$; by Lemma 3.37, $xy \in P$. We claim that this implies that $x \in P$ or $y \in P$.

Assume to the contrary that $x, y \notin P$. For any $z \in \langle x \rangle \langle y \rangle$, we have $z = \sum_{k=1}^{m} (h_k x)(q_k y)$ for an appropriate choice of $m \in \mathbb{N}^+$ and $h_k, q_k \in R$. Recall that R is commutative, which means $z = xy \sum (h_k q_k)$. We determined above that $xy \in P$, so by absorption, $z \in P$. Since z was arbitrary in $\langle x \rangle \langle y \rangle$, we conclude that $\langle x \rangle \langle y \rangle \subseteq P$. Now P is a prime ideal, so $\langle x \rangle \subseteq P$ or $\langle y \rangle \subseteq P$; without loss of generality, $\langle x \rangle \subseteq P$. Since R has unity, $x \in \langle x \rangle$, and thus $x \in P$. Lemma 3.37 now implies that $x + P = P$. Thus $X = 0_{R/P}$.

We took two arbitrary elements of R/P, and showed that if their product was the zero element of R/P, then one of those elements had to be P, the zero element of R/P. That is, R/P is an integral domain.

(⇐) For the converse, assume that R/P is an integral domain. Let A, B be two ideals of R, and assume that $AB \subseteq P$. Assume that $A \not\subseteq P$ and let $a \in A \backslash P$; by coset equality, $a + P \neq P$. Let $b \in B$ be arbitrary. By hypothesis, $ab \in AB \subseteq P$, so here coset equality implies that

$$(a+P)(b+P) = (ab) + P = P \quad \forall b \in B.$$

Since R/P is an integral domain, $P = 0_{R/P}$, and $a + P \neq P$, we conclude that $b + P = P$. By coset equality, $b \in P$. Since b was arbitrary, this holds for all $b \in B$; hence, $B \subseteq P$.

We took two arbitrary ideals of R, and showed that if their product was a subset of P, then one of them had to be a subset of P. Thus P is a prime ideal. □

Have you noticed that this gives us an alternate proof of Theorem 8.60?

Corollary 8.65. In a ring with unity, every maximal ideal is prime, but the converse is not necessarily true.

Proof. Let R be a ring with unity, and M a maximal ideal. By Theorem 8.63, R/M is a field. By Theorem 7.24, R/M is an integral domain. By Theorem 8.64, M is prime.

The converse is not necessarily true, as not every integral domain is a field. □

Chapter Exercises.

Exercise 8.66. Determine necessary and sufficient conditions on a ring R such that in $R[x,y]$:
(a) the ideal $I = \langle x \rangle$ is prime;

(b) the ideal $I = \langle x, y \rangle$ is maximal.

Exercise 8.67. Let $n \in \mathbb{Z}$ be an integer that is not prime. Show that $n\mathbb{Z}$ is not a prime ideal.

Exercise 8.68. Show that $\{[0], [4]\}$ is a proper ideal of $\mathbb{Z}_8$, but that it is not maximal. Then find a maximal ideal of $\mathbb{Z}_8$.

Exercise 8.69. Find all the maximal ideals of $\mathbb{Z}_{12}$. Are they prime? How do you know?

Exercise 8.70. Let $\mathbb{F}$ be a field, and $a_1, a_2 \ldots, a_n \in \mathbb{F}$.

(a) Show that the ideal $\langle x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n \rangle$ is both a prime ideal and a maximal ideal of $\mathbb{F}[x_1, x_2, \ldots, x_n]$.

(b) Use Exercise 8.24 to describe the common root(s) of this ideal.

Exercise 8.71. Consider the ideal $I = \left\langle x^2 + 1 \right\rangle$ in $R = \mathbb{R}[x]$. The purpose of this exercise is to show that I is maximal.

(a) Explain why $\left(x^2 + x\right) + I = (x - 1) + I$.

(b) Explain why every $f \in R/I$ has the form $r + I$ for some $r \in R$ such that $\deg r < 2$.

(c) Part (b) implies that every element of R/I can be written in the form $f = (ax + b) + I$ where $a, b \in \mathbb{R}$. Show that if $f + I$ is a nonzero element of R/I, then $a^2 + b^2 \neq 0$.

(d) Let $f + I \in R/I$ be nonzero, and find $g + I \in R/I$ such that $g + I = (f + I)^{-1}$; that is, $(fg) + I = 1_{R/I}$.

(e) Explain why part (d) shows that I is maximal.

(f) Explain why $\left\langle x^2 + 1 \right\rangle$ is not even prime if $R = \mathbb{C}[x]$, let alone maximal. Show further that this is because the observation in part (c) no longer holds in $\mathbb{C}$.

Exercise 8.72. Let $\mathbb{F}$ be a field, and $f \in \mathbb{F}[x]$ be any polynomial that does not factor in $\mathbb{F}[x]$. Show that $\mathbb{F}[x] / \langle f \rangle$ is a field.

Exercise 8.73. Recall the ideal $I = \left\langle x^2 + y^2 - 4, xy - 1 \right\rangle$ of Exercise 8.5. We want to know whether this ideal is maximal. The purpose of this exercise is to show that it is not so "easy" to accomplish this as it was in Exercise 8.71.

(a) Explain why someone might think naïvely that every $f \in R/I$ has the form $r + I$ where $r \in R$ and $r = bx + p(y)$, for appropriate $b \in \mathbb{C}$ and $p \in \mathbb{C}[y]$; in the same way, someone might think naïvely that every distinct polynomial r of that form represents a distinct element of R/I.

(b) Show that, to the contrary, $1 + I = \left(x + y^3 - 4y + 1\right) + I$.

Exercise 8.74. The **Krull dimension** of a ring is the length of the longest chain of prime ideals. If R has chains $\{0\} \subsetneq P_1 \subsetneq P_2 \subsetneq R$ and $\{0\} \subsetneq Q \subsetneq R$, where P_1, P_2, and Q are all prime ideals of R, then the Krull dimension of R is 2.

(a) Show that the Krull dimension of $\mathbb{Z} = 1$.

(b) Show that the Krull dimension of $\mathbb{Q} = 0$.

(c) Show that the Krull dimension of $\mathbb{Z}[x] = 2$.

(d) Show that the Krull dimension of $\mathbb{C}[x] = 1$.

(e) Show that the Krull dimension of $\mathbb{C}[x, y] = 2$.

8.5: Ring isomorphisms

As with groups and rings, it is often useful to show that two rings have the same ring structure. With monoids and groups, we defined *isomorphisms* to do this. We will do the same thing with rings. However, ring homomorphisms are a little more complicated, as rings have two operations, rather than one.

Ring homomorphisms and their properties

Definition 8.75. Let R and S be rings. A function $f : R \to S$ is a **ring homomorphism** if for all $a, b \in R$

$$f(a+b) = f(a) + f(b)$$

and

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

If, in addition, f is one-to-one and onto, we call it a **ring isomorphism**.

Right away, you should see that a ring homomorphism is a special type of group homomorphism with respect to addition. Even if the ring has unity, however, *it might not* be a monoid homomorphism with respect to multiplication, because there is no guarantee that $f(1_R) = 1_S$.

Example 8.76. Let $f : \mathbb{Z} \to \mathbb{Z}_2$ by $f(x) = [x]$. The homomorphism properties are satisfied:

$$f(x+y) = [x+y] = [x] + [y] = f(x) + f(y)$$

and

$$f(xy) = [xy] = [x][y] = f(x) f(y).$$

Notice that f is onto, but it is certainly not one-to-one, inasmuch as $f(0) = f(2)$.

On the other hand, consider Example 8.77.

Example 8.77. Let $f : \mathbb{Z} \to 2\mathbb{Z}$ by $f(x) = 4x$. In Example 4.3 on page 142, we showed that this was a homomorphism of groups. However, it is *not* a homomorphism of rings, because it does not preserve multiplication:

$$f(xy) = 4xy \quad \text{but} \quad f(x) f(y) = (4x)(4y) \neq f(xy).$$

Example 8.77 drives home the point that rings are more complicated than groups on account of having two operations. It is harder to show that two rings are homomorphic, and therefore harder to show that they are isomorphic. This is especially interesting in this example, since we had shown earlier that $\mathbb{Z} \cong n\mathbb{Z}$ *as groups* for all nonzero n. If this is the case with rings, then we have to find some other function between the two. Theorem 8.78 shows that this is not possible, in a way that should not surprise you.

Theorem 8.78. Let R be a ring with unity. If there exists an onto homomorphism between R and another ring S, then S is also a ring with unity.

Proof. Let S be a ring such that there exists a homomorphism f between R and S. We claim that $f(1_R)$ is an identity for S.

Let $y \in S$; the fact that R is onto implies that $f(x) = y$ for some $x \in R$. Applying the homomorphism property,

$$y = f(x) = f(x \cdot 1_R) = f(x) f(1_R) = y \cdot f(1_R).$$

A similar argument shows that $y = f(1_R) \cdot y$. Since y was arbitrary in S, $f(1_R)$ is an identity for S. □

We can deduce from this that $\mathbb{Z}$ and $n\mathbb{Z}$ are not isomorphic as rings whenever $n \neq 1$:

- to be isomorphic, there would have to exist an onto function from $\mathbb{Z}$ to $n\mathbb{Z}$;
- $\mathbb{Z}$ has a multiplicative identity;
- by Theorem 8.78, $n\mathbb{Z}$ would also have to have a multiplicative identity;
- but $n\mathbb{Z}$ does not have a multiplicative identity when $n \neq 1$.

Here are more useful properties of a ring homomorphism.

Theorem 8.79. Let R and S be rings, and f a ring homomorphism from R to S. Each of the following holds:
(A) $f(0_R) = 0_S$;
(B) for all $x \in R$, $f(-x) = -f(x)$;
(C) for all $x \in R$, if x has a multiplicative inverse and f is onto, then $f(x)$ has a multiplicative inverse, and $f(x^{-1}) = f(x)^{-1}$.

Proof. You do it! See Exercise 8.89. □

We have not yet encountered an example of a ring isomorphism, so let's consider one.

Example 8.80. Let $\mathbb{F}$ be any field, and $p = ax + b \in \mathbb{F}[x]$, where $a \neq 0$. Recall from Exercise 8.72 that $\langle p \rangle$ is maximal in $\mathbb{F}[x]$. For convenience, we will write $R = \mathbb{F}[x]$ and $I = \langle p \rangle$; by Theorem 8.63, R/I is a field.

Are $\mathbb{F}$ and R/I isomorphic? Let $f : \mathbb{F} \to R/I$ in the following way: let $f(c) = c + I$ for every $c \in \mathbb{F}$. Is f a homomorphism?

Homomorphism property? Let $c, d \in \mathbb{F}$; using the definition of f and the properties of coset addition,

$$\begin{aligned} f(c+d) &= (c+d) + I \\ &= (c+I) + (d+I) = f(c) + f(d). \end{aligned}$$

Similarly,

$$f(cd) = (cd) + I = (c+I)(d+I) = f(c) f(d).$$

One-to-one? Let $c, d \in \mathbb{F}$ and suppose that $f(c) = f(d)$. Then $c + I = d + I$; by Lemma 3.37, $c - d \in I$. By closure, $c - d \in \mathbb{F}$, while $I = \langle ax + b \rangle$ is the set of all multiples of $ax + b$. Since $a \neq 0$, the only rational number in I is 0, which implies that $c - d = 0$, so $c = d$.

Onto? Let $X \in R/I$; let $p \in R$ such that $X = p + I$. Divide p by $ax + b$ to obtain

$$p = q(ax + b) + r$$

where $q, r \in R$ and $\deg r < \deg(ax+b) = 1$. Since $ax+b \in I$, absorption tells us that $q(ax+b) \in I$, so

$$\begin{aligned} p + I &= [q(ax+b) + r] + I \\ &= [q(ax+b) + I] + (r + I) \\ &= I + (r + I) \\ &= r + I. \end{aligned}$$

Now, $\deg r < 1$ implies that $\deg r = 0$, or in other words, r is a constant. The constants of $R = \mathbb{F}[x]$ are elements of $\mathbb{F}$, so $r \in \mathbb{F}$. Hence

$$f(r) = r + I = p + I,$$

and f is onto.

We have shown that there exists a one-to-one, onto ring homomorphism from $\mathbb{F}$ to R/I; as a consequence, $\mathbb{F}$ and R/I are isomorphic as rings.

The isomorphism theorem for rings

We now consider the isomorphism theorem for groups (Theorem 4.46) in the context of rings. To do this, we need to revisit the definition of a kernel.

Definition 8.81. Let R and S be rings, and $f : R \to S$ a homomorphism of rings. The **kernel** of f, denoted $\ker f$, is the set of all elements of R that map to 0_S. That is,

$$\ker f = \{x \in R : f(x) = 0_S\}.$$

You will show in Exercise 8.91 that $\ker f$ is an ideal of R, and that the function $g : R \to R/\ker f$ by $g(x) = x + \ker f$ is a homomorphism of rings.

Theorem 8.82. Let R, S be rings, and $f : R \to S$ an onto homomorphism. Let $g : R \to R/\ker f$ be the natural homomorphism $g(r) = r + \ker f$. There exists an isomorphism $h : R/\ker f \to S$ such that $f = h \circ g$.

Proof. Define h by $h(X) = f(x)$ where $X = x + \ker f$. Is f an isomorphism? Since its domain consists of cosets, we must show first that it's well-defined:

well-defined? Let $X \in R/\ker f$ and let $x, y \in R$ such that $X = x + \ker f = y + \ker f$ — that is, $x + \ker f$ and $y + \ker f$ are two representations of the same coset, X. We must show that $h(X)$ has the same value regardless of which representation we use. By Lemma 3.37, $x - y \in \ker f$. From the definition of the kernel, $f(x-y) = 0_S$. We can apply Theorem 8.79 to see that

$$\begin{aligned} 0_S &= f(x-y) \\ &= f(x+(-y)) \\ &= f(x)+f(-y) \\ &= f(x)+[-f(y)] \\ f(y) &= f(x). \end{aligned}$$

By substitution, we have $h(y+\ker f) = f(y) = f(x) = h(x+\ker f)$. In other words, the representation of X does not affect the value of h, and h is well-defined.

homomorphism property? Let $X, Y \in R/\ker f$ and consider the representations $X = x+\ker f$ and $Y = y+\ker f$. Since f is a ring homomorphism,

$$\begin{aligned} h(X+Y) &= h((x+\ker f)+(y+\ker f)) \\ &= h((x+y)+\ker f) \\ &= f(x+y) \\ &= f(x)+f(y) \\ &= h(x+\ker f)+f(y+\ker f) \\ &= h(X)+h(Y). \end{aligned}$$

Similarly,

$$\begin{aligned} h(XY) &= h((x+\ker f)\cdot(y+\ker f)) \\ &= h((xy)+\ker f) \\ &= f(xy) \\ &= f(x)f(y) \\ &= h(x+\ker f)\cdot f(y+\ker f) \\ &= h(X)\cdot h(Y). \end{aligned}$$

Thus h is a ring homomorphism.

one-to-one? Let $X, Y \in R/\ker f$ and suppose that $h(X) = h(Y)$. Let $x, y \in R$ such that $X = x+\ker f$ and $Y = y+\ker y$. By the definition of h, $f(x) = f(y)$. Applying Theorem 8.79, we see that

$$\begin{aligned} f(x) = f(y) &\Longrightarrow f(x)-f(y) = 0_S \\ &\Longrightarrow f(x-y) = 0_S \\ &\Longrightarrow x-y \in \ker f \\ &\Longrightarrow x+\ker f = y+\ker f, \end{aligned}$$

so $X = Y$. We have shown that if $h(X) = h(Y)$, then $X = Y$. By definition, h is one-to-one.

onto? Let $y \in S$. Since f is onto, there exists $x \in R$ such that $f(x) = y$. Then $h(x+\ker f) = f(x) = y$. We have shown that an arbitrary element of the range S has a preimage in the domain. By definition, h is onto.

We have shown that h is a well-defined, one-to-one, onto homomorphism of rings. Thus h is an isomorphism from $R/\ker f$ to S. □

Example 8.83. Let $f : \mathbb{Q}[x] \to \mathbb{Q}$ by $f(p) = p(2)$ for any polynomial $p \in \mathbb{Q}[x]$. That is, f maps any polynomial to the value that polynomial gives for $x = 2$. For example, if $p = 3x^3 - 1$, then $p(2) = 3(2)^3 - 1 = 23$, so $f(3x^3 - 1) = 23$.

Is f a homomorphism? For any polynomials $p, q \in \mathbb{Q}[x]$, we have

$$f(p+q) = (p+q)(2);$$

applying a property of polynomial addition, we have

$$f(p+q) = (p+q)(2) = p(2) + q(2) = f(p) + f(q).$$

A similar property of polynomial multiplication gives

$$f(pq) = (pq)(2) = p(2) \cdot q(2) = f(p)f(q),$$

so f is a homomorphism.

Is f onto? Let $a \in \mathbb{Q}$; we need a polynomial $p \in \mathbb{Q}[x]$ such that $p(2) = a$. The easiest way to do this is to use a linear polynomial, and $p = x + (a-2)$ will work, since

$$f(p) = p(2) = 2 + (a-2) = a.$$

We took an arbitrary element of the range $\mathbb{Q}$, and showed that it has a preimage in the domain. By definition, f is onto.

Is f one-to-one? The answer is *no*. We already saw that $f(3x^3 - 1) = 23$, and from our work showing that f is onto, we deduce that $f(x+21) = 23$, so f is not one-to-one.

Let's apply Theorem 8.82 to obtain an isomorphism. First, identify $\ker f$: it consists of all the polynomials $p \in \mathbb{Q}[x]$ such that $p(2) = 0$. The Factor Theorem (7.45) implies that $x - 2$ must be a factor of any such polynomial. In other words,

$$\ker f = \{p \in \mathbb{Q}[x] : (x-2) \text{ divides } p\} = \langle x-2 \rangle.$$

Since $\ker f = \langle x-2 \rangle$, Theorem 8.82 tells us that there exists an isomorphism between the quotient ring $\mathbb{Q}[x] / \langle x-2 \rangle$ and $\mathbb{Q}$.

Notice, as in Example 8.80, that $x-2$ is a linear polynomial. Linear polynomials do not factor. By Exercise 8.72, $\langle x-2 \rangle$ is a maximal ideal; so $\mathbb{Q}[x] / \langle x-2 \rangle$ must be a field—as is $\mathbb{Q}$.

A construction of the complex numbers

We conclude this section by showing that the complex numbers can be viewed not only as an "abstract" extension of $\mathbb{R}$ by an "imaginary" number $i = \sqrt{-1}$, but also as a "concrete" construction: a quotient ring of $\mathbb{R}[x]$. This not only gives you an exciting new view of the complex numbers, but also suggests how we can "solve" polynomial equations in general.

I assume you already know the basics of the complex number system: namely, $i^2 = -1$, and any complex number number takes the form $a + bi$ for some $a, b \in \mathbb{R}$. Addition and multiplication of complex numbers follow very simple rules:

$$(a+bi)+(c+di)=(a+c)+(b+d)i \quad \text{and} \quad (a+bi)(c+di)=(ac-bd)+(ad+bc)i.$$

We can get the same behavior out of a quotient ring. Let $I = \langle x^2+1 \rangle$ and $\mathcal{C} = \mathbb{R}[x]/I$. We claim that $\mathcal{C} \cong \mathbb{C}$.

We will construct an explicit isomorphism in just a moment, but first let's look at how arithmetic in $\mathcal{C}$ mimics the properties of $\mathbb{C}$. Start off by considering the elements of R.

Proposition 8.84. Let $P \in R$; by definition of a quotient ring, P has the form $p+I$, where $p \in \mathbb{R}[x]$. Without loss of generality, we may assume that $\deg p < 2$.

Proof. Since $\mathbb{R}[x]$ is a Euclidean domain, we can find $q, r \in \mathbb{R}[x]$ such that $p = q(x^2+1)+r$ and $r = 0$ or $\deg r < 2$. Rewrite the equation as $r = p - q(x^2+1)$. By substitution,

$$p+I = \left(q\left(x^2+1\right)+r\right)+I.$$

Arithmetic in a quotient ring allows us to rewrite this as

$$p+I = \left[q\left(x^2+1\right)+I\right]+(r+I).$$

Recall that $I = \langle x^2+1 \rangle$. By absorption, $q(x^2+1) \in I$, so $I = q(x^2+1)+I$. Since $I = 0_{\mathcal{C}}$, we can rewrite the above equation as $p+I = r+I$. In other words, we can write r in place of p, and obtain the same result as using p. Thus, we can assume that $\deg p = \deg r < 2$. □

Thanks to Proposition 8.84, we can write any element of $\mathcal{C}$ as $(bx+a)+I$, where $a, b \in \mathbb{R}$: its degree is less than 2, and its coefficients are real. Even this is a bit much work, though. To simplify the writing further, we notice that one element of $\mathcal{C}$ has a very nice property.

Proposition 8.85. $(x+I)^2 = -1+I$.

Proof. You do it! See Exercise 8.90. □

This motivates us to adopt a highly suggestive notation.

Notation 8.86. We will write each $(bx+a)+I \in \mathcal{C}$ as $a+b\mathbf{i}$. If $b = 0$, we will write a and understand that we mean $a+0\mathbf{i}$ or $a+I$.

This means that we can write $\mathbf{i} = x+I$ and $\mathbf{i}^2 = -1$. Exploring the resulting arithmetic, we find some astonishing parallels to complex arithmetic:

$$\begin{aligned}(a+b\mathbf{i})+(c+d\mathbf{i}) &= [(bx+a)+I]+[(dx+c)+I]\\ &= [(b+d)x+(a+c)]\\ &= (a+c)+(b+d)\mathbf{i}\end{aligned}$$

and

$$\begin{aligned}(a+b\mathbf{i})(c+d\mathbf{i}) &= [(bx+a)+I][(dx+c)+I]\\ &= (bx+a)(dx+c)+I\\ &= \left[bdx^2+(bc+ad)x+ac\right]+I\\ &= bd\left(x^2+I\right)+[((bc+ad)x+ac)+I]\\ &= (-bd+I)+[((bc+ad)x+ac)+I]\\ &= [(bc+ad)x+(ac-bd)]+I\\ &= (ac-bd)+(ad+bc)\mathbf{i}.\end{aligned}$$

Things are looking rather encouraging at this point, so let's try to build an explicit isomorphism. We will build on the notation we have used thus far.

Theorem 8.87. $\mathbb{R}[x]/\left\langle x^2+1\right\rangle \cong \mathbb{C}$.

Proof. Let I, $\mathcal{C}$, and $\mathbf{i}$ hold the same meanings as above. To use the isomorphism theorem, start by defining a map $f : \mathbb{R}[x] \to \mathbb{C}$ in the following way:

1. Let $p \in \mathbb{R}[x]$.
2. Let $bx+a$ be the remainder of division of p by x^2+1.
3. Let $f(p) = a+bi$.

We claim that f is a homomorphism. To see why, let $p, q \in \mathbb{R}[x]$. Let $bx+a$ and $cx+d$ be the remainders of division of p and q (respectively) by x^2+1. It is pretty clear that

$$f(p)+f(q) = (a+bi)+(c+di) = (a+c)+(b+d)i$$

and

$$f(p)f(q) = (a+bi)(c+di) = (ac-bd)+(ad+bc)i.$$

It's a little harder to show that these equal $f(p+q)$ and $f(pq)$, respectively. To see that they do, consider $f(p+q)$ first. Since the remainders of division were $bx+a$ and $dx+c$, we know that there exist $h_p, h_q \in \mathbb{R}[x]$ such that

$$\begin{aligned}p+q &= \left[h_p\left(x^2+1\right)+(bx+a)\right]+\left[h_q\left(x^2+1\right)+(dx+c)\right]\\ &= \left(h_p+h_q\right)\left(x^2+1\right)+[(b+d)x+(a+c)].\end{aligned}$$

We see that the remainder of division of $p+q$ by x^2+1 is $(b+d)x+(a+c)$, so by definition,

$$f(p+q) = (a+c)+(b+d)i = f(p)+f(q).$$

As for multiplication,

$$\begin{aligned}pq &= \left[h_p\left(x^2+1\right)+(bx+a)\right]\left[h_q\left(x^2+1\right)+(dx+c)\right]\\ &= h'\left(x^2+1\right)+\left[bdx^2+(bc+ad)x+ac\right],\end{aligned}$$

where

$$h' = h_p h_q \left(x^2+1\right) + h_p \left(dx+c\right) + h_q \left(bx+a\right).$$

(We don't really care much for the details of h', but there they are.) We can rewrite this again as

$$\begin{aligned} pq &= \left(h'+bd\right)\left(x^2+1\right) + \left[bdx^2 + (bc+ad)x + ac\right] - bd\left(x^2+1\right) \\ &= h''\left(x^2+1\right) + \left[(bc+ad)x + (ac-bd)\right], \end{aligned}$$

where $h'' = h' + bd$. (Again, we don't really care much for the details of h''.) We have now written pq in a form that allows us to apply the definition of f:

$$f(pq) = (ac-bd) + (bc+ad)i = f(p)f(q).$$

We have shown that f is indeed a ring homomorphism. It is *not* an isomorphism, since $f\left(x^2\right) = i = f\left(2x^2+1\right)$ (and a bunch more, besides). However, did you notice something? We also have

$$\ker f = \left\langle x^2+1\right\rangle = I,$$

since the remainder of division of p by x^2+1 is zero if and only if p is a multiple of x^2+1, and hence, in its principal ideal! By the isomorphism theorem, then, there exists an isomorphism from $\mathcal{C} = \mathbb{R}[x]/\ker f$ to $\mathbb{C}$, as claimed by the theorem. □

Exercises.

Exercise 8.88. Construct an explicit isomorphism to show that the Boolean ring R of Exercise 7.17 is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Exercise 8.89. Prove Theorem 8.79.

Exercise 8.90. Prove Proposition 8.85.

Exercise 8.91. Let R and S be rings, and $f : R \to S$ a homomorphism of rings.

(a) Show that $\ker f$ is an ideal of R.

(b) Show that the function $g : R \to R/\ker f$ by $g(x) = x + \ker f$ is a homomorphism of rings.

Exercise 8.92. Let R be a ring and $a \in R$. The **evaluation map with respect to** a is $\varphi_a : R[x] \to R$ by $\varphi_a(f) = f(a)$; that is, φ_a maps a polynomial to its value at a.

(a) Suppose $R = \mathbb{Q}[x]$ and $a = 2/3$, find $\varphi_a\left(2x^2-1\right)$ and $\varphi_a(3x-2)$.

(b) Show that the evaluation map is a ring homomorphism.

(c) Recall from Example 8.80 that $\mathbb{Q}$ is isomorphic to the quotient ring $\mathbb{Q}[x]/\langle ax+b\rangle$ where $ax+b \in \mathbb{Q}[x]$ is non-zero. Use Theorem 8.82 to show this a different way.

Exercise 8.93. Use Theorem 8.82 to show that $\mathbb{Q}[x]/\left\langle x^2\right\rangle$ is isomorphic to

$$\left\{\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}\right\} \subset \mathbb{Q}^{2\times 2}.$$

Note: $\mathbb{Q}^{2\times 2}$ is not commutative! However, $\mathbb{Q}[x]/\langle x^2\rangle$ is commutative, so this isomorphism shows that the given subset of $\mathbb{Q}^{2\times 2}$ is, too. (It might not be the most efficient way of showing that, of course.)

Exercise 8.94. In this exercise we show that $\mathbb{R}$ is not isomorphic to $\mathbb{Q}$ as rings, and $\mathbb{C}$ is not isomorphic to $\mathbb{R}$ as rings.

(a) Assume to the contrary that there exists an isomorphism f from $\mathbb{R}$ to $\mathbb{Q}$.
   (i) Use the properties of an onto homomorphism to find $f(1)$.
   (ii) Use the properties of a homomorphism with the result of (i) to find $f(2)$.
   (iii) Use the properties of a homomorphism to obtain a contradiction with $f\left(\sqrt{2}\right)$.

(b) Find a similar proof that $\mathbb{C}$ and $\mathbb{R}$ are not isomorphic.

Exercise 8.95. Show that if R is an integral domain, then $\operatorname{Frac}(R)$ is isomorphic to the intersection of all fields containing R as a subring.

Exercise 8.96. Recall from Exercise 8.15 that if A is an ideal of R and B is an ideal of S, then $A\times B$ is an ideal of $R\times S$. Show that $(R\times S)/(A\times B)\cong(R/A)\times(S/B)$.

Part III

Applications

Chapter 9:
Roots of univariate polynomials

In this chapter, we take very preliminary steps into a field called *Galois theory*. This brings the two major threads of these notes to a culmination: here, group theory and ring theory come together to produce very powerful tools to analyze the behavior of roots of polynomial equations. The development of this subject was originally to find a way 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. What you should notice about this formula is that it requires arithmetic operations (addition, subtraction, multiplication, division) *and* one radical. This very elegant approach, called *solving by radicals,* can be extended to cubic and quartic polynomials, as Renaissance mathematicians discovered. Quintic polynomials turned out to be more difficult — because, as Ruffini discovered and Abel proved, it is *impossible* to solve every quintic polynomial by radicals; you have to introduce new techniques for that. We will explore only the theory developed by Galois which explains this failure.

To keep things simple, we will keep our polynomials over a ground field $\mathbb{F}$. In addition, we will assume for any $n\in\mathbb{Z}$ and any $a\in\mathbb{F}$, $na=0$ implies $n=0$ or $a=0$. This assumption rules out all the clockwork fields, since $a\in\mathbb{Z}_p$ implies that $pa=0$. This is related to the notion of *characteristic*, which we take up in more detail in Section 10.3.

9.1: Radical extensions of a field

Section 8.5 ended with an example; using the polynomial x^2+1 over $\mathbb{R}$, we built a new field, "$\mathbb{C}$", over which the polynomial x^2+1 factored. In addition, we showed that we could view "$\mathbb{C}$" as an extension of $\mathbb{R}$; that is, we can find a subfield of "$\mathbb{C}$" that is isomorphic to $\mathbb{R}$.

Can we generalize this phenomenon to arbitrary polynomials? No! Consider, for example, x^2-1; in this case, $\mathbb{F}/\left\langle x^2-1\right\rangle$ is not a field! After all,

$$\left[(x+1)+\left\langle x^2-1\right\rangle\right]\cdot\left[(x-1)+\left\langle x^2-1\right\rangle\right]=\left(x^2-1\right)+\left\langle x^2-1\right\rangle=\left\langle x^2-1\right\rangle,$$

which violates the zero-product property.

Extending a field by a root

The problem in the example above is that x^2-1 factors into two polynomials, both of smaller degree. To create extensions that have roots, we cannot do this.

Definition 9.1. Let $f\in\mathbb{F}[x]$ be nonzero, and suppose that we can find $p,q\in\mathbb{F}[x]$ such that $f=pq$. We say that f is **irreducible** if one of $\deg p$ or $\deg q$ is zero.

It turns out that irreducible polynomials can be characterized in terms of ideals.

Proposition 9.2. For any $f \in \mathbb{F}[x]$, the following are equivalent.
(A) f is irreducible;
(B) $\langle f \rangle$ is maximal;
(C) $\mathbb{F}[x] / \langle f \rangle$ is a field.

Proof. The equivalence between (B) and (C) is a special case of Theorem 8.63, so we focus on showing the equivalence between (A) and (B).

Let $f \in \mathbb{F}[x]$ and A an ideal of $\mathbb{F}[x]$ such that $\langle f \rangle \subseteq A \subseteq \mathbb{F}[x]$. Recall from Exercise 8.36 that $\mathbb{F}[x]$ is a principal ideal domain; thus, we can find $g \in \mathbb{F}[x]$ such that $A = \langle g \rangle$. By substitution, $\langle f \rangle \subseteq \langle g \rangle \subseteq \mathbb{F}$. By Exercise 8.17(b), $g \mid f$. Choose $q \in \mathbb{F}[x]$ such that $f = qg$.

Now, f is irreducible if and only if $\deg q = 0$ or $\deg g = 0$. In the first case, g is a constant multiple of f, so $\langle f \rangle = \langle g \rangle$. In the second case, g is constant, so $g \in \mathbb{F}$, so $1 \in \langle g \rangle$, and by Exercise 8.20, $\langle g \rangle = \mathbb{F}$.

In other words, f is irreducible if and only if $\langle f \rangle$ is maximal. □

Suppose f is an irreducible polynomial over a field $\mathbb{F}$. Since f is irreducible, we know that $\langle f \rangle$ is maximal, so $\mathbb{F} / \langle f \rangle$ is a field. Call this new field $\mathbb{E}$, and let $\alpha = x + \langle f \rangle$. Coset arithmetic shows that

$$f(\alpha) = f(x) + \langle f \rangle = \langle f \rangle = 0_{\mathbb{E}}.$$

So, just as $x + \langle x^2 + 1 \rangle$ was a root of $x^2 + 1$ in "$\mathbb{C}$", so is $\alpha = x + \langle f \rangle$ a root of f in $\mathbb{E}$. We have just proved the following:

Theorem 9.3. If $f \in \mathbb{F}[x]$ is irreducible, then $x + \langle f \rangle$ is a root of f in the field $\mathbb{F}[x] / \langle f \rangle$.

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; in fact, we will define it as the intersection of all fields that contain both $\mathbb{F}$ and α. We will write $\mathbb{F}(\alpha)$ for this field, and in the time-honored tradition of abusing notation, we will act as if $\mathbb{F}(\alpha) = \mathbb{E}$, even though $\mathbb{E}$ was defined above as something else. This interesting property gives rise to a new idea.

Definition 9.4. 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** $\mathbb{F}$.

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. To see why, let $f, g \in \mathbb{F}[x]$ be two polynomials with a common root $\alpha \notin \mathbb{F}$. Recall that $\mathbb{F}[x]$ is a Euclidean domain, and compute a gcd p of f and g. By Bezout's Identity, we can find $h_1, h_2 \in \mathbb{F}[x]$ such that $p = h_1 f + h_2 g$. By substitution, $p(\alpha) = 0$, so $p \neq 1$. On the other hand, if $\deg p < \deg f$, then f is not irreducible, which would be a contradiction. We conclude that $\deg p = \deg f$. A similar argument shows that $\deg p = \deg g$, so the degree of an algebraic extension is well-defined by an irreducible polynomial that produces that root.

Example 9.5. Let $f = x^5 - 2x^3 - 3x^2 + 6$. This factors over $\mathbb{Q}$ as $\left(x^2-2\right)\left(x^3-3\right)$. 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 $\mathbb{Q}[x] / \left\langle x^2-2\right\rangle$, 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:

$$\begin{aligned}
x^2 + x\sqrt[3]{3} + \sqrt[3]{9} &= 0 \\
x &= \frac{-\sqrt[3]{3} \pm \sqrt{\left(\sqrt[3]{3}\right)^2 - 4\sqrt[3]{9}}}{2} \\
&= \frac{-\sqrt[3]{3} \pm \sqrt{-3\sqrt[3]{9}}}{2} \\
&= \frac{-\sqrt[3]{3} \pm i\sqrt{3}\sqrt[3]{3}}{2} \\
&= -\sqrt[3]{3}\left(\frac{1}{2} \pm i\frac{\sqrt{3}}{2}\right).
\end{aligned}$$

We just encountered cube roots of unity!

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

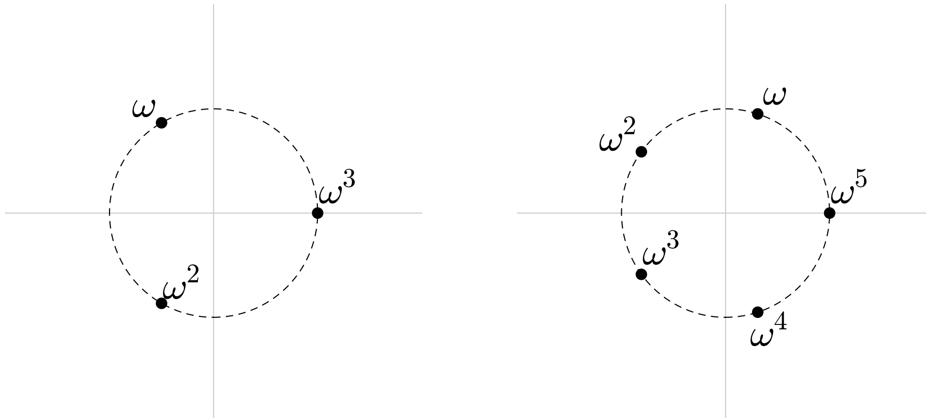


Figure 9.1. The cube and fifth roots of unity on the complex plane

degree of $\mathbb{Q}\left(\sqrt{2}, \sqrt[3]{3}\right)$ over $\mathbb{Q}$ as 6.

Definition 9.6. 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 $\left[\mathbb{E}_i : \mathbb{E}_{i-1}\right]$; we define the **degree** of $\mathbb{E}_m$ over $\mathbb{F}$ as

$$\left[\mathbb{E}_m : \mathbb{E}_{m-1}\right]\left[\mathbb{E}_{m-1} : \mathbb{E}_{m-2}\right]\cdots\left[\mathbb{E}_2 : \mathbb{E}_1\right]\left[\mathbb{E}_1 : \mathbb{E}_0\right].$$

Complex roots

The previous example shows that the roots we need are related to the roots of unity. We will see that, in fact, we can obtain radical roots by adjoining both a "principal" root, and a sensible "root of unity." This relates closely to some geometry, so let's take a brief glance at the **complex plane**, which we already met in Section 2.4. The plane maps any number $a + bi \in \mathbb{C}$ to the point $(a, b) \in \mathbb{R}^2$. It can be shown that this map is an isomorphism between the additive groups $\mathbb{C}$ and $\mathbb{R}^2$, which shouldn't startle you too much if you think about it long enough. (Hint: $\mathbb{C} \cong \mathbb{R}[x] / \left\langle x^2 + 1\right\rangle$.)

Graphing the roots of unity gives us a spectacular pattern, which applies to radical extensions, as well. Recall from Theorems 2.73 and 2.75 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 2.74 we see further that

$$\omega^m = \cos\left(\frac{2\pi m}{n}\right) + i\sin\left(\frac{2\pi m}{n}\right).$$

Figure 9.1 shows some of these roots on the complex plane.

The amazing thing is how this pattern extends beyond the roots of unity, and why.

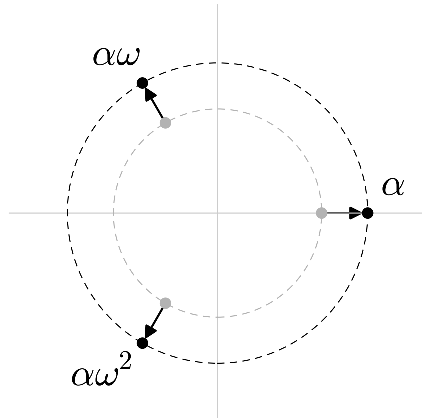


Figure 9.2. The roots of x^3-3, obtained using one root and the cube roots of unity.

Theorem 9.7. 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 roots,

$$\alpha,\ \alpha\omega,\ \alpha\omega^2,\ \ldots,\ \alpha\omega^{n-1}.$$

□

Example 9.8. 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 9.2 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 9.7 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 that 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.

Exercises

Exercise 9.9. Show that the function $f:\mathbb{C}\longrightarrow\mathbb{R}^2$ by $f(a+bi)=(a,b)$ is an isomorphism of additive groups.

Exercise 9.10. Find the smallest extension field of $\mathbb{Q}$ where $f(x)=x^7-2x^4-x^3+2$ factors completely.

Exercise 9.11. In the discussion of whether x^3-3 factored over $\mathbb{Q}\left(\sqrt{2}\right)$, we stated that we "knew" that $\sqrt{2}$, $\sqrt[3]{3}$, and $\sqrt{-2/3}$ were not rational. Explain how we know this; in other words, prove that they are irrational.

Exercise 9.12. Let α be a root of an irreducible polynomial $f\in\mathbb{F}[x]$, with $\deg f\geq 2$. Show that the set $\mathbb{F}(\alpha)$, as defined in this section, is non-empty, satisfies the properties of a field, and satisfies $\mathbb{F}\subsetneq\mathbb{F}(\alpha)\subseteq\mathbb{E}$, where $\mathbb{E}$ is any field that contains both $\mathbb{F}$ and α.

Exercise 9.13. Suppose that $\alpha^n\in\mathbb{Q}$, $\alpha^i\notin\mathbb{Q}$ for $1\leq i<n$, and ω is a primitive n-th root of unity. Show that $\mathbb{Q}\left(\alpha,\omega\alpha,\ldots,\omega^{n-1}\alpha\right)=\mathbb{Q}(\omega,\alpha)$.

Exercise 9.14. Let f be an irreducible polynomial over a field $\mathbb{F}$, of degree n. Let α be a root of f. Show that $\mathbb{F}(\alpha)$ is a vector space of dimension n over $\mathbb{F}$, with basis elements $\left\{1,\alpha,\ldots,\alpha^{n-1}\right\}$.

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 9.24.) 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

$$f(x) \quad = \quad x^n + \left(\sum_{i=1}^{n} \alpha_i\right)x^{n-1} + \left(\sum_{i<j} \alpha_i\alpha_j\right)x^{n-2} + \cdots + \alpha_1\alpha_2\cdots\alpha_n.$$

In particular, every coefficient is a sum of terms, and if we were to change any term in this sum by permuting the roots, we end up with another term in the same sum.

Example 9.15. 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 $\alpha_2\alpha_1$, which also appears in that coefficient, although in a different order.

This gives rise to a special class of polynomial.

Definition 9.16. 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 9.17. 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 9.18. 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.

Theorem 9.19. 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 $n = \deg f$.

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 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 9.19 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 9.20. 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 9.14, $\mathbb{E}$ 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 $w,z \in \mathbb{E}$, $\varphi(w)\varphi(z) = \varphi(wz)$.

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, we are interested in the automorphisms that fix $\mathbb{Q}$, so we will assume that $\varphi(a) = a$ 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 we already found that $\varphi(q) = q$, and φ is an automorphism, which means it has to be one-to-one: 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.

Exercises.

Exercise 9.21. The polynomial $f(x) = x^4 - 7x^2 + 10$ factors over $\mathbb{Q}$ as $\left(x^2-2\right)\left(x^2-5\right)$, 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.

Exercise 9.22. Show that the conjugation map $\varphi(a + bi) = a - bi$ is a ring homomorphism in $\mathbb{C}$.

Exercise 9.23. 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) $\mathbb{F} = \mathbb{Q}$; $\mathbb{E} = \mathbb{Q}(2)$

(c) $\mathbb{F} = \mathbb{Q}$; $\mathbb{E} = \mathbb{Q}\left(\sqrt{2}, \sqrt{5}\right)$

(d) $\mathbb{F} = \mathbb{Q}$; $\mathbb{E} = \mathbb{Q}\left(i, \sqrt{2}\right)$

Exercise 9.24. Let $f \in \mathbb{F}$ of degree n. Use the Factor Theorem to show that f has at most n roots in $\mathbb{F}$.

9.3: Galois groups

In the previous section, we found a narrow avenue into studying the solution of a polynomial via the structure of the coefficients of the terms. In particular, we observed that permuting the roots does not change the coefficients, which 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, *the order in which we add the roots should not matter*. This becomes a condition on σ, as well.

Example 9.25. 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 9.14, 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 $\mathbb{E}$.

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 any such function that fixes $\mathbb{Q}$ must satisfy

$$\varphi\left(\sqrt{2} \cdot \sqrt{2}\right) = \varphi(2) = 2,$$

but the homomorphism property implies instead that

$$\varphi\left(\sqrt{2} \cdot \sqrt{2}\right) = \varphi\left(\sqrt{2}\right)\varphi\left(\sqrt{2}\right) = \sqrt[3]{3} \cdot \sqrt[3]{3} \neq 2,$$

a contradiction.

This example illustrates an important property.

Theorem 9.26. 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 there is no isomorphism over $\mathbb{E}$ that fixes $\mathbb{F}$ and exchanges α and β.

You will generalize this result in Exercise 9.35.

Proof. By way of contradiction, suppose that 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 $\alpha^m \neq \beta^m$. □

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 9.27. 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 $\text{Gal}(\mathbb{E}/\mathbb{F})$.

Our first observation of the Galois set is that it's actually a Galois *group*.

Theorem 9.28. 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 $\mathbb{E}$ certainly acts as the identity over $\mathbb{F}$, 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.

We have shown that G satisfies the definition of a group. By definition, $\text{Gal}(\mathbb{E}/\mathbb{F}) = G$ is a group. □

Our second observation is that the Galois group of a radical extension has a wonderfully simple form.

Theorem 9.29. 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 $\text{Gal}(\mathbb{F}(\alpha)/\mathbb{F}) \cong \mathbb{Z}_p$.

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 9.33. (See also Exercise 9.13.)

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 9.26, any $\varphi \in G$ satisfies $\varphi(\alpha) = \beta$ only if β is another root of α^p. By Theorem 9.7, $\beta = \omega^m \alpha$ 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 $0 \le i < p-1$ and define, for any $m \in \mathbb{N}^+$, $\varphi\left(\sum_{j=0}^m b_j \alpha^j\right) = \sum_{j=0}^m b_j \left(\omega^i \alpha\right)^j$. 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 $\varphi(a + 0 \cdot \alpha) = a + 0\left(\omega^i \alpha\right) = \alpha$. To see why φ is a homomorphism, observe that for any $a, b, c, d \in \mathbb{F}$, we have

$$\begin{aligned}\varphi\left(\left(\sum_{j=0}^{p-1} b_j \alpha^j\right)\left(\sum_{k=0}^{p-1} c_k \alpha^k\right)\right) &= \varphi\left(\sum_{j=0}^{2p-2}\left[\sum_{k+\ell=j}(b_k c_\ell)\right]\alpha^j\right)\\ &= \sum_{j=0}^{2p-2}\left[\sum_{k+\ell=j}(b_k c_\ell)\right]\left(\omega^i\alpha\right)^j\end{aligned}$$

and

$$\begin{aligned}\varphi\left(\sum_{j=0}^{p-1} b_j \alpha^j\right)\varphi\left(\sum_{k=0}^{p-1} c_k \alpha^k\right) &= \left[\sum_{j=0}^{p-1} b_j\left(\omega^i\alpha\right)^j\right]\left[\sum_{k=0}^{p-1} c_k\left(\omega^i\alpha\right)^k\right]\\ &= \sum_{j=0}^{p-1}\sum_{k=0}^{p-1} b_j c_k\left(\omega^i\alpha\right)^{j+k}\\ &= \sum_{j=0}^{2p-2}\left[\sum_{k+\ell=j}(b_k c_\ell)\right]\left(\omega^i\alpha\right)^j.\end{aligned}$$

(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? It is not hard to see that the definitino 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 $0 \le j, k < p$. 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 (Exercise 3.61), $\mathbb{Z}_p$. Hence, $\operatorname{Gal}(\mathbb{E}/\mathbb{F}) \cong \mathbb{Z}_p$. □

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, $\operatorname{Gal}(\mathbb{E}/\mathbb{Q})$. 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**, 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 $\operatorname{Gal}(\mathbb{F}_3/\mathbb{F}_i)$ 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.

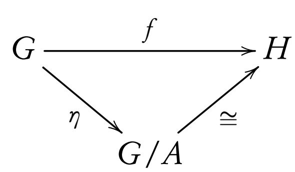
Theorem 9.30. If $\mathbb{E} \supseteq \mathbb{F}(\alpha) \supseteq \mathbb{F}$ is a tower of extensions of $\mathbb{F}$, where $\mathbb{F}(\alpha)$ is a radical extension of degree p, p is irreducible, and

- α is a primitive p-th root of unity, or
- $\mathbb{F}$ contains a primitive p-th root of unity,

then

- $\operatorname{Gal}(\mathbb{E}/\mathbb{F}(\alpha)) \triangleleft \operatorname{Gal}(\mathbb{E}/\mathbb{F})$, and
- the corresponding quotient group is abelian.

Proof of Theorem 9.30. The basic idea is to use the Isomorphism Theorem for Groups (Theorem 4.46 on page 156): for a homomorphism f from G onto H, with $A = \ker f$, we have the relationships in the following diagram.



(Theorem 4.14 guarantees that $\ker f$ is a normal subgroup of G.) Suppose we set $H = \operatorname{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$. Depending on whether α is a primitive p-th root of unity or $\mathbb{F}$ contains a primitive p-th root of unity, $H = \operatorname{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$ is isomorphic either to $\mathbb{Z}_p$ (Theorem 9.29) or $\mathbb{Z}_{p-1}$ (Exercise 9.33). If we can find a way to set $G = \operatorname{Gal}(\mathbb{E}/\mathbb{F})$ and map G *onto* H in such a way that $\ker f =$

$\text{Gal}(\mathbb{E}/\mathbb{F}(\alpha))$, we would first have

$$\text{Gal}(\mathbb{E}/\mathbb{F}(\alpha)) = \ker f \quad \triangleleft \quad G = \text{Gal}(\mathbb{E}/\mathbb{F}),$$

and since H would be abelian, we would have

$$\text{Gal}(\mathbb{E}/\mathbb{F}) / \text{Gal}(\mathbb{E}/\mathbb{F}(\alpha)) \quad \cong \quad \text{Gal}(\mathbb{F}(\alpha)/\mathbb{F}) = H,$$

so that the quotient group is abelian, as desired.

To this end, define $f : \text{Gal}(\mathbb{E}/\mathbb{F}) \to \text{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$ by *restriction to* $\mathbb{F}(\alpha)$, which means that f assigns each $\sigma \in \text{Gal}(\mathbb{E}/\mathbb{F})$ to $\tau \in \text{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$ 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 $\sigma, \widehat{\sigma} \in \text{Gal}(\mathbb{E}/\mathbb{F})$, $\tau = f(\sigma)$, $\widehat{\tau} = f(\widehat{\sigma})$, 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 $\tau \in \text{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$, and define

$$\sigma(x) = \begin{cases} \tau(x), & x \in \mathbb{F}(\alpha); \\ 0, & \text{otherwise.} \end{cases}$$

Exercise 9.36 shows that $\sigma \in \text{Gal}(\mathbb{E}/\mathbb{F})$, 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 $\text{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$. 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, $\sigma \in \text{Gal}(\mathbb{E}/\mathbb{F}(\alpha))$! Since σ was arbitrary, $\ker f = \text{Gal}(\mathbb{E}/\mathbb{F}(\alpha))$.

We have shown that f is a function from $\text{Gal}(\mathbb{E}/\mathbb{F})$ onto $\text{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$ whose kernel is $\text{Gal}(\mathbb{E}/\mathbb{F}(\alpha))$. As explained in the first paragraph of the proof, this completes the theorem. □

We rely on the following corollary in subsequent sections.

Corollary 9.31. 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

$$\begin{aligned} \{e\} = G_0 &\triangleleft G_1 \\ G_1 &\triangleleft G_2 \\ &\vdots \\ G_{n-1} &\triangleleft G_n = \mathrm{Gal}(\mathbb{F}(\alpha_1,\ldots,\alpha_n)/\mathbb{F}) \end{aligned}$$

and the corresponding quotient rings 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

$$\begin{gathered} \mathrm{Gal}(\mathbb{F}(\alpha_1,\ldots,\alpha_n)/\mathbb{F})/\mathrm{Gal}(\mathbb{F}(\alpha_1,\ldots,\alpha_n)/\mathbb{F}(\alpha_1)) \\ \mathrm{Gal}(\mathbb{F}(\alpha_1,\ldots,\alpha_n)/\mathbb{F}(\alpha_1))/\mathrm{Gal}(\mathbb{F}(\alpha_1,\ldots,\alpha_n)/\mathbb{F}(\alpha_1,\alpha_2)) \\ \vdots \\ \mathrm{Gal}\big(\mathbb{F}(\alpha_1,\ldots,\alpha_n)/\mathbb{F}\big(\alpha_1,\ldots,\alpha_{n-1}\big)\big)/\mathrm{Gal}(\mathbb{F}(\alpha_1,\ldots,\alpha_n)/\mathbb{F}(\alpha_1,\ldots\alpha_n)). \end{gathered}$$

From these groups, the following assignments satisfy the claim:

$$\begin{aligned} G_0 &= \mathrm{Gal}(\mathbb{F}(\alpha_1,\ldots,\alpha_n)/\mathbb{F}(\alpha_1,\ldots,\alpha_n)) \\ G_1 &= \mathrm{Gal}\big(\mathbb{F}(\alpha_1,\ldots,\alpha_n)/\mathbb{F}\big(\alpha_1,\ldots,\alpha_{n-1}\big)\big) \\ &\vdots \\ G_{n-1} &= \mathrm{Gal}(\mathbb{F}(\alpha_1,\ldots,\alpha_n)/\mathbb{F}(\alpha_1)). \end{aligned}$$

□

Exercises.

Exercise 9.32. Show that every non-trivial element of Ω_p is a primitive root of unity when p is irreducible.

Exercise 9.33. Suppose that ω is a primitive p-th root of unity, where p is irreducible. Show that if $\omega \notin \mathbb{F}$, then $\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F}) \cong \mathbb{Z}_{p-1}$.

Exercise 9.34. Theorem 9.29 classifies the Galois group for radical extensions of *one* radical of *prime* order; that is, $\alpha^p \in \mathbb{Q}$. Why does that also take care of radical extensions of *one* radical of *composite* order? In other words, how can we deal with $\sqrt[6]{a}$ using the same theorem?

Exercise 9.35. 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.

Exercise 9.36. Suppose $\mathbb{E} \supsetneq \mathbb{K} \supsetneq \mathbb{F}$ is a tower of fields. Let $\tau \in \operatorname{Gal}(\mathbb{K}/\mathbb{F})$. Define $\sigma : \mathbb{E} \to \mathbb{E}$ by

$$\begin{cases} \sigma(x) = \tau(x), & x \in \mathbb{K}; \\ \sigma(x) = 0, & \text{otherwise.} \end{cases}$$

Show that $\sigma \in \operatorname{Gal}(\mathbb{E}/\mathbb{F})$.

9.4: "Solvable" groups

We found in the previous section that the Galois groups corresponding to each step of a tower of radical extensions had a special property. We study this property in some detail in this section, and start by generalizing the property to arbitrary groups.

Definition 9.37. If a group G contains subgroups G_0, G_1, …, G_n such that
- $G_0 = \{e\}$;
- $G_n = G$;
- $G_{i-1} \triangleleft G_i$; and
- G_i / G_{i-1} is abelian,

then G is a **solvable group**. The chain of subgroups G_0, …, G_n is called a **normal series**.

Example 9.38. Any finite abelian group G is solvable: let $G_0 = \{e\}$ and $G_1 = G$. Subgroups of an abelian group are always normal, so $G_0 \triangleleft G_1$. In addition, $X, Y \in G_1/G_0$ implies that $X = x\{e\}$ and $Y = y\{e\}$ for some $x, y \in G_1 = G$. Since G is abelian,

$$XY = (xy)\{e\} = (yx)\{e\} = YX.$$

Example 9.39. The group D_3 is solvable. To see this, let $n = 2$ and $G_1 = \langle \rho \rangle$:
- By Exercise 3.76 on page 131, $\{e\} \triangleleft G_1$. To see that $G_1/\{e\}$ is abelian, note that for any $X, Y \in G_1/\{e\}$, we can write $X = x\{e\}$ and $Y = y\{e\}$ 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

$$\begin{aligned} XY &= (xy)\{e\} = \rho^{a+b}\{e\} \\ &= \rho^{b+a}\{e\} = (yx)\{e\} = YX. \end{aligned}$$

- By Exercise 3.88 on page 134 and the fact that $|G_1| = 3$ and $|G_2| = 6$, we know that $G_1 \triangleleft G_2$. The same exercise tells us that G_2/G_1 is abelian.

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 9.40. Every quotient group of a solvable group is solvable.

Proof. Let G be a group and $A \triangleleft G$. 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 $A_{i-1} \triangleleft A_i$ for each $i = 1, \ldots, n$. 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 $G_{i-1} \triangleleft G_i$, 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 $A_{i-1} \triangleleft A_i$.

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 $T = tA$. Let $U \in A_{i-1}$; we can likewise write $U = uA$ for some $u \in G_{i-1}$. Since G_i/G_{i-1} is abelian, $(st)G_{i-1} = (ts)G_{i-1}$; thus, $(st)u = (ts)v$ for some $v \in G_{i-1}$. By definition, $vA \in A_{i-1}$. By substitution and the definition of coset arithmetic, we have

$$\begin{aligned} XY &= (ST)A_{i-1} = ((st)A)A_{i-1} \\ &= [(st)A](uA) = ((st)u)A \\ &= ((ts)v)A = [(ts)A](vA) \\ &= ((ts)A)A_{i-1} = (TS)A_{i-1} \\ &= YX. \end{aligned}$$

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. □

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 Exercises 2.40 on page 83 and 3.89 on page 135, and a few properties of commutator subgroups.

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 Exercise 3.22.

Notation 9.43. We wrote G' as $[G, G]$ in Exercise 3.89.

Lemma 9.44. For any group G, $G' \triangleleft G$. In addition, G/G' is abelian.

Proof. You showed that $G' \triangleleft G$ in Exercise 3.89 on page 135. 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 9.45. If $H \subseteq G$, then $H' \subseteq G'$.

Proof. You do it! See Exercise 9.49. □

Notation 9.46. 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 9.47. A group is solvable if and only if $G^{(n)} = \{e\}$ for some $n \in \mathbb{N}$.

Proof. ($\Longrightarrow$) 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 $G^{(n-0)} \subseteq G_0 = \{e\}$, 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 9.45, $\left(G^{(n-i)}\right)' \subseteq G_i'$. Hence

$$G^{(n-i+1)} \subseteq G_i'. \tag{29}$$

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

$$\begin{aligned}(xy)G_{i-1} &= \left(xG_{i-1}\right)\left(yG_{i-1}\right)\\ &= \left(yG_{i-1}\right)\left(xG_{i-1}\right) = (yx)G_{i-1}.\end{aligned}$$

By Lemma 3.37 on page 121, $(yx)^{-1}(xy) \in G_{i-1}$; 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 (29), this implies that $G^{(n-(i-1))} = G^{(n-i+1)} \subseteq G_{i-1}$.

We have shown the claim; thus, $G^{(n)} = \{e\}$ for some $n \in \mathbb{N}$.

($\Longleftarrow$) Suppose that $G^{(n)} = \{e\}$ for some $n \in \mathbb{N}$. By Lemma 9.44, the subgroups form a normal series; that is,

$$\{e\} = G^{(n)} \triangleleft G^{(n-1)} \triangleleft \cdots \triangleleft G^{(0)} = G$$

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. □

We can now prove Theorem 9.41.

Proof of Theorem 9.41. Let $H < G$. Assume G is solvable; by Lemma 9.47, $G^{(n)} = \{e\}$. By Lemma 9.45, $H^{(i)} \subseteq G^{(i)}$ for all $n \in \mathbb{N}$, so $H^{(n)} \subseteq \{e\}$. By the definition of a group, $H^{(n)} \supseteq \{e\}$, so the two are equal. By the same lemma, H is solvable. □

Exercises.

Exercise 9.48. Explain why Ω_n is solvable for any $n \in \mathbb{N}^+$.

Exercise 9.49. Show that if $H \subseteq G$, then $H' \subseteq G'$.

Exercise 9.50. Show that Q_8 is solvable.

Exercise 9.51. In the textbook *God Created the Integers...* the theoretical physicist Stephen Hawking reprints 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 $n \leq 4$ 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.[25] Recall from Section 5.1 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?

9.5: The Theorem of Abel and Ruffini

In this section, we use the characterization of solution by radicals in Theorem 9.30 and Definition 9.37 to show that some polynomials cannot be solved by radicals. The basic idea is that S_5

[25]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.

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, $\operatorname{ord} g \mid |G|$. You might wonder whether the reverse is true; that is, if m is an integer that divides $|G|$, then can we always find $g \in G$ such that $\operatorname{ord}(g) = m$? Of course not; if so, then we could find $g \in G$ such that $\operatorname{ord} g = |G|$, and every group would be cyclic. Nevertheless, some interesting properties *do* hold, and one of them is critical to the result we want.

Theorem 9.52 (Cauchy's Theorem)**.** Let $p \in \mathbb{N}^+$ is irreducible, and let G be a group. If $p \mid |G|$, then we can find $g \in G$ such that $\operatorname{ord}(g) = p$.

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 true "vacuously".

Inductive hypothesis: Let $n \in \mathbb{N}^+$, and suppose that all abelian groups whose size is at most n, and where $p \mid n$, contain at least one element whose order is p.

Inductive step: Let $g \in G \setminus \{e\}$. If $p \mid \operatorname{ord}(g)$, then let $d = \operatorname{ord}(g)/p$, and the group

$$\left\langle g^d \right\rangle = \left\{ g^d, \left(g^d\right)^2, \ldots, \left(g^d\right)^{p-1}, \left(g^d\right)^p = g^{\operatorname{ord}(g)} = e \right\}$$

will have order p. Otherwise, $p \nmid \operatorname{ord}(g)$. Let $Q = G/\langle g \rangle$; the size of Q is, by definition, the number of cosets of $\langle g \rangle$, which is $|G|/\operatorname{ord}(g)$. Since $p \mid |G|$ but $p \nmid \langle g \rangle$, we see 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 $g \neq e$, 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 $m = \operatorname{ord}(x)$ in G. Since X has order p, we know that $x^m = e$, so

$$X^m = x^m \langle g \rangle = e \langle g \rangle = \langle g \rangle .$$

By Exercise 2.67, $p \mid m$. 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 Exercise 9.62 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 9.61 that this is a subgroup of G. Since $p \nmid Z(G)$, $Z(G) \neq G$, so we can find $x \in G \backslash Z(G)$, so that $|C_x| < |G|$. If $p \mid C_x$, then the inductive hypothesis applies.

Assume, therefore, that p does not divide the size of any centralizers. 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; 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 $y \in x^G$. That means 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 $x^G \cap y^G = \emptyset$. From the partition we can see that $\sum_{x \in \mathcal{P}} \left|x^G\right| = |G|$.

On the other hand, we claim that each x^G satisfies $\left|x^G\right| = |G/C_x|$. 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 well-defined, 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 $\left(h^{-1}g\right)x\left(g^{-1}h\right) = x$, or $\left(h^{-1}g\right)x\left(h^{-1}g\right)^{-1} = x$, so $h^{-1}g \in C_x$. The Lemma on coset equality then gives us $hC_x = gC_x$, as needed; φ is, indeed, well-defined. 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 $\sum_{x \in \mathcal{P}} \left|x^G\right| = |G|$. 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 equation as $|G| = |Z(G)| + \sum_{x \in \mathcal{P} \backslash Z(G)} \left|x^G\right|$. We have also seen that $\left|x^G\right| = |G/C_x|$ for all

$x \in G$, so by substitution, $|G| = |Z(G)| + \sum_{x \in \mathcal{P} \setminus Z(G)} |G/C_x|$. Rewrite this as

$$|G| - \sum_{x \in \mathcal{P} \setminus Z(G)} |G/C_x| = |Z(G)|. \tag{30}$$

Recall that if $p \nmid |C_x|$ for each $x \in \mathcal{P} \setminus Z(G)$, 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 30. 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 $\operatorname{ord}(g) = p$. □

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 9.54. If α and β are roots of an irreducible polynomial $f \in \mathbb{F}[x]$, then there exists a unique isomorphism $\sigma : \mathbb{F}(\alpha) \rightarrow \mathbb{F}(\beta)$ with $\sigma(\alpha) = \beta$.

Proof. Let $m = \deg f$. Let $\sigma : \mathbb{F}(\alpha) \rightarrow \mathbb{F}(\beta)$ by $\sigma\left(\sum_{j=0}^{m-1} a_j \alpha^j\right) = a_j \beta^j$. 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:

$$\sigma\left(\sum_{j=0}^{m-1} a_j \alpha^j \cdot \sum_{j=0}^{m-1} b_j \alpha^j\right) = \sigma\left(\sum_{j=0}^{2m-2} \left[\sum_{k+\ell=j} \left(a_k b_\ell\right)\right] \alpha^j\right) = \sum_{j=0}^{2m-2} \left(\sum_{k+\ell=j} a_k b_\ell\right) \beta^j,$$

while

$$\sigma\left(\sum_{j=0}^{m-1} a_j \alpha^j\right) \cdot \sigma\left(\sum_{j=0}^{m-1} b_j \alpha^j\right) = \sum_{j=0}^{m-1} a_j \beta^j \cdot \sum_{j=0}^{m-1} b_j \beta^j = \sum_{j=0}^{2m-2}\left(\sum_{k+\ell=j} a_k b_\ell\right)\beta^j,$$

where the js in the last equality do not have the same meaning in the left and right expressions.

To show that σ is unique, consider how an isomorphism can map roots. Let $\tau : \mathbb{F}(\alpha) \to \mathbb{F}(\beta)$ be any isomorphism that fixes $\mathbb{F}$. By Exercise 9.35, $\tau(\alpha)$ 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 9.55. A_5 is not solvable.

Proof. Use conjugates to show that any non-trivial normal subgroup contains all the three-cycles, which generate A_5.

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 $\sigma \in H \setminus \{(1)\}$ and $\tau \in A_5$. Since H is normal, $\tau\sigma\tau^{-1} \in H$. Consider the possible simplifications.

- 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 Exercise 9.64. □

Corollary 9.56. 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. □

Lemma 9.57 (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 $p^2 \mid a_0$. Hence, 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. □

Theorem 9.58. There exists a quintic polynomial over $\mathbb{Q}$ that is not solvable by radicals.

Abel and (arguably) Ruffini showed that there was no *formula* that would solve a quintic polynomial by radicals; they used many of these ideas. We will show instead that there is a quintic *polynomial* that cannot be solved by any formula by radicals. So, while they showed something general, we are showing something quite specific.

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 Exercise 9.59), so f has 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 $\beta_2 = c-di$. Consider the automorphisms of the final extension field.

- 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. The "homomorphic" construction of φ guarantees that it is a homomorphism; it remains to show that φ is one-to-one and onto. Let $z, w \in \mathbb{C}$ and write $z = a+bi$, $w = \widehat{a}+\widehat{b}i$. Assume that $\varphi(z) = \varphi(w)$; applying the homomorphism property, we see that

$$\varphi(a)+\varphi(b)\varphi(i) = \varphi(\widehat{a})+\varphi\left(\widehat{b}\right)\varphi(i)$$
$$a-bi = \widehat{a}-\widehat{b}i.$$

  Complex numbers are equal if and only if their real parts and their imaginary parts are equal; thus, $a = \widehat{a}$ and $b = \widehat{b}$. Substitution shows that $z = w$, so φ is one-to-one. As for the onto property, $\varphi(a-bi) = z$; since z was chosen arbitrarily, φ is onto. Thus, φ is an automorphism.
- 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, we 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 $\left\{1, x+I, \ldots, x^4+I\right\}$, 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 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 9.54, 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 $\mathrm{Gal}(\mathbb{E}/\mathbb{Q}) \cong S_5$ (Exercise 9.60). We know from Corollary 9.56 that S_5 is not solvable. Apply the contrapositive of Theorem 9.30 to see that f cannot be solved by radicals. □

Exercises

Exercise 9.59. Use the product rule of Calculus to show that $x = a$ is a repeated root of a polynomial f if and only $f'(a) = 0$.

Exercise 9.60. Suppose a subgroup H of S_5 has a two-cycle and a five-cycle. Show that $H = S_5$.

Exercise 9.61. Show that the centralizer C_x of an element x in a group G is a subgroup of G.

Exercise 9.62. Show that the center $Z(G)$ of a group G is a subgroup of G.

Exercise 9.63. Show that S_4 is solvable, and explain why this means any degree-four polynomial can be solved by radicals.

Exercise 9.64. Show that if a subgroup H of A_5 contains all the three-cycles, then in fact $H = A_5$.

9.6: The Fundamental Theorem of Algebra

Carl Friedrich Gauß proved the Fundamental Theorem of Algebra in his doctoral thesis.

Theorem 9.65 (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 calculus and 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 [FR97]. 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.

Background from Calculus

The first result we need should be well-known to every first-semester calculus student.

Theorem 9.66 (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 9.67. 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 9.68. 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

We need two final, 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 9.69. 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 9.70. 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 where repeated addition can give 0, but as noted at the beginning of the chapter, we assume that this is not the case.)

Theorem 9.71. 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 $[\mathbb{E}:\mathbb{C}] \leq n$, 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 $\varphi \in \text{Gal}(\mathbb{E}/\mathbb{C})$. By definition, $\varphi(\alpha)$ is another root of q. We showed above that extensions of $\mathbb{C}$ are separable, so φ has a choice of m roots, where $m = \deg q$. The *only* choice of a target for φ is another root of q, and $\mathbb{C}(\alpha)$. Hence, $|\text{Gal}(\mathbb{C}(\alpha)/\mathbb{C})| = \deg q = [\mathbb{C}(\alpha) : \mathbb{C}]$. Apply the inductive hypothesis to $[\mathbb{E} : \mathbb{C}(\alpha)]$ to obtain the rest. □

We now turn to the First Sylow Theorem. We can view this as a generalization of Cauchy's Theorem.

Theorem 9.72 (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 of order smaller than $|G|$, 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 30,

$$|G| - \sum_{x \in \mathcal{P} \setminus Z(G)} |G/C_x| = |Z(G)|.$$

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 $\mu\left(y^{-1}\right) = \mu(y)^{-1} \in B$, so now closure and properties of homomorphisms guarantee that $\mu\left(xy^{-1}\right) = \mu(x)\,\mu(y)^{-1} \in B$.

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, 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 $\mathbb{E} = \mathbb{C}$.

Note that $\mathbb{E}$ is a finite extension of $\mathbb{C}$, which is a finite extension of $\mathbb{R}$. Hence, $\mathbb{E}$ is also a finite extension of $\mathbb{R}$. If $\mathbb{E}$ is an odd-degree extension of $\mathbb{R}$, then we can 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 $\mathbb{C} = \mathbb{E}$.

Suppose, therefore, that the degree of $\mathbb{E}$ over $\mathbb{R}$ is $2^m q$, where $m, q \in \mathbb{N}^+$ and $2 \nmid q$. Let $G = \operatorname{Gal}(\mathbb{E}/\mathbb{R})$ 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 $|G| = 2^m$.

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 $r = a^2 + b^2$. In the exercises, you will show 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, $\mathbb{E} = \mathbb{C}$.

Exercises

Exercise 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: Factorization

In this chapter we begin a turn toward applications of ring theory. In particular, here we will build up some basic algorithms for factoring polynomials. To do this, we will study more precisely the rings that factor, then delve into the algorithms themselves.

Remark 10.1. In this chapter, every ring is an integral domain, unless otherwise specified.

10.1: The link between factoring and ideals

We start with two important problems for factorization: the link between factoring and ideals, and the distinction between irreducible and prime elements of a ring.

As for the latter, we mentioned in Chapter 6 that although irreducible integers are prime and vice-versa, the same would not hold true later. Here we want to explore the question,

When is a prime element of a ring irreducible, and vice-versa?

Before answering that question, we should first define what are meant by the two terms. In fact, their definitions are identical to the definitions in Chapter 6. Compare the definitions below to Definitions 6.28 and 6.31.

Definition 10.2. Let R be a commutative ring with unity, and $a, b, c \in R \setminus \{0\}$. We say that

- a is a **unit** if a has a multiplicative inverse;
- a and b are **associates** if $a = bc$ and c is a unit;
- a is **irreducible** if a is not a unit and for every factorization $a = bc$, one of b or c is a unit; and
- a is **prime** if a is not a unit and whenever $a \mid bc$, we can conclude that $a \mid b$ or $a \mid c$.

Example 10.3. Consider the ring $\mathbb{Q}[x]$.

- The only units are the rational numbers, since no polynomial has a multiplicative inverse.
- $4x^2 + 6$ and $6x^2 + 9$ are associates, since $4x^2 + 6 = \frac{2}{3}(6x^2 + 9)$, and $\frac{2}{3}$ is a unit. Notice that they are *not* associates in $\mathbb{Z}[x]$, however.
- $x + q$ is irreducible for every $q \in \mathbb{Q}$. $x^2 + q$ is also irreducible for every $q \in \mathbb{Q}$ such that $q > 0$.

The link between divisibility and principal ideals that you studied in Exercise 8.17(b) implies that we can rewrite Definition 10.2 in terms of ideals.

Theorem 10.4. 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 $\langle a \rangle = \langle b \rangle$.

(C) In a principal ideal domain, a is irreducible if and only if $\langle a \rangle$ is maximal.

(D) In a principal ideal domain, a is prime if and only if $\langle a \rangle$ is prime.

Proof. We show (A), (B), and (C), and leave (D) to the exercises.

(A) This is a straightforward chain: a is a unit if and only if there exists $b \in R$ such that $ab = 1_R$ if and only if $1_R \in \langle a \rangle$ if and only if $R = \langle a \rangle$ (Exercise 8.20 and 8.20).

(B) Assume that a and b are associates. Let $c \in R \setminus \{0\}$ be a unit such that $a = bc$. By definition, $a \in \langle b \rangle$. Since any arbitrary $x \in \langle a \rangle$ 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.

(C) 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) $\langle a \rangle = \langle b \rangle = B$. Otherwise, b is a unit, and by part (A) $B = \langle b \rangle = R$. 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 8.27 and part (B) of this lemma, $\langle a \rangle \subsetneq \langle b \rangle \subsetneq R$. In other words, $\langle a \rangle$ is not maximal. By the contrapositive, then, if $\langle a \rangle$ is maximal, then a is irreducible. □

Remark 10.5. In the proof, we *do* need R to be an integral domain 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.

We did *not* need the assumption that R be a principal ideal domain to show that if $\langle a \rangle$ is maximal, then a is irreducible. So in fact this remains true even when R is not a principal ideal domain.

On the other hand, if R is not a principal ideal domain, then it can happen that a is irreducible, but $\langle a \rangle$ is not maximal. Returning to the example $\mathbb{C}[x, y]$ that we exploited in Theorem 8.62 on page 270, x is irreducible, but $\langle x \rangle \subsetneq \langle x, y \rangle \subsetneq \mathbb{C}[x, y]$.

In a similar way, the proof you develop of part (D) should show that if $\langle a \rangle$ is prime, then a is prime even if R is not a principal ideal domain. The converse, however, might not be true. In any case, we have the following result.

Theorem 10.6. 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.

It is now easy to answer part of the question that we posed at the beginning of the section.

Corollary 10.7. In a principal ideal domain, if an element p is irreducible, then it is prime.

Proof. You do it! See Exercise 10.12. □

The converse is true even if we are not in a principal ideal domain.

Theorem 10.8. If R is an integral domain and $p \in R$ is prime, then p is irreducible.

Proof. Let R be a ring with unity, and $p \in R$. Assume that p is prime. Suppose that there exist $a, b \in R$ such that p factors as $p = ab$. Since $p \cdot 1 = ab$, the definition of prime implies that $p \mid a$ or $p \mid b$. Without loss of generality, there exists $q \in R$ such that $pq = a$. By substition, $p = ab = (pq)b$. Since we are in an integral domain, it follows that $1_R = qb$; that is, b is a unit.

We took an arbitrary prime p that factored, and found that one of its factors is a unit. By definition, then, p is irreducible. □

To resolve the question, we must still decide whether:

1. an irreducible element is prime even when the ring is not a principal ideal domain; or
2. a prime element is irreducible even when the ring is not an integral domain.

The answer to both question is, "only sometimes". We can actually get there with a more sophisticated structure, but we don't have the information yet.

Example 10.9. Let

$$\mathbb{Z}\left[\sqrt{-5}\right] = \left\{a + b\sqrt{-5} : a, b \in \mathbb{Z}\right\}.$$

Past results show that this is a ring; we leave the precise identification of those results to an exercise. However, it is not a principal ideal domain. Rather than show this directly, consider the fact that

$$6 = 2 \cdot 3 = \left(1 + \sqrt{-5}\right)\left(1 - \sqrt{-5}\right).$$

We claim that 2 is irreducible. By way of contradiction, suppose that it is not; then we can find $x, y \in \mathbb{Z}\left[\sqrt{-5}\right]$ such that $xy = 2$ and neither x nor y is a unit. It cannot be that $x, y \in \mathbb{Z}$, since then 2 would not be irreducible in $\mathbb{Z}$ — but it is! So at least one of x or y has the form $a + b\sqrt{-5}$ with a or b nonzero. In fact, *both* of them must have the form $a + b\sqrt{-5}$ with a **and** b nonzero; otherwise, we have the contradiction $2 = a + b\sqrt{-5}$ for some $a, b \in \mathbb{Z}$.

Can it be the case, then, that $2 = \left(a + b\sqrt{-5}\right)\left(c + d\sqrt{-5}\right)$? Expanding the product, we have

$$2 = (ac - 5bd) + (ad + bc)\sqrt{-5}.$$

This implies the system of equations

$$\begin{aligned} ac - 5db &= 2, \text{ and} \\ ad + bc &= 0. \end{aligned}$$

Since $a \neq 0$, we can rewrite the latter equation as

$$d = -\frac{bc}{a},$$

and substitute into the first equation to obtain

$$2 = ac - 5b\left(-\frac{bc}{a}\right) = -\frac{a^2c}{a} - \frac{5b^2c}{a} = -\frac{c}{a}\left(a^2 + 5b^2\right).$$

Since 2 and $a^2 + 5b^2$ are both positive, we must have c/a negative. In that case, ac is negative, as well. Since $ac - 5bd = 2$, one of b or d must be negative, but not both. That is, b and d have different signs — but this contradicts the equation $d = -bc/a$. Finding no way out of contradiction, we conclude that 2 is irreducible.

However, 2 cannot be prime in $\mathbb{Z}\sqrt{-5}$, since 2 divides the product $(1+\sqrt{-5})(1-\sqrt{-5})$, but neither of its factors. We know from Corollary 10.7 that irreducibles are prime in a principal ideal domain; hence, $\mathbb{Z}[\sqrt{-5}]$ must not be a principal ideal domain.

Example 10.10. Consider the ring $\mathbb{Z}_6$. It is not hard to verify that 2 is a prime element of $\mathbb{Z}_6$; we discussed this at the beginning of Section 8.4. However, 2 is not irreducible, since $2 = 20 = 4 \cdot 5$, neither of which is a unit. This should not surprise us, since $\mathbb{Z}_6$ is not an integral domain.

We have now answered the 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 (as many authors do) to,

A prime element is always irreducible, but an irreducible element is not always prime.

Exercises.

Exercise 10.11. Prove part (D) of Theorem 10.4.

Exercise 10.12. Prove Corollary 10.7.

Exercise 10.13. Prove that $\mathbb{Z}[\sqrt{-5}]$ is a ring.

Exercise 10.14. 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 $i \geq n$.

Exercise 10.15. Show that in a principal ideal domain R, a greatest common divisor d of $a, b \in R$ always exists, and:

(a) $\langle d \rangle = \langle a, b \rangle$; and

(b) there exist $r, s \in R$ such that $d = ra + sb$.

10.2: Unique Factorization domains

An important fact about the integers is that every integer factors *uniquely* into a product of irreducible elements. We saw this in Chapter 6 with the Fundamental Theorem of Arithmetic (Theorem 6.34). This is not true in every ring. For example, consider $\mathbb{Z}[-\sqrt{5}]$ from Exercise 7.18; here $6 = 2 \cdot 3$, but $6 = (1+\sqrt{-5})(1-\sqrt{-5})$. In this ring, 2, 3, $1+\sqrt{-5}$, and $1-\sqrt{-5}$

are all irreducible, so 6 factors two different ways as a product of irreducibles. We are interested in unique factorization, so we will start with a definition:

Definition 10.16. An integral domain is a **unique factorization domain** if every $r \in R$ factors into irreducibles $r = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_n^{\alpha_n}$, and if this factorization is unique up to order and associates.

The Fundamental Theorem of Arithmetic tells us that $\mathbb{Z}$ is a unique factorization domain. What are some others?

Example 10.17. $\mathbb{Z}[x]$ is a unique factorization domain. To see this takes two major steps. Let $f \in \mathbb{Z}[x]$. If the terms of f have a common divisor, we can factor that out easily; for example, $2x^2 + 4x = 2x(x+2)$. So we may assume, without loss of generality, that the terms of f have no common factor. If f is not irreducible, then we claim it must factor as two polynomials of smaller degree. Otherwise, f would factor as ag where $\deg a = 0$, which implies $a \in \mathbb{Z}$, which implies that a is a common factor of the terms of f, contradicting the hypothesis. Since the degrees of the factors of f are integers, and they decrease each time we factor a polynomial further, the well-ordering property of $\mathbb{Z}$ implies that this process must eventually end with irreducibles; that is, $f = p_1 p_1 \cdots p_n$ — but $i \neq j$ does *not* imply that $p_i \neq p_j$.

Suppose that we can also factor f into irreducibles by $f = q_1 \cdots q_m$. Consider f as an element of $\mathbb{Q}[x]$, which by Exercise 8.36 is a principal ideal domain. Corollary 10.7 tells us that irreducible elements of $\mathbb{Q}[x]$ are prime. Hence p_1 divides q_j for some $j = 1, \ldots, m$. 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 $a_1 = \frac{b}{c}$ where $\gcd(b,c) = 1$; we have $p_1 = \frac{b}{c} \cdot q_1$. We can rewrite this as $cp_1 = bq_1$. Lemma 6.18 implies both that $c \mid q_1$ and that $b \mid p_1$. However, we assumed that p_1 and q_1 were irreducible. If $b \mid p_1$ (resp. $c \mid q_1$), then the greatest common divisor of the coefficients of p_1 (resp. q_1) is not 1, so p_1 (resp. q_1) would not be irreducible in $\mathbb{Z}[x]$! So $b, c = \pm 1$, which implies that $a_1 = \pm 1$. Hence p_1 and q_1 are associates in $\mathbb{Z}[x]$.

The same argument can be applied to the remaining irreducible factors. Thus, the factorization of f was unique up to order and associates.

This result generalizes to an important class of rings.

Theorem 10.18. Every principal ideal domain is a unique factorization domain.

Proof. Let R be a principal ideal domain, and $f \in R$.

First we show that f has a factorization. Suppose f is not irreducible; then there exist $p_1, p_2 \in R$ such that $f = p_1 p_2$ and f is not an associate of either. By Theorem 10.4, $\langle f \rangle \subsetneq \langle p_1 \rangle$ and $\langle f \rangle \subsetneq \langle p_2 \rangle$. If p_1 is not irreducible, then there exist $p_3, p_4 \in R$ such that $p_1 = p_3 p_4$ and p_1 is not an associate of either. Again, $\langle p_1 \rangle \subsetneq \langle p_3 \rangle$ and $\langle p_1 \rangle \subsetneq \langle p_4 \rangle$. Continuing in this fashion, we obtain an ascending chain of ideals

$$\langle f\rangle \subsetneq \langle p_1\rangle \subsetneq \langle p_3\rangle \subsetneq \cdots .$$

By Theorem 8.33, a principal ideal domain satisfies the ascending chain condition; thus, this chain must terminate eventually. It can terminate only if we reach an irreducible polynomial. This holds for each chain, so they must all terminate with irreducible polynomials. Combining the results, we obtain $f = p_1^{\alpha_1}\cdots p_m^{\alpha_m}$ where each p_i is irreducible.

Now we show the factorization is unique. Suppose that $f = p_1^{\alpha_1}\cdots p_m^{\alpha_m}$ and $f = q_1^{\beta_1}\cdots q_n^{\beta_n}$ where $m \leq n$ and the p_i and q_j are irreducible. Recall that irreducible elements are prime in a principal ideal domain (Corollary 10.7). 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 $p_1 = a_1 q_1$ for some unit $a_1 \in R$. Since we are in an integral domain, we can cancel p_1 and q_1 from $f = f$, obtaining

$$p_2^{\alpha_2}\cdots p_m^{\alpha_m} = a_1^{-1} q_1^{\beta_1-\alpha_1} q_2^{\beta_2}\cdots q_n^{\beta_n}.$$

Since p_2 is irreducible, hence prime, we can continue this process until we conclude with $1_R = a_1^{-1}\cdots a_m^{-1} q_1^{\gamma_1}\cdots q_n^{\gamma_n}$. By definition, irreducible elements are not units, so $\gamma_1,\ldots,\gamma_n$ are all zero. 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. □

Corollary 10.19. Every Euclidean domain is a unique factorization domain.

Proof. This is a consequence of Theorem 10.18 and Theorem 8.30. □

The converse is false; see Example 7.60. However, the definition of a greatest common divisor that we introduced with Euclidean domains certainly generalizes to unique factorization domains.

We can likewise extend a result from a previous section.

Theorem 10.20. In a unique factorization domain, irreducible elements are prime.

Proof. You do it! See Exercise 10.24. □

Corollary 10.21. In a unique factorization domain:
- an element is irreducible iff it is prime; and
- an ideal is maximal iff it is prime.

In addition, we can say the following:

Theorem 10.22. 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 $d = p_1^{\alpha_1}\cdots p_m^{\alpha_m}$ be an irreducible factorization of d, and $\widehat{d} = q_1^{\beta_1}\cdots q_n^{\beta_n}$ 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 10.20, 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. □

Exercises.

Exercise 10.23. Use $\mathbb{Z}[x]$ to show that even if R is a unique factorization domain but not a principal ideal domain, then we cannot write always find $r, s \in R$ such that $\gcd(a, b) = ra + sb$ for every $a, b \in R$.

Exercise 10.24. Prove Theorem 10.20.

Exercise 10.25. 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 $a_1, a_2, \ldots$ clearly.

Exercise 10.26. Theorem 10.22 says that gcds 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 (Exercise 7.71), it is a unique factorization domain, and gcds are unique up to associates (Theorem 10.22). 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.

Exercise 10.27. Generalize the argument of Example 10.17 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 $R[x]$.

10.3: Finite Fields I

Most of the fields you have studied in the past have been infinite, such as $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$. Not all fields are infinite, however; we saw in Exercises 7.29 and 7.31 that $\mathbb{Z}_n$ is a field when n is irreducible. This tells us that

- not all fields are infinite, like $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$; and
- the finite fields that we have worked with so far are of the form $\mathbb{Z}_p$, where p is irreducible.

Do finite fields of other forms exist? In fact, they do, and we can state their structure with precision! We will find that any finite field has p^n elements where $p, n \in \mathbb{N}$ and p is irreducible.

The characteristic of a ring

Before we proceed, we will need the following definition.

Definition 10.28. Let R be a ring with unity. If there exists a smallest positive integer c such that $c1_R = 0_R$, then, R has **characteristic** c. Otherwise, R has **characteristic zero**.

Example 10.29. The rings $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ have characteristic zero. Why? Let r be any nonzero element of one of these sets. The zero product property of the integers tells us that $c \cdot 1 = 0$ if and only if $c = 0$. Since 0 is not a *positive* integer, we conclude that these rings have characteristic 0.

Example 10.30. The ring $\mathbb{Z}_8$ has characteristic 8. Why? Certainly $8 \cdot [1] = [8] = [0]$, and no smaller integer n gives us $n \cdot [1] = [0]$.

Example 10.31. Let $p \in \mathbb{Z}$ be irreducible. By Exercise 7.32, $\mathbb{Z}_p$ is a field. The same argument we used in Example 10.30 shows that the characteristic of $\mathbb{Z}_p$ is p. In fact, the characteristic of $\mathbb{Z}_n$ is n for *any* $n \in \mathbb{N}^+$.

In the last two examples, the characteristic of a finite ring turned out to be the number of elements in the ring. This is not always the case.

Example 10.32. 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:

$$(a, b) + (c, d) = (a + c, b + d)$$
$$(a, b) \cdot (c, d) = (ac, bd).$$

It is not hard to show that R is a ring; we leave it to Exercise 10.37. 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.

In case you are wondering why we have dedicated this much time to characteristic, which is about *rings*, whereas this section is supposedly about fields, don't forget that a field is a commutative ring with a multiplicative identity and a little more. Thus we *have* been talking about fields, but we have also been talking about other kinds of rings as well. This is one of the nice things about abstraction: later, when we talk about other kinds of rings that are not fields but are commutative and have a multiplicative identity, we can still use these ideas.

Example

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 10.33. Our notation for a finite field with n elements is $\mathbb{F}_n$.

Example 10.34. We will build finite fields with four and sixteen elements. In the exercises, you will use the same technique to build fields of nine and twenty-seven elements.

Case 1. $\mathbb{F}_4$

Start with the polynomial ring $\mathbb{Z}_2[x]$. We claim that $f(x) = x^2 + x + 1$ does not factor in $\mathbb{Z}_2[x]$. If it did, it would have to factor as a product of linear polynomials; that is,

$$f(x) = (x+a)(x+b)$$

where $a, b \in \mathbb{Z}_2$. This implies that a is a root of f (remember that in $\mathbb{Z}_2$, $a = -a$), but f has no zeroes:

$$f(0) = 0^2 + 0 + 1 = 1 \text{ and}$$
$$f(1) = 1^2 + 1 + 1 = 1.$$

Thus f does not factor. By Exercise 8.72, $I = \langle f \rangle$ is a maximal ideal in $R = \mathbb{Z}_2[x]$, and by Theorem 8.63, R/I is a field.

How many elements does this field have? Let $X \in R/I$; choose a representation $g + I$ of X where $g \in R$. I claim that we can assume that $\deg g < 2$. Why? If $\deg g \geq 2$, then we can subtract multiples of f; since $f + I$ is the zero element of R/I, this does not affect X. After all, absorption tells us that $hf \in I$ for each $h \in \mathbb{Z}_2[x]$, so $hf + I = I$, and thus

$$(g - hf) + I = (g + I) + (-hf + I) = (g + I) + I = g + I.$$

Given that $\deg g < 2$, there must be two terms in g: x^1 and x^0. Each of these terms can have one of two coefficients: 0 or 1. This gives us $2 \times 2 = 4$ distinct possibilities for the representation of X; thus there are 4 elements of R/I. We can write them as

$$I, \quad 1 + I, \quad x + I, \quad x + 1 + I.$$

Case 2. $\mathbb{F}_{16}$

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 zero of f. As for the second, suppose that $f = \left(x^2 + ax + b\right)\left(x^2 + cx + d\right)$, where $a, b, c, d \in \mathbb{Z}_2$. Let's consider this possibility: If

$$\begin{aligned} x^4 + x + 1 = x^4 + (a + c)x^3 + (ac + b + d)x^2 \\ + (ad + bc)x + db, \end{aligned}$$

then since equal polynomials must have the same coefficients for like terms, we have the system of linear equations

$$\begin{aligned} a + c &= 0 && (31) \\ ac + b + d &= 0 \\ ad + bc &= 1 \\ bd &= 1. && (32) \end{aligned}$$

Recall that $b, d \in \mathbb{Z}_2$, so (32) means that $b = d = 1$; after all, the only other choice would be 0, which would contradict $bd = 1$. The system now simplifies t

$$a + c = 0 \tag{33}$$
$$ac + 1 + 1 = ac = 0 \tag{34}$$
$$a(1+1) = a = 1 \tag{35}$$

Equation (35) states flatly that $a = 1$. Equation 34 and substitution tell us that $c = 0$. However, this contradicts equation 31!

We have confirmed that f does not factor. By Exercise 8.72, $I = \langle f \rangle$ is a maximal ideal in $R = \mathbb{Z}_2[x]$, and by Theorem 8.63, R/I is a field.

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 $\deg g \geq 4$ then we can subtract multiples of f; since $f + I$ is the zero element of R/I, this does not affect X. Since $\deg g < 4$, there are four terms in g: x^3, x^2, x^1, and x^0. Each of these terms can have one of two coefficients: $[0]$ or $[1]$. This gives us $2^4 = 16$ distinct possibilities for the representation of X; thus there are 16 elements of R/I. We can write them as

$$\begin{array}{rr}
I, & 1+I, \\
x^2+I & x^2+1+I, \\
x^3+I, & x^3+1+I, \\
x^3+x^2+I, & x^3+x^2+1+I, \\
\\
x+I, & x+1+I, \\
x^2+x+I, & x^2+x+1+I, \\
x^3+x+I, & x^3+x+1+I, \\
x^3+x^2+x+I, & x^3+x^2+x+1+I.
\end{array}$$

You may have noticed that in each case we ended up with p^n elements where $p = 2$. Since we started with $\mathbb{Z}_p$, you might wonder if the generalization of this to arbitrary finite fields starts with $\mathbb{Z}_p[x]$, finds a polynomial that does not factor in that ring, then builds the quotient ring. Yes and no. One does start with $\mathbb{Z}_p$, and if we could find an irreducible polynomial of degree n over $\mathbb{Z}_p$, then we would be finished. Unfortunately, finding an irreducible polynomial of $\mathbb{Z}_p$ is not easy.

Instead, one considers $f(x) = x^{p^n} - x$; from Euler's Theorem (6.51) we deduce (via induction) that $f(a) = 0$ for all $a \in \mathbb{Z}_p$. One can then use *field extensions* from *Galois Theory* to construct p^n roots of f, so that f factors into linear polynomials. Extend $\mathbb{Z}_p$ by those roots; the resulting field has p^n elements. We will take that question up in Section 10.4. For the time being, we settle for the following.

Main result

Theorem 10.35. 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. The proof has three steps.[26]

First, we show that $\mathbb{F}_n$ has characteristic p, where p is an irreducible integer. Let p be the characteristic of $\mathbb{F}_n$. Since $\mathbb{F}_n$ is finite, $p \neq 0$. Suppose that $p = ab$ for some $a, b \in \mathbb{N}^+$. Now

$$0_{\mathbb{F}_n} = p \cdot 1_{\mathbb{F}_n} = (ab) \cdot 1_{\mathbb{F}_n} = \left(a \cdot 1_{\mathbb{F}_n}\right)\left(b \cdot 1_{\mathbb{F}_n}\right).$$

Recall that a field is an integral domain; by definition, it has no zero divisors. Hence $a \cdot 1_{\mathbb{F}_n} = 0_{\mathbb{F}_n}$ or $b \cdot 1_{\mathbb{F}_n} = 0_{\mathbb{F}_n}$; without loss of generality, $a \cdot 1_{\mathbb{F}_n} = 0_{\mathbb{F}_n}$. By definition, p is the smallest positive integer c such that $c \cdot 1_{\mathbb{F}_n} = 0_{\mathbb{F}_n}$; thus $p \leq a$. However, a divides p, so $a \leq p$. This implies that $a = p$ and $b = 1$; since $p = ab$ was an arbitrary factorization of p, p is irreducible.

Second, we claim that for any irreducible $q \in \mathbb{N}$ that divides n (the size of the field), we can find $x \in \mathbb{F}_n$ such that $q \cdot x = 0_{\mathbb{F}_n}$. Let $q \in \mathbb{N}$ such that q is irreducible and q divides $n = |\mathbb{F}_n|$. Consider the additive group of $\mathbb{F}_n$. Let

$$\mathcal{L} = \left\{ \left(a_1, a_2, \ldots, a_q\right) : a_i \in \mathbb{F}_n, \sum_{i=1}^{q} a_i = 0 \right\};$$

that is, $\mathcal{L}$ is the set of all lists of q elements of $\mathbb{F}_n$ such that the sum of those elements is the additive identity. For example,

$$q \cdot 0_{\mathbb{F}_n} = 0_{\mathbb{F}_n} + 0_{\mathbb{F}_n} + \cdots + 0_{\mathbb{F}_n} = 0_{\mathbb{F}_n},$$

so $\left(0_{\mathbb{F}_n}, 0_{\mathbb{F}_n}, \ldots, 0_{\mathbb{F}_n}\right) \in \mathcal{L}$.

Recall the group of permutations S_q. Let $\sigma \in S_q$ and $\left(a_1, a_2, \ldots, a_q\right) \in \mathcal{L}$; the commutative property implies $\sigma\left(a_1, a_2, \ldots, a_q\right) \in \mathcal{L}$. Let

- $\mathcal{M}_1$ be the subset of $\mathcal{L}$ that is invariant under S_q — that is, if $A \in \mathcal{M}_1$ and $\sigma \in S_q$, then $\sigma(A) = A$; and
- $\mathcal{M}_2$ be a subset of $\mathcal{L}$ containing *exactly one* permutation of any $A \in \mathcal{L}$ that is not invariant under S_q — that is, if $A \in \mathcal{M}_2$, then there exists $\sigma \in S_q$ such that $\sigma(A) \neq A$, but only one of A or $\sigma(A)$ is in S_q, not both.

We pause a moment in the proof, to consider an example of this.

Example 10.36. In $\mathbb{F}_6$ with $q = 3$, we know that $1 + 3 + 2 = 0$, so $A = (1, 3, 2) \in \mathcal{L}$. Certainly A is not invariant under S_3, since if $\sigma = \left(1\ \ 2\right)$, we have $\sigma(A) = (3, 1, 2) \neq A$. *This does not mean* that $A \in \mathcal{M}_2$; rather, *exactly one* of

$$(1,3,2), (3,1,2), (2,3,1), (1,2,3), (3,2,1), (2,1,3)$$

is in $\mathcal{M}_2$. Notice, therefore, that each element of $\mathcal{M}_2$ corresponds to $q! = 6$ elements of $\mathcal{L}$.

Back to the proof!

Which elements are in $\mathcal{M}_1$? Let $A \in \mathcal{L}$, and notice that if $a_i \neq a_j$, then $\sigma = \left(i\ \ j\right)$ swaps those two elements in $\sigma(A)$. So if $a_i \neq a_j$ for any i and j, then A is not invariant under S_q.

[26] Adapted from the proofs of Theorems 31.5, 42.4, and 46.1 in [AF05].

Thus, the elements of $\mathcal{M}_1$ are those tuples whose entries are identical; that is, $\left(a_1,\ldots,a_q\right)\in\mathcal{M}_1$ iff $a_1=\cdots=a_q$. In particular, $\left(0_{\mathbb{F}_n},0_{\mathbb{F}_n},\ldots,0_{\mathbb{F}_n}\right)\in\mathcal{M}_1$.

Let $|\mathcal{M}_1|=r$ and $|\mathcal{M}_2|=s$. Can we show that q divides either of these numbers? Let's look at $\mathcal{M}_2$ first. Let $A\in\mathcal{M}_2$; by definition of $\mathcal{M}_2$, A must have at least two distinct entries. How many lists $B\in\mathcal{L}$ does A represent? For one of its distinct entries, we can choose any of the q positions to rewrite it, so the number of possible B must be $q\cdot m$, where m is some integer that counts the number of positions we can choose for the remaining distinct entries. Inasmuch as A was arbitrary in $\mathcal{M}_2$, *every* element A of $\mathcal{M}_2$ represents qm_A elements of $\mathcal{L}$. Since every element of $\mathcal{L}$ is represented by only one element of $\mathcal{M}_2$, the number of elements of $\mathcal{L}$ that can be permuted into any $A\in\mathcal{M}_2$ is

$$\sum_{A\in\mathcal{M}_2} qm_A = q\sum_{A\in\mathcal{M}_2} m_A.$$

Unfortunately, this does *not* say that $q\mid s$. It does, however, imply that $q\mid r$.

Why? Each element A of $\mathcal{L}$ is either invariant under S_q or modified by some permutation of S_q. If A is invariant, it ends up in $\mathcal{M}_1$. If A is not invariant, it ends up in $\mathcal{M}_2$. So, we can count the elements of $\mathcal{L}$ in the following way:

$$|\mathcal{L}| \quad = \quad \#\text{ of els of }\mathcal{M}_1 \quad + \quad q\sum_{A\in\mathcal{M}_2} m_A.$$

We can rewrite this as

$$|\mathcal{L}| = |\mathcal{M}_1| + q\sum_{A\in\mathcal{M}_2} m_A. \tag{36}$$

How many elements are in $\mathcal{L}$ total? Any element of $\mathcal{L}$ satisfies

$$a_q = -\left(a_1+a_2+\cdots+a_{q-1}\right),$$

so we can choose any elements at all for a_1, …, a_{q-1}, while the final, qth element is determined. Since $\mathbb{F}_n$ has n elements, we have n choices for each of these, so

$$|\mathcal{L}| = |\mathbb{F}_n|^{q-1} = n^{q-1}.$$

Recall that $q\mid n$; choose $d\in\mathbb{N}$ such that $n=qd$. By substitution, $|\mathcal{L}|=(qd)^{q-1}$. Substitute into 36, and we find that

$$(qd)^{q-1} = r + q\sum_{A\in\mathcal{M}_2} m_A$$

$$q\left[d\,(qd)^{q-2} - \sum_{A\in\mathcal{M}_2} m_A\right] = r,$$

so q does in fact divide r.

Now recall that $\left(0_{\mathbb{F}_n},0_{\mathbb{F}_n},\ldots,0_{\mathbb{F}_n}\right)\in\mathcal{L}$; this tells us that $r\geq 1$. Since q is irreducible, $q\neq 1$; so $r\neq 1$. Since $r=|\mathcal{M}_1|$, it must be that $\mathcal{M}_1$ contains a non-zero element; call it X. Since $\mathcal{M}_1\subseteq\mathcal{L}$,

$$x_1+x_2+\cdots+x_q = 0.$$

Recall that elements of $\mathcal{M}_1$ are those whose entries are identical; that is, $x_1 = x_2 = \cdots = x_q$. Let's agree to write x instead. By substitution, then,

$$\underbrace{x + x + \cdots + x}_{q \text{ times}} = 0.$$

In other words,

$$q \cdot x = 0_{\mathbb{F}_n},$$

as claimed.

Third, recall that, in the first claim, we showed that the characteristic of $\mathbb{F}_n$ is an irreducible positive integer, p. We claim that for any irreducible $q \in \mathbb{Z}$ that divides n, $q = p$. To see this, let q be an irreducible integer that divides n. Recall that in the second claim, we showed that there existed some $x \in \mathbb{F}_n$ such that $q \cdot x = 0_{\mathbb{F}_n}$. Choose such an x. Since the characteristic of $\mathbb{F}_n$ is p, we also have $px = 0$. Consider the additive cyclic group $\langle x \rangle$; by Exercise 2.67 on page 101, $\operatorname{ord}(x) \mid p$, but p is irreducible, so $\operatorname{ord}(x) = 1$ or $\operatorname{ord}(x) = p$. Since $x \neq 0_{\mathbb{F}_n}$, $\operatorname{ord}(x) \neq 1$; thus $\operatorname{ord}(x) = p$. Likewise, $p \mid q$, and since both p and q are irreducible, this implies that $q = p$.

We have shown that if $q \mid n$, then $q = p$. Thus all the irreducible divisors of n are p, so n is a power of p. □

A natural question to ask is whether $\mathbb{F}_{p^n}$ exists for every irreducible p and every $n \in \mathbb{N}^+$. You might think that the answer is yes; after all, it suffices to find an polynomial of degree n that is irreducible over $\mathbb{F}_p$. However, it is not obvious that such polynomials exist for every possible p and n. That is the subject of Section 10.4.

Exercises.

Exercise 10.37. Recall $R = \mathbb{Z}_2 \times \mathbb{Z}_4$ from Example 10.32.

(a) Show that R is a ring, but not an integral domain.

(b) Show that for any two rings R_1 and R_2, $R_1 \times R_2$ is a ring with addition and multiplication defined in the natural way.

(c) Show that even if the rings R_1 and R_2 are fields, $R_1 \times R_2$ is not even an integral domain, let alone a field. In other words, we cannot construct direct products of integral domains and fields.

(d) Show that for any n rings $R_1, R_2, \ldots, R_n$, $R_1 \times R_2 \times \cdots \times R_n$ is a ring with addition and multiplication defined in the natural way. In other words, we can construct direct products of rings.

Exercise 10.38. Build the addition and multiplication tables of the field of four elements that we constructed in Example 10.34.

Exercise 10.39. Construct a field with 9 elements, and list them all.

Exercise 10.40. Construct a field with 27 elements, and list them all.

Exercise 10.41. Does every infinite field have characteristic 0?

10.4: Finite fields II

We saw in Section 10.3 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 $n \in \mathbb{N}^+$.

Let $\mathbb{F}$ be a field.

Theorem 10.42. 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 $\widehat{f} \in \mathbb{E}[x]$ such that the coefficient of x^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 $\widehat{f}$.

Proof. Denote $I = \langle f \rangle$.

(A) Let $\mathbb{E} = \mathbb{F}[x]/I$. In Exercise 8.72, you showed that if f is irreducible in $\mathbb{F}[x]$, then I is maximal in $\mathbb{F}[x]$. By Theorem 8.63, the quotient ring $\mathbb{E} = \mathbb{F}[x]/I$ is a field.

(B) To see that $\mathbb{F}$ is isomorophic to

$$\mathbb{F}' = \{a + I : a \in \mathbb{F}\} \subsetneq \mathbb{E},$$

use the function $\varphi : \mathbb{F} \to \mathbb{F}'$ by $\varphi(a) = a + I$. You will show in the exercises that φ is a ring isomorphism.

(C) Let $\alpha = x + I$. 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,

$$\widehat{f}(\alpha) = (a_0 + I) + (a_1 + I)\alpha + \cdots + (a_n + I)\alpha^n.$$

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 3.37, $f + I = I$, so $\widehat{f}(\alpha) = I$. Recall that $\mathbb{E} = \mathbb{F}[x]/I$; it follows that $\widehat{f}(\alpha) = 0_{\mathbb{E}}$. □

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 $\mathbb{F}$.

Corollary 10.43 (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 Exercise 10.27, $\mathbb{F}[x]$ 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 10.42, 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 $\mathbb{E}$. □

Example 10.44. 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 $\mathbb{E}$.

Before we proceed to the third topic of this section, we need a concept that we borrow from Calculus.

Definition 10.45. Let $f \in \mathbb{F}[x]$, and write $f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$. The **formal derivative of** f is

$$f' = a_1 + 2a_2 x + \cdots + n a_n x^{n-1}.$$

Proposition 10.46 (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 Definition 7.46 and 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

$$f=\sum_{k=0}^{m+n}\left[\left(\sum_{i+j=k}a_ib_j\right)x^k\right].$$

We can now examine the claim. By definition,

$$f'=\sum_{k=1}^{m+n}\left[k\left(\sum_{i+j=k}a_ib_j\right)x^{k-1}\right].$$

On the other hand,

$$\begin{aligned}
p'q+pq'&=\left(\sum_{i=1}^{m}ia_ix^{i-1}\right)\left(\sum_{j=0}^{n}b_jx^j\right)\\
&\quad+\left(\sum_{i=0}^{m}a_ix^i\right)\left(\sum_{j=1}^{n}jb_jx^{j-1}\right)\\
&=\sum_{k=1}^{m+n}\left[\left(\sum_{i+j=k}ia_ib_j\right)x^{k-1}\right]\\
&\quad+\sum_{k=1}^{m+n}\left[\left(\sum_{i+j=k}ja_ib_j\right)x^{k-1}\right]\\
&=\sum_{k=1}^{m+n}\left[\left(\sum_{i+j=k}(i+j)a_ib_j\right)x^{k-1}\right]\\
&=\sum_{k=1}^{m+n}\left[\left(\sum_{i+j=k}ka_ib_j\right)x^{k-1}\right]\\
&=f'.
\end{aligned}$$

□

We need one more result: a generalization of Euler's Theorem.

Lemma 10.47. 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. Euler's Theorem tells us that $a^{p-1} = 1$. Thus,

$$\begin{aligned} a^{p^n} - a &= a\left(a^{p^n-1} - 1\right) \\ &= a\left(a^{(p-1)\left(p^{n-1}+p^{n-2}+\cdots+1\right)} - 1\right) \\ &= a\left(\left(a^{(p-1)}\right)^{\left(p^{n-1}+p^{n-2}+\cdots+1\right)} - 1\right) \\ &= a\left(1^{p^{n-1}+p^{n-2}+\cdots+1} - 1\right) \\ &= 0. \end{aligned}$$

Since $a^p = a$, $a^p - a = 0$, so a is a root of $x^p - x$; applying the Factor Theorem gives us the factorization claimed. □

We can now prove the final assertion of this section.

Theorem 10.48. 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 $\mathbb{Z}_2/\left\langle x^2 + x + 1\right\rangle$ 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. Suppose to the contrary that it has at least one such factor, $x - a$. We can write

$$f = (x-a)^m \cdot g$$

for some $g \in \mathbb{E}[x]$ where $(x-a) \nmid g$. By Proposition 10.46,

$$\begin{aligned} f' &= m(x-a)^{m-1} \cdot g + (x-a)^m \cdot g' \\ &= (x-a)^{m-1} \cdot \left(mg + (x-a)g'\right). \end{aligned}$$

That is, $x - a$ divides f'.

Is $f' \neq 0$? certainly $x - a \neq 0$, so we would have to have $mg + (x-a)g' = 0$. This implies either $g' = 0$ and $m = p$, or $(x-a) \mid g$. However, neither of these can hold. On the one hand, if we let $b \in \mathbb{F}\backslash\{a\}$, then Lemma 10.47 tells us that $f(b) = 0$. By the Factor Theorem, $x - b$ is a factor of f; since it is irreducible, hence prime, it is a factor of one of $(x-a)^m$ or g. Since $a \neq b$, $x - b$ has no common factor with $x - a$, so $x - b$ must divide g. Thus, $g' \neq 0$. On the other hand, we chose m to be large enough that $(x-a) \nmid g$. Hence $f' \neq 0$.

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 Exercise 10.51(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

$$(a-b)^p=a^p+\sum_{i=1}^{p-1}(-1)^i\binom{p}{i}a^ib^{p-i}+(-1)^p\,b^p.$$

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 Exercise 10.51(b), we can rewrite $(a-b)^p$ as

$$\begin{aligned}(a-b)^p&=a^p+\sum_{i=1}^{p-1}(-1)^i\,\frac{p!}{i!\,(p-i)!}a^ib^{p-i}+(-1)^p\,b^p\\&=a^p+p\cdot\sum_{i=1}^{p-1}(-1)^i\,\frac{(p-1)!}{i!\,(p-i)!}a^ib^{p-i}+(-1)^p\,b^p\\&=a^p+0+(-1)^p\,b^p\\&=a^p+(-1)^p\,b^p.\end{aligned}$$

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 $a\in\mathbb{E}\backslash\{0\}$; 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 $\mathbb{E}\backslash\{0\}$, *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.

In a finite field, we can generalize Euler's Theorem a little further. □

Theorem 10.49 (Fermat's Little Theorem)**.** In $\mathbb{Z}_{p^d}[x]$, we have

$$x^{p^d}-x=\prod_{a\in\mathbb{Z}_{p^d}}(x-a).$$

Proof. Let $a\in\mathbb{Z}_{p^d}$. If $a=0$, it is clear that $x-a=x$ is a factor of $x^{p^d}-x$. Otherwise, a lies in the multiplicative group $\mathbb{Z}_{p^d}\backslash\{0\}$. By Lagrange's Theorem, its order divides $\left|\mathbb{Z}_{p^d}\backslash\{0\}\right|=p^d-1$, so $a^{p^d-1}=1$. Multiplying both sides by a, we have $a^{p^d}=a$, which we can rewrite as $a^{p^d}-a=0$, showing that a is a root of $x^{p^d}-x$. By the Factor Theorem, $x-a$ is a factor of $x^{p^d}-x$.

Now let $b \in \mathbb{Z}_{p^d} \setminus \{a\}$. A similar argument shows that $x - b$ is a factor of $x^{p^d} - x$. Since $b \neq a$, $x - b$ and $x - a$ can have no common factors. Thus, every element of $\mathbb{Z}_{p^d}$ corresponds to a unique factor of $x^{p^d} - x$, proving the theorem. □

Exercises.

Exercise 10.50. Show that the function φ defined in part (B) of the proof of Theorem 10.42 is an isomorphism between $\mathbb{F}$ and $\mathbb{F}'$.

Exercise 10.51. 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 $f(ab) = 0$.

10.5: Extending a ring by a root

Let R and S be rings, with $R \subseteq S$ and $\alpha \in S$. In Exercise 7.18, you showed that $R[\alpha]$ was also a ring, called a **ring extension** of R. Sometimes, this is equivalent to a polynomial ring over R, but in one important case, it is more interesting.

Example 10.52. Let $R = \mathbb{R}$, $S = \mathbb{C}$, and $\alpha = i = \sqrt{-1}$. Then $\mathbb{R}[i]$ is a ring extension of $\mathbb{C}$. Moreover, $\mathbb{R}[i]$ is not really a polynomial ring over $\mathbb{R}$, since $i^2 + 1 = 0$, but $x^2 + 1 \neq 0$ in $\mathbb{R}[x]$.

In fact, since every element of $\mathbb{R}[i]$ has the form $a + bi$ for some $a, b \in \mathbb{R}$, 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 $a + bi = a\mathbf{u} + b\mathbf{v}$.

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.

Definition 10.53. Let $f \in R[x]$. If $\operatorname{lc}(f) = 1$, we say that f is **monic**.

Notation 10.54. We write $\mathbb{F}(\alpha)$ for the smallest field containing both $\mathbb{F}$ and α.

Example 10.55. In the previous example, $\mathbb{R}[i] = \mathbb{R}(i) = \mathbb{C}$. This is not always the case, though; if $\alpha = \sqrt{2}$, then $\mathbb{R}\left[\sqrt{2}\right] \subsetneq \mathbb{R}\left(\sqrt{2}\right) \subsetneq \mathbb{C}$.

Theorem 10.56. 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 $I = \langle f \rangle$. Notice that $\mathbb{F} \subseteq \mathbb{E}$. 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 $\deg r < \deg f = d$. Hence, we may assume, without loss of generality, that any element of $\mathbb{E}$ can be written in the form $g + I$ where $g \in \mathbb{F}[x]$ and $\deg g < d$. In other words, every element of $\mathbb{E}$ has the form

$$\left(a_{d-1}x^{d-1} + \cdots + a_1 x^1 + a_0 x^0\right) + I$$

where $a_{d-1}, \ldots, a_1, a_0 \in \mathbb{F}$. Since $\mathbb{F}$ is a field, and $x^i + I$ cannot be written as a linear combination of the $x^j + I$ where $j \neq i$, we have proved that $\mathbb{E}$ is a vector space over $\mathbb{F}$ with basis

$$B = \left\{x^0 + I, x^1 + I, \ldots, x^{d-1} + I\right\}.$$

□

It turns out that the field described in the previous theorem has an important relationship to the roots of the irreducible polynomial f.

Corollary 10.57. Let f be an irreducible, monic polynomial of degree d over a field $\mathbb{F}$. Let $I = \langle f \rangle$ and $\alpha = x + I \in \mathbb{F}[x]/I$. Then $f(\alpha) = 0$; that is, α is a root of f.

Proof. Choose $a_0, \ldots, a_d$ as in Theorem 10.56. Then

$$\begin{aligned} f(\alpha) &= a_d (x+I)^d + \cdots + a_1 (x+I)^1 + a_0 (x+I)^0 \\ &= a_d \left(x^d + I\right) + \cdots + a_1 \left(x^1 + I\right) + a_0 \left(x^0 + I\right) \\ &= \left(a_d x^d + \cdots + a_1 x^1 + a_0 x^0\right) + I \\ &= f(x) + \langle f \rangle \\ &= \langle f \rangle = 0_{\mathbb{E}} \end{aligned}$$

where $\mathbb{E} = \mathbb{F}[x]/I$, as before. □

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 Corollary 10.57, $f_i(\alpha_i) = 0$, 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 (Exercise 0.37). We have thus described a way to factor irreducible polynomials.

Definition 10.58. Let f and α be as in Corollary 10.57. We say that $\deg f$ is the **degree of** α, and write $\mathbb{F}(\alpha) = \mathbb{F}[x] / \langle f \rangle$.

It is sensible to say that $\deg f = \deg \alpha$ since we showed in Theorem 10.56 that $\deg f = \dim(\mathbb{F}[x] / \langle f \rangle)$.

We need one last result.

Theorem 10.59. Suppose $\mathbb{F}$ is a field, $\mathbb{E} = \mathbb{F}(\alpha)$, and $\mathbb{D} = \mathbb{E}(\beta)$. Then $\mathbb{D}$ 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 an irreducible polynomial over $\mathbb{F}$.

Proof. By Theorem 10.56, $B_1 = \left\{\alpha^0, \ldots, \alpha^{d_1-1}\right\}$ 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 $B_3 = \left\{\alpha^{(i)}\beta^{(j)} : 0 \le i < d_1, 0 \le j < d_2\right\}$ 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 $b_0, \ldots, b_{d_2-1} \in \mathbb{E}$ 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 $\mathbb{E}$ 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 cound the number of elements of B_3, we find that there are $d_1 \cdot d_2$ elements of the basis. Hence,

$$\dim_{\mathbb{F}} \mathbb{D} = |B_3| = d_1 \cdot d_2 = \deg \alpha \cdot \deg \beta.$$

□

Exercises

Exercise 10.60. Let $\mathbb{F} = \mathbb{R}\left(\sqrt{2}\right)\left(\sqrt{3}\right)$.

(a) Find an irreducible polynomial $f \in \mathbb{R}[x]$ that factors in $\mathbb{F}$.

(b) What is $\dim_{\mathbb{R}} \mathbb{F}$?

Exercise 10.61. 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 $x^n + a = (x+b)\left(x^{n-1} - bx^{n-2} + \cdots + b^{n-1}\right)$.

10.6: Polynomial factorization in finite fields

We now turn to the question of factoring polynomials in $R[x]$. This material comes primarily from [vzGG99].

Suppose that $f \in R[x]$; factorization requires the following steps.

- **Squarefree factorization** is the process of removing multiples of factors p of f; that is, if $p^a \mid f$, then we want to work with $\frac{f}{p^{a-1}}$, for which only p is a factor.
- **Distinct degree factorization** is the process of factoring a squarefree polynomial f into polynomials $p_1, \ldots, p_m$ such that if p_i factors as $p_i = q_1 \cdots q_n$, then $\deg q_1 = \cdots \deg q_n$.
- **Equal degree factorization** is the process of factoring each distinct degree factor p_i into its equal degree factors $q_1, \ldots, q_n$.

The algorithms we develop in this chapter only work in finite fields. To factor a polynomial in $\mathbb{Z}[x]$, we will first factor over several finite fields $\mathbb{Z}_p[x]$, then use the Chinese Remainder Theorem to recover a factorization in $\mathbb{Z}[x]$. We discuss this in Section 10.7.

The goal of this section is merely to show you how the ideas studied so far combine into this problem. The algorithm we will study is not an inefficient algorithm, but more efficient ones exist.

For the rest of this section, we assume that $p \in \mathbb{N}$ is irreducible and $f \in \mathbb{Z}_p[x]$.

Distinct degree factorization.

Distinct-degree factorization can be accomplished using Fermat's Little Theorem.

Example 10.62. Suppose $p = 5$. You already know from basic algebra that

$$\begin{aligned} x^5 - x &= x\left(x^4 - 1\right) \\ &= x\left(x^2 - 1\right)\left(x^2 + 1\right) \\ &= x\left(x - 1\right)\left(x + 1\right)\left(x^2 + 1\right). \end{aligned}$$

We are working in $\mathbb{Z}_5$, so $1 = -4$. Thus $x + 1 = x - 4$, and $(x-2)(x-3) = \left(x^2 - 5x + 6\right) = \left(x^2 + 1\right)$. 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.

We can generalize this to the following.

> **Theorem 10.63.** Let $d, d' \in \mathbb{N}^+$, and $a = d^{d'}$. Then $x^{p^a} - x$ is the product of all monic irreducible polynomials in $\mathbb{F}_{p^d}[x]$ whose degree divides d'.

Proof. We will show that if $f \in \mathbb{Z}_{p^d}[x]$ is monic and irreducible of degree n, then satisfies

$$f \mid \left(x^{p^a} - x\right) \iff n \mid d'.$$

Assume first that f divides $x^{p^a} - x$. By Fermat's Little Theorem on the field $\mathbb{F}_{p^a}$, the factors of f are of the form $x - c$, where $c \in \mathbb{F}_{p^a}$. Let α be any one of the corresponding roots, and let $\mathbb{E} = \mathbb{F}(\alpha)$. Using the basis B of Theorem 10.56, we see that $|\mathbb{E}| = p^{d^n}$, since it has $|B| = n$ basis elements, and p^d choices for each coefficient of a basis element.

Now, $\mathbb{Z}_{p^a}$ 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 $p^a = \left|\mathbb{Z}_{p^a}\right| = p^{d^{nb}}$ for some $b \in \mathbb{N}^+$. Rewriting the extreme sides of that equation, we have

$$p^{d^{d'}} = p^a = p^{d^{nb}}.$$

Since $nb = d'$, we see that $n \mid d'$.

Algorithm 4. 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 $i = 0$
  repeat while $f_i \neq 1$
    Increment i
    Let h_i be the remainder of division of h_{i-1}^p by f
    Let $p_i = \gcd\left(h_i - x, f_{i-1}\right)$
    Let $f_i = \frac{f_{i-1}}{p_i}$
  Let $m = i$
return $p_1, \ldots, p_m$

Conversely, assume that $n \mid d'$. We construct $\mathbb{F}_{p^{d^n}} = \mathbb{F}[x] / \langle f \rangle$, and let α be the corresponding root $x + \langle f \rangle$ of f. Fermat's Little Theorem tells us that $\alpha^{p^{d^n}} = \alpha$. Notice that

$$p^a - 1 = \left(p^{d^n} - 1\right)\left(p^{a-d^n} + p^{a-2d^n} + \cdots + 1\right).$$

Let $r = p^{a-d^n} + p^{a-2d^n} + \cdots + 1$; we have

$$x^{p^a - 1} - 1 = \left(x^{p^{d^n} - 1} - 1\right)\left(x^{r-1} + \cdots + 1\right).$$

Rewrite this as

$$x^{p^a} - x = \left(x^{p^{d^n}} - x\right)\left(x^{r-1} + \cdots + 1\right).$$

Hence, $x^{p^{d^n}} - x$ divides $x^{p^a} - x$, so $x - \alpha$ is a root of $x^{p^a} - x$, as well. Since α was an arbitrary root of f, every root of f is a root of $x^{p^a} - x$, and unique factorization guarantees us that f divides $x^{p^a} - x$. □

Theorem 10.63 suggests an "easy" algorithm to compute the distinct degree factorization of $f \in \mathbb{Z}_p[x]$. See algorithm 4.

Theorem 10.64. algorithm 4 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 $\deg h_i^p > \deg f_i$: Theorem 10.63 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 10.65. Returning to $\mathbb{Z}_5[x]$, let's look at

$$f = x(x+3)\left(x^3+4\right).$$

Notice that *we do not know* whether this factorization is into irreducible elements. Expanded, $f = x^5 + 3x^4 + 4x^2 + 2x$. When we plug it into algorithm 4, 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.

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, $\left\langle q_j \right\rangle$ is a maximal ideal in $\mathbb{Z}_p[x]$, so $\mathbb{Z}_p[x] / \left\langle q_j \right\rangle$ is a field with p^i elements. Denote it by $\mathbb{F}$.

Lemma 10.66. Let G be the set 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 $\varphi(g) = g^e$.
(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 3.55 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 Theorem 7.45 on page 238, 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, $|\ker\varphi| \le a$. 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, Exercise 4.26 on page 147 gives us

$$p^i - 1 = |G| = |\ker\varphi|\,|\varphi(G)| \le a \cdot 2 = \frac{p^i-1}{\cancel{2}} \cdot \cancel{2} = p^i - 1.$$

The inequality is actually an equality, forcing $|\ker\varphi| = a$. □

To see how Lemma 10.66 is useful, consider a nonzero coset in $\mathbb{F}$,

$$[h] = h + \langle q_j \rangle \in \mathbb{F}.$$

Since $\gcd(h, q_j) = 1$, $h \notin \langle q_j \rangle$, 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

$$(h^a - 1) + \langle q_j \rangle = [h]^a - 1_{\mathbb{F}} = 0_{\mathbb{F}} = \langle q_j \rangle.$$

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 5. Note that there we have written f instead of p_i and d instead of i.

Algorithm 5. 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 $q = 1$
  repeat while $q = 1$
    Let $h \in \mathbb{Z}_p[x] \backslash \mathbb{Z}_p$, with $\deg h < n$
    Let $q = \gcd(h, f)$
    if $q = 1$
      Let h be the remainder from division of $h^{\frac{p^d-1}{2}}$ by f
      Let $q = \gcd(h-1, f)$
  return q

algorithm 5 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 $\varphi([h]) = 1_{\mathbb{F}}$. So to get $q \neq 1$, we have to be "lucky". If we're extraordinarily unlucky, algorithm 5 might never terminate. But this is highly unlikely, for two reasons. First, Lemma 10.66(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 5 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 of $\mathbb{N}$ implies that we will eventually conclude with an irreducible factor.

Example 10.67. Recall from Example 10.65 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 5 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 $q = \gcd\left(\left(3x^2 + 4x + 4\right) - 1, p_1\right) = x + 4$.
  · Return $x + 4$ as a factor of p_1.

We did not know this factor from the outset! In fact, $f = x(x+3)(x+4)\left(x^2 + x + 1\right)$.

As with algorithm 4, we need efficient algorithms to compute gcd's and exponents in order to perform algorithm 5. 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 1 on page 194) and fast exponentiation (Section 6.5).

Squarefree factorization

We can take two approaches to squarefree factorization. The first, which works fine for any polynomial $f \in \mathbb{C}[x]$, is to compute its derivative f', then to compute $g = \gcd\left(f, f'\right)$, and finally to factor $\frac{f}{g}$, which (as you will show in the exercises) is squarefree.

Another approach 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.

Example 10.68. In Exercise 10.72 you will try (and fail) to perform a distinct degree factorization on $f = x^5 + x^3$ using only algorithm 4. Suppose that we use algorithm 6 to factor f instead.

– Since f is monic, $b = 1$.
– With $i = 1$, distinct-degree factorization gives us $h_1 = 4x^3$, $q_1 = x^3 + x$, $f_1 = x^2$.
  · Suppose that the first factor that algorithm 5 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 5 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 $\alpha_2 = \alpha_3 = 1$.

Exercises.

Exercise 10.69. Show that $\frac{f}{g}$ is squarefree if $f \in \mathbb{C}[x]$, f' is the usual derivative from Calculus, and $g = \gcd\left(f, f'\right)$.

Exercise 10.70. Use the distinct degree factorization of Example 10.65 and the fact that $f = x(x+3)\left(x^3 + 4\right)$ to find a complete factorization of f, using only the fact that you now know three irreducible factors f (two linear, one quadratic).

Exercise 10.71. Compute the distinct degree factorization of $f = x^5 + x^4 + 2x^3 + 2x^2 + 2x + 1$ in $\mathbb{Z}_5[x]$. Explain why you know this factorization is into irreducible elements.

Exercise 10.72. Explain why you might think that algorithm 4 might not work for $f = x^5 + x^3$. Then try using the algorithm to factor f in $\mathbb{Z}_5[x]$, and explain why the result is incorrect.

Algorithm 6. Squarefree factorization in $\mathbb{Z}_p[x]$

inputs
  $f \in \mathbb{Z}_p[x]$
outputs
  An irreducible factorization $f = b\,p_1^{\alpha_1} \cdots p_m^{\alpha_m}$
do
  Let $b = \mathrm{lc}(f)$
  Let $h_0 = x$
  Let $f_0 = b^{-1} \cdot f$ — After this step, f is monic
  Let $i = j = 0$
  repeat while $f_i \neq 1$
    — 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
    repeat while $q_i \neq 1$
      Increment j
      Find a degree-i factor p_j of q_i using algorithm 5
      Let $q_i = \frac{q_i}{p_j}$
      — Divide out all copies of p_j from f_i
      Let $\alpha_j = 1$
      repeat while p_j divides f_i
        Increment α_j
        Let $f_i = \frac{f_i}{p_j}$
  Let $m = j$
  return $b, p_1, \ldots, p_m, \alpha_1, \ldots, \alpha_m$

Exercise 10.73. 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 10.71 to illustrate your argument.

10.7: Factoring integer polynomials

We conclude, at the end of this chapter, with factorization in $\mathbb{Z}[x]$. In the previous section, we showed how one could 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 coefficeints 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.

We will approach factorization by two different routes: using one big irreducible p, or several small irreducibles along with the Chinese Remainder Theorem.

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 p exist, but for our purposes it will suffice to choose p that is approximately twice the size of the maximum coefficient of $\widehat{f}$.

Example 10.74. 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 10.75. 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 certain easy enough to verify whether a potential factor of p_i 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 $\gcd\left(m_i, m_j\right) = 1$ for each $1 \leq i < j \leq n$, 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 $\gcd\left(m_i, m_j\right) = 1$; 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 10.76. 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

$$f = (x+a)(x+b)(x^2+c).$$

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 $c = [15] \in \mathbb{Z}_{1155}$.

The problem with recovering a and b is that we have to guess "correctly" which arrangement of the coefficients in the finite fields give 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 will turn 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.

Exercises.

Exercise 10.77. Factor $x^7 + 8x^6 + 5x^5 + 53x^4 - 26x^3 + 93x^2 - 96x + 18$ using each of the two approaches described here.

Chapter 11: Roots of multivariate polyomials

This chapter is about the roots of polynomial equations. However, rather than investigate the *computation* of roots, it considers the *analysis* of roots, 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.

A chemist once emailed me about a problem he was studying that involved microarrays. Microarrays measure gene expression, and he was using some data to build a system of equations of this form:

$$\begin{aligned} axy - b_1x - cy + d_1 &= 0 \\ axy - b_2x - cy + d_2 &= 0 \\ axy - b_2x - b_1y + d_3 &= 0 \end{aligned} \tag{37}$$

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. A— wanted to find values for x and y that made all the equations true.

This already is an interesting problem, and it is well-studied. In fact, A— had a fancy software program that sometimes solved the system. However, it didn't *always* solve the system, and 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 turned red, which meant that it found no solution.

The software the chemist was using relied on well-known *numerical techniques* to look for a solution. There are many reasons that numerical techniques can fail; most importantly, they can fail *even when a solution exists.*

Analyzing these systems with an *algebraic* technique, I was able to give him some glum news: the reason the software failed to find a solution is that, in fact, no *real* solution existed. Instead, the solutions were *complex*. So, the problem wasn't with the software's numerical techniques.

This chapter develops and describes the algebraic techniques that allowed me to reach this conclusion. Most of the material in these notes are relatively "old": at least a century old. Gröbner bases, however, are relatively new: they were first described in 1965 [Buc65]. We will develop Gröbner bases, and finally explain how they answer the following important questions for any system of polynomial equations

$$f_1(x_1, x_2, \ldots, x_n) = 0, \quad \cdots \quad f_m(x_1, x_2, \ldots, x_n) = 0$$

whose coefficients are in $\mathbb{R}$:

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 will refer to these as ***five natural questions about the roots of a polynomial system***. To answer them, we first review a little linear algebra, then study monomials a bit more, before concluding with a foray into Hilbert's Nullstellensatz and Gröbner bases, fundamental results and tools of commutative algebra and algebraic geometry.

Remark 11.1. From here on, all rings are polynomial rings over a field $\mathbb{F}$, *unless we say otherwise.*

11.1: Gaussian elimination

Let's look again at the system (37) described in the introduction:

$$\begin{aligned} axy - b_1x - cy + d_1 &= 0 \\ axy - b_2x - cy + d_2 &= 0 \\ axy - b_2x - b_1y + d_3 &= 0. \end{aligned}$$

It is *almost* a linear system, and you've studied linear systems in the past. In fact, you've even studied how to answer the five natural questions about the roots of a linear polynomial system. Let's review that.

A generic system of m linear equations in n variables looks like

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

where the a_{ij} and b_i are elements of a field $\mathbb{F}$. Linear algebra can be done over *any* field $\mathbb{F}$, although it is typically taught with $\mathbb{F} = \mathbb{Q}$. Since that's old hat to you, let's try some linear algebra over a finite field!

Example 11.2. A linear system with $m = 3$ and $n = 5$ 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

$$\begin{aligned} 5x_1 + x_2 + 7x_5 + 6 &= 0 \\ x_3 + 11x_4 + 2x_5 + 12 &= 0 \\ 3x_1 + 7x_2 + 8x_3 + 11 &= 0. \end{aligned}$$

In these notes, we favor the latter form.

A technique called *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.

Definition 11.3. 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 $\operatorname{lv}(g_i)$ for this variable.

The leading variable of the zero polynomial is undefined.

Since this ordering guarantees $x_1 > x_2 > \ldots > x_n$, something like a dictionary, we refer to it as the **lexicographic term ordering**.

Example 11.4. Using the example from 11.2,

$$\operatorname{lv}(5x_1 + x_2 + 7x_5 + 8) = x_1,$$
$$\operatorname{lv}(x_3 + 11x_4 + 2x_5 + 12) = x_3.$$

Definition 11.5. 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 $\operatorname{lv}(f_i) > \operatorname{lv}(f_j)$.

Example 11.6. Using the example from 11.2,the list

$$F = (5x_1 + x_2 + 7x_5 + 8, \quad x_3 + 11x_4 + 2x_5 + 12, \quad 3x_1 + 7x_2 + 8x_3 + 11)$$

is not in triangular form, since $\operatorname{lv}(f_1) = \operatorname{lv}(f_3) = x_1$, whereas we want $\operatorname{lv}(f_1) > \operatorname{lv}(f_3)$.

The list

$$G = (x_1 + 6, \quad 0, \quad x_2 + 3x_4)$$

is also not in triangular form, because g_2 is zero while $g_3 \neq 0$.

However, the list

$$H = (x_1 + 6, \quad x_2 + 3x_4, \quad 0)$$

is in triangular form, because $h_3 = 0$ and $\operatorname{lv}(h_1) > \operatorname{lv}(h_2)$.

Algorithm 7 describes one way to apply the method.

Theorem 11.7. Algorithm 7 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 claim that each 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 $\operatorname{lv}(g_j)$ in $g_{i+1}, \ldots, g_m$ is 0.

Algorithm 7. Gaussian elimination

1: **inputs**
2: $F = (f_1, f_2, \ldots, f_m)$, a list of linear polynomials in n variables, with coefficients from a field $\mathbb{F}$.
3: **outputs**
4: $G = (g_1, g_2, \ldots, g_m)$, a list of linear polynomials in n variables, in triangular form, whose roots are precisely the roots of F.
5: **do**
6: Let $G := F$
7: **for** $i = 1, 2, \ldots, m-1$
8: Rearrange $g_i, g_{i+1}, \ldots, g_m$ so that for each $k < \ell$, $g_\ell = 0$, or $\mathrm{lv}(g_k) \geq \mathrm{lv}(g_\ell)$
9: **if** $g_i \neq 0$
10: Denote the coefficient of $\mathrm{lv}(g_i)$ by a
11: **for** $j = i+1, i+2, \ldots m$
12: **if** $\mathrm{lv}(g_j) = \mathrm{lv}(g_i)$
13: Denote the coefficient of $\mathrm{lv}(g_j)$ by b
14: Replace g_j with $a g_j - b g_i$
15: **return** G

Note that G is in triangular form if and only if G is in i-subtriangular form for all $i = 1, 2, \ldots, m$. This is fairly straightforward, since line 8 ensures that all the zero polynomials occur at the end of the list, as well as $\mathrm{lv}(g_i) > \mathrm{lv}(g_{i+j})$ for any $j \geq 1$.

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

$$g_j = c_{1,j} f_1 + c_{2,j} f_2 + \cdots + a_{m,j} f_m.$$

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 11.8. There are other ways to define both triangular form and Gaussian elimination. Our method is perhaps stricter than necessary, but we have chosen this definition first to keep matters relatively simple, and second to assist us in the development of Gröbner bases.

Example 11.9. We use Algorithm 7 to illustrate Gaussian elimination for the system of equations described in Example 11.2.

- We start with the input,

$$F = (\quad 5x_1 + x_2 + 7x_5 + 8, \quad x_3 + 11x_4 + 2x_5 + 12, \quad 3x_1 + 7x_2 + 8x_3 + 11 \quad).$$

- Line 6 tells us to set $G = F$, so now

$$G \quad = \quad (\quad 5x_1 + x_2 + 7x_5 + 8, \quad x_3 + 11x_4 + 2x_5 + 12, \quad 3x_1 + 7x_2 + 8x_3 + 11 \quad).$$

- We now enter an *outer* loop:
  - In the first iteration, $i = 1$.
  - We rearrange G, obtaining

$$G \quad = \quad (\quad 5x_1 + x_2 + 7x_5 + 8, \quad 3x_1 + 7x_2 + 8x_3 + 11, \quad x_3 + 11x_4 + 2x_5 + 12 \quad).$$

  - Since $g_i \neq 0$, Line 10 tells us to denote a as the coefficient of $\operatorname{lv}(g_i)$, so $a = 5$.
  - We now enter an *inner* loop:
    - In the first iteration, $j = 2$.
    - Since $\operatorname{lv}\left(g_j\right) = \operatorname{lv}(g_i)$, denote b as the coefficient of $\operatorname{lv}\left(g_j\right)$; since $\operatorname{lv}\left(g_j\right) = x_1$, $b = 3$.
    - Replace g_j with

$$\begin{aligned} a g_j - b g_i &= 5\,(3x_1 + 7x_2 + 8x_3 + 11) \\ &\quad - 3\,(5x_1 + x_2 + 7x_5 + 8) \\ &= 32x_2 + 40x_3 - 21x_5 + 31. \end{aligned}$$

      Recall that the field is $\mathbb{Z}_{13}$, so we can rewrite this as

$$6x_2 + x_3 + 5x_5 + 5.$$

      We now have

$$G \quad = \quad (\quad 5x_1 + x_2 + 7x_5 + 8, \quad 6x_2 + x_3 + 5x_5 + 5, \quad x_3 + 11x_4 + 2x_5 + 12 \quad).$$

  - We continue with the inner loop:
    - In the second iteration, $j = 3$.
    - Since $\operatorname{lv}\left(g_j\right) \neq \operatorname{lv}(g_i)$, we do not proceed further.
  - Now $j = 3 = m$, and the inner loop is finished.
- We continue with the outer loop:
  - 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 $\operatorname{lv}(g_i)$; since $\operatorname{lv}(g_i) = x_2$, $a = 6$.
  - We now enter an *inner* loop:
    - In the first iteration, $j = 2$.
    - Since $\operatorname{lv}\left(g_j\right) \neq \operatorname{lv}(g_i)$, 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 11.10. 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 constant.

(B) G has finitely many common solutions if and only if G has a solution and $m = n$. In this case, there is exactly one solution.

(C) G has common solutions of dimension d if and only if G has a solution and $d = n - m$.

A proof of Theorem 11.10 can be found in any textbook on linear algebra, although probably not in one place.

Example 11.11. Continuing with the system that we have used in this section, we found that a triangular form of

$$F = (\ 5x_1 + x_2 + 7x_5 + 8, \quad x_3 + 11x_4 + 2x_5 + 12, \quad 3x_1 + 7x_2 + 8x_3 + 11\ )$$

is

$$G = (\ 5x_1 + x_2 + 7x_5 + 8, \quad 6x_2 + x_3 + 5x_5 + 5, \quad x_3 + 11x_4 + 2x_5 + 12\ ).$$

Theorem 11.10 implies that

(A) G has a solution, because none of the g_i is a constant.

(B) G has infinitely many solutions, because the number of polynomials ($m = 3$) is *not* the same as the number of variables ($n = 5$).

(C) G has solutions of dimension $d = n - m = 2$.

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 $g_2 = 0$,
$$6x_2 = 12x_3 + 8t + 8. \tag{38}$$
The Euclidean algorithm helps us derive the multiplicative inverse of 6 in $\mathbb{Z}_2$; we get 11. Multiplying both sides of (38) by 11, we have
$$x_2 = 2x_3 + 10t + 10.$$
Recall that we found $x_3 = 2s + 11t + 1$, so
$$x_2 = 2(2s + 11t + 1) + 10t + 10 = 4s + 6t + 12.$$
- From $g_1 = 0$,
$$5x_1 = 12x_2 + 6x_5 + 5.$$
Repeating the process that we carried out in the previous step, we find that
$$x_1 = 7s + 9.$$

We can verify this solution by substituting it into the original system:

$$\begin{aligned}
f_1: &\quad 5(7s+9)+(4s+6t+12)+7t+8\\
&= (9s+6)+4s+20\\
&= 0\\
f_2: &\quad (2s+11t+1)+11s+2t+12\\
&= 0\\
f_3: &\quad 3(7s+9)+7(4s+6t+12)+8(2s+11t+1)+11\\
&= (8s+1)+(2s+3t+6)+(3s+10t+8)+11\\
&= 0.
\end{aligned}$$

Before proceeding to the next section, study the proof of Theorem 11.7 carefully. Think about how we might relate these ideas to non-linear polynomials.

Exercises.

Exercise 11.12. 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.

Exercise 11.13. 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 $\mathbb{Q}$.

Exercise 11.14. 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 previous exercise.

Exercise 11.15. Does Algorithm 7 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?

11.2: Monomial orderings

Before looking at how we might analyze systems of nonlinear polynomial equations, we consider the question of identifying the "most important" monomial in this more general setting. With linear polynomials, it was relatively easy; we picked the variable with the smallest index.

But which monomial should be the *leading* monomial of $x + y^3 - 4y$? It seems clear enough that y should not be the leading term, since it divides y^3, and as such does not "lead" even if there were no x's to reckon with. With x and y^3, however, things are less clear.

Recall from Section 7.3 the definition of $\mathbb{M}$, the set of monomials over $x_1, x_2, \ldots, x_n$.

Definition 11.16. Let $t, u \in \mathbb{M}$. 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
- $\ldots$
- $\deg_{x_i} t = \deg_{x_i} u$ for $i = 1, 2, \ldots, n-1$ 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 $j = 1, 2, \ldots, i-1$, 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 11.17. We denote the leading monomial of a polynomial p as $\operatorname{lm}(p)$.

Example 11.18. Using the lexicographic ordering over x, y,

$$\operatorname{lm}\left(x^2 + y^2 - 4\right) = x^2$$
$$\operatorname{lm}(xy - 1) = xy$$
$$\operatorname{lm}\left(x + y^3 - 4y\right) = x.$$

Before proceeding, we should prove a few simple, but important, properties of the lexicographic ordering.

Proposition 11.19. The lexicographic ordering on $\mathbb{M}$

(A) is a linear ordering;
(B) is **compatible with divisibility**: for any $t, u \in \mathbb{M}$, if $t \mid u$, then $t \leq u$;
(C) is **compatible with multiplication**: for any $t, u, v \in \mathbb{M}$, if $t < u$, then for any monomial v over $\mathbf{x}$, $tv < uv$;
(D) orders $1 \leq t$ for any $t \in \mathbb{M}$; and
(E) is a well ordering.

(Recall that we defined a monoid way back in Section 1.1, and used $\mathbb{M}$ as an example.)

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; then $\deg_{x_j} t = \deg_{x_j} u$ for $j = 1, 2, \ldots, i-1$. 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), we know that $t \mid u$ iff $\deg_{x_i} t \leq \deg_{x_i} u$ for all $i = 1, 2, \ldots, m$. Hence $t \leq u$.

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 $M \subset \mathbb{M}$. 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 $E \subset \mathbb{N}$. Recall that $\mathbb{N}$ is well-ordered. Hence E has a least element; call it e. By definition of E, e is the exponent of some monomial m of M. Since $e \leq \alpha$ 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

$$P = \{t \cdot x_n^e : \ t \cdot x_n^e \in M \ \exists e \in \mathbb{N}\}.$$

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, $E \subset \mathbb{N}$. 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 $e \leq \alpha$ 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

$$\deg_{x_j} u = \deg_{x_j} t = \deg_{x_j} v \quad \forall j = 1, \ldots, i-1$$

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. Otherwise, since e is minimal in E,

$$\deg_{x_n} u = e \leq \deg_{x_n} v,$$

in which case $u \leq v$. Hence u is a least element of M.

Since M is arbitrary in $\mathbb{M}$, every subset of $\mathbb{M}$ has a least element. Hence $\mathbb{M}$ is well-ordered. □

Before we start looking for a triangular form of non-linear systems, let's observe one more thing.

Proposition 11.20. Let p be a polynomial in the variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. If $\operatorname{lm}(p) = x_i^\alpha$, then every other monomial u of p has the form

$$u = \prod_{j=i}^{n} x_j^{\beta_j}$$

where $\beta_i < \alpha$.

Proof. Assume that $\operatorname{lm}(p) = x_i^\alpha$. Let u be any monomial of p. Write

$$u = \prod_{j=1}^{n} x_j^{\beta_j}$$

for appropriate $\beta_j \in \mathbb{N}$. Since $u < \operatorname{lm}(p)$, the definition of the lexicographic ordering implies that

$$\deg_{x_j} u = \deg_{x_j} \operatorname{lm}(p) = \deg_{x_j} x_i^\alpha \quad \forall j = 1, 2, \ldots, i-1$$

and

$$\deg_{x_i} u < \deg_{x_i} t.$$

Hence u has the form claimed. □

We now identify and generalize the properties of Proposition 11.19 to a generic ordering on monomials.

Definition 11.21. An **admissible ordering** $<$ on $\mathbb{M}$ is a linear ordering which is compatible with divisibility and multiplication.

By definition, properties (A) and (B) of Proposition 11.19 hold for an admissible ordering. What of the others?

Proposition 11.22. The following properties of an admissible ordering all hold.

(A) $1 \leq t$ for all $t \in \mathbb{M}$.

(B) The set $\mathbb{M}$ of all monomials over $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is well-ordered by any admissible ordering. That is, every subset M of $\mathbb{M}$ has a least element.

Proof. Let $<$ be any admissible ordering.

For (A), you do it! See Exercise 11.32.

For (B), let $t, u \in \mathbb{M}$. By (A), we know that $1 \leq u$. By the ordering's compatibility with multiplication, we know that $t \cdot 1 \leq t \cdot u$, or $t \leq tu$, satisfying compatibility with divisibility.

For (C), let $M \subseteq \mathbb{M}$ and let A be the smallest absorbing subset of $\mathbb{M}$ that contains M (you might want to refamiliarize yourself with absorbing subsets, which we studied in Section 1.4). Dickson's Lemma (Theorem on page 67) tells us that A has a finite generating set; call it T. In fact, $T \subseteq M$, as the definition of absorption means that *every* element of A is divisible by an element of M. There are only finitely many elements of T, so the linear ordering property of $<$ implies that that we can identify a smallest element, t. Let $u \in M$; by definition, $u \in A$, so we can find $v \in T$ such that v divides u. Since $t \leq v$, we use compatibility with divisibility to see that $t \leq v \leq u$. We chose u as an arbitrary element of M, so t is minimal in M. We chose M as an arbitrary subset of $\mathbb{M}$, so $\mathbb{M}$ is well-ordered by $<$. □

We can now introduce an ordering that you haven't seen before.

Definition 11.23. 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 11.24. The total degree of x^3y^2 is 5, and $x^3y^2 < xy^5$.

A simple total degree ordering is not itself admissible, because not it is not linear.

Example 11.25. 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 $x^3y^2 \neq x^2y^3$.

When there is a tie in the total degree, we need to fall back on another method. An interesting way of doing this is the following.

Definition 11.26. For two 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
  - $\deg_{x_i} t > \deg_{x_i} u$.

Notice that 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 11.27. Under grevlex, $x^3y^2 > x^2y^3$ because the total degrees are the same and $y^2 < y^3$.

Theorem 11.28. The graded reverse lexicographic ordering is an admissible ordering.

Proof. Let $t, u \in \mathbb{M}$.

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 $t \mid u$. By definition, $\deg_{x_i} t \leq \deg_{x_i} u$ for all $i = 1, \ldots, n$. If $t = u$, then we're done. Otherwise, $t \neq u$. If $\operatorname{tdeg}(t) > \operatorname{tdeg}(u)$, then the fact that the degrees are all natural numbers implies (see Exercise) that for some $i = 1, \ldots, n$ we have $\deg_{x_i} t > \deg_{x_i} u$, contradicting the hypothesis that $t \mid u$! Hence $\text{tdeg}(t) = \text{tdeg}(u)$. Since $t \neq u$, 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 $j = i+1, \ldots, n$. Since $t \mid u$, $\deg_{x_i} t < \deg_{x_i} u$, and $\deg_{x_j} t \leq \deg_{x_j} u$. Hence

$$\begin{aligned}
\text{tdeg}(t) &= \sum_{j=1}^{i-1} \deg_{x_j} t + \deg_{x_i} t + \sum_{j=i+1}^{n} \deg_{x_j} t \\
&= \sum_{j=1}^{i-1} \deg_{x_j} t + \deg_{x_i} t + \sum_{j=i+1}^{n} \deg_{x_j} u \\
&\leq \sum_{j=1}^{i-1} \deg_{x_j} u + \deg_{x_i} t + \sum_{j=i+1}^{n} \deg_{x_j} u \\
&< \sum_{j=1}^{i-1} \deg_{x_j} u + \deg_{x_i} u + \sum_{j=i+1}^{n} \deg_{x_j} u \\
&= \text{tdeg}(u).
\end{aligned}$$

Hence $t < u$.

Compatible with multiplication? Assume $t < u$, and let $v \in \mathbb{M}$. By definition, $\text{tdeg}(t) < \text{tdeg}(u)$ 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, you will show in the exercises that

$$\begin{aligned}
\text{tdeg}(tv) &= \text{tdeg}(t) + \text{tdeg}(v) \\
&< \text{tdeg}(u) + \text{tdeg}(v) = \text{tdeg}(uv).
\end{aligned}$$

In the second,

$$\deg_{x_i} tv = \deg_{x_i} t + \deg_{x_i} v > \deg_{x_i} u + \deg_{x_i} v = \deg_{x_i} uv$$

while

$$\deg_{x_j} tv = \deg_{x_j} t + \deg_{x_j} v = \deg_{x_j} u + \deg_{x_j} v = \deg_{x_j} uv.$$

In either case, $tv < uv$ as needed. □

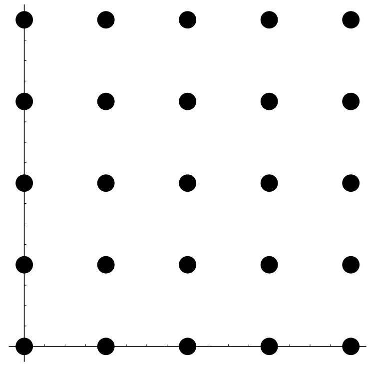
A useful tool when dealing with monomial orderings is a **monomial diagram**. These are most useful for monomials in a bivariate polynomial ring $\mathbb{F}[x, y]$, but we can often imagine important aspects of these diagrams in multivariate rings, as well. We discuss the bivariate case here.

Definition 11.29. Let $t \in \mathbb{M}$. Define the **exponent vector** $(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ where $\alpha_i = \deg_{x_i} t$.

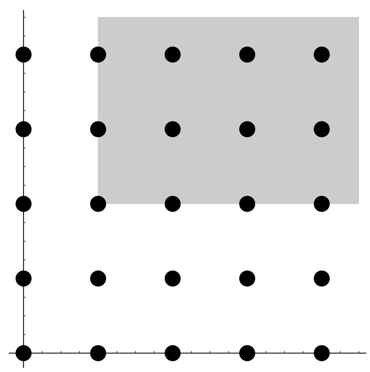
Let $t \in \mathbb{F}[x, y]$ be a monomial, and (α, β) its exponent vector. That is,

$$t = x^{\alpha} y^{\beta}.$$

We can consider (α, β) as a point in the x-y plane. If we do this with all the monomials of $\mathbb{M} \subset \mathbb{F}[x, y]$, and we obtain the following diagram:



This diagram is not especially useful, aside from pointing out that the monomial x^2 is the third point on the left in the bottom row, and the monomial 1 is the point in the lower left corner. What does make diagrams like this useful is the fact that 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. We often shade the points corresponding monomials divisible by a given monomial; for example, the points corresponding to monomials divisible by xy^2 lie within the shaded region of the following diagram:

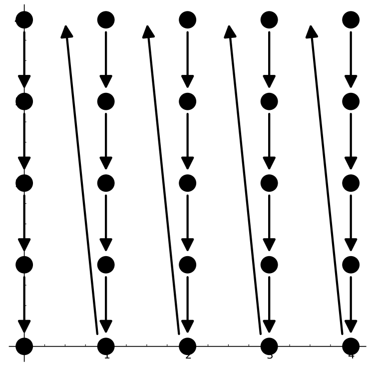


As we will see later, diagrams such as the one above can come in handy when visualizing certain features of an ideal.

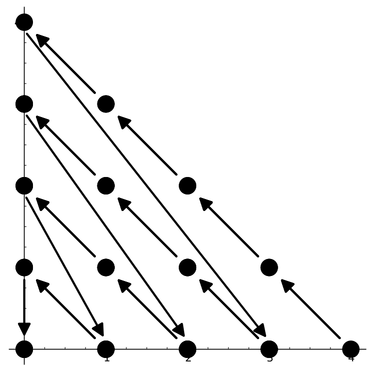
What interests us most for now is that we can sketch vectors on a monomial diagram that show the ordering of the monomials.

Example 11.30. We sketch monomial diagrams that show how lex and grevlex order $\mathbb{M}$. We already know that the smallest monomial is 1. The next smallest will always be y.

For the *lex* order, $y^a < x$ for *every* choice of $a \in \mathbb{N}$, no matter how large. Hence the next largest monomial is y^2, followed by y^3, etc. Once we have marked every power of y, the next largest monomial is x, followed by xy, by xy^2, etc., for $xy^a < x^2$ for all $a \in \mathbb{N}$. Continuing in this fashion, we have the following diagram:



With the *grevlex* order, by contrast, the next largest monomial after y is x, since $\text{tdeg}(x) < \text{tdeg}(y^2)$. After x come y^2, xy, and x^2, in that order, followed by the degree-three monomials y^2, xy^2, x^2y, and x^3, again in that order. This leads to the following monomial diagram:



These diagrams illustrate an important and useful fact.

Theorem 11.31. Let $t \in \mathbb{M}$.

(A) In the lexicographic order, there are infinitely many monomials smaller than t if and only if t is not a power of x_n alone.

(B) In the grevlex order, there are finitely many monomials smaller than t.

Proof. You do it! See Exercise 11.36. □

Exercises.

Exercise 11.32. Show that for any admissible ordering and any $t \in \mathbb{M}$, $1 \leq t$.

Exercise 11.33. The **graded lexicographic order**, which we will denote by **gralex**, orders $t < u$ if

- $\text{tdeg}(t) < \text{tdeg}(u)$, or
- $\text{tdeg}(t) = \text{tdeg}(u)$ 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 $\mathbb{M}$.

Exercise 11.34. Define $\pi_{\leq i}$ as the map from $\mathbb{M}$ to itself that "projects" a monomial in n variables to a monomial in i variables. For example,

$$\pi_{\leq 3}\left(x_1^5 x_2^4 x_4 x_5^2\right) = x_1^5 x_2^4.$$

We can think of $\pi_{\leq i}$ as "chopping" variables $x_{i+1}, x_{i+2}, \ldots, x_n$ off the monomial. More formally, if $0 < i \leq n$, then

$$\pi_{\leq i} : \mathbb{M}_m \to \mathbb{M}_i \quad \text{by} \quad \pi_{\leq i}\left(x_1^{a_1} \cdots x_n^{a_n}\right) = x_1^{a_1} \cdots x_i^{a_i}.$$

Show that the definition of the grevlex ordering is equivalent to the following:

> **Definition 11.35** (Alternate definition of grevlex)**.** We say that $t < u$ if $\operatorname{tdeg}\left(\pi_{\leq i}(t)\right) = \operatorname{tdeg}\left(\pi_{\leq i}(u)\right)$ for $i = n, n-1, \ldots, k+1$ but $\operatorname{tdeg}\left(\pi_i(t)\right) < \operatorname{tdeg}\left(\pi_i(u)\right)$.

Exercise 11.36. Prove Theorem 11.31.

11.3: Matrix representations of monomial orderings

In fact, there are limitless ways to design an admissible ordering.

Example 11.37. Consider the matrix

$$M = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ & & & & -1 \\ & & & -1 & \\ & & \cdots & & \\ & -1 & & & \end{pmatrix}$$

where the empty entries are zeroes. We claim that M represents the grevlex ordering, and weighted vectors computed with M can be read from top to bottom, where the first entry that does not tie determines the larger monomial.

Why? The top row of M adds all the elements of the exponent vector, so the top entry of the weighted vector is the total degree of the monomial. If the two monomials have different total degrees, the top entry of the weighted vector determines the larger monomial. In case they have the same total degree, the second entry of $M\mathbf{t}$ contains $-\deg_{x_n} t$, so if they have different degree in the smallest variable, the second entry determines the larger monomial. And so forth.

The monomials $t = x^3y^2$, $u = x^2y^3$, and $v = xy^5$ have exponent vectors $\mathbf{t} = (3,2)$, $\mathbf{u} = (2,3)$, and $\mathbf{v} = (1,5)$, respectively. We have

$$M\mathbf{t} = \begin{pmatrix} 5 \\ -2 \end{pmatrix}, \qquad M\mathbf{u} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}, \qquad M\mathbf{v} = \begin{pmatrix} 6 \\ -5 \end{pmatrix},$$

from which we conclude that $v > t > u$.

Definition 11.38. Let $M \in \mathbb{R}^{n \times n}$. If $\mathbf{t} \in \mathbb{N}^n$, the **weight** of $\mathbf{t}$ is $w(\mathbf{t}) = M\mathbf{t}$. Similarly, if $t \in \mathbb{M}_n$, the **weight** of t is the weight of its exponent vector.

Not all matrices can represent admissible orderings.

Theorem 11.39. Let $M \in \mathbb{R}^{m \times m}$. M represents a admissible ordering if and only if its rows are linearly independent over $\mathbb{Z}$ and the topmost nonzero entry in each column is positive.

To prove the theorem, we need the following lemma.

Lemma 11.40. If M satisfies the criteria of Theorem 11.39, then there exists a matrix N that satisfies (B), whose entries are all nonnegative, and for all $\mathbf{t} \in \mathbb{Z}^n$ comparison from top to bottom implies that $N\mathbf{t} > N\mathbf{u}$ if and only if $M\mathbf{t} > M\mathbf{u}$.

Example 11.41. In Example 11.37, we saw that grevlex could be represented by

$$M = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ & & & & -1 \\ & & & -1 & \\ & & \cdots & & \\ & -1 & & & \end{pmatrix}.$$

However, it can also be represented by

$$N = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 & \\ & \cdots & & & \\ 1 & 1 & & & \\ 1 & & & & \end{pmatrix}$$

where the empty entries are, again, zeroes. Notice that the first row operates exactly the same, while the second row adds all the entries *except the last*. If $t_n < y_n$ then from $t_1 + \cdots + t_n = u_1 + \cdots + u_n$ we infer that $t_1 + \cdots + t_{n-1} > u_1 + \cdots + u_{n-1}$, so the second row of $N\mathbf{t}$ and $N\mathbf{u}$ would break the tie in exactly the same way as the second row of $M\mathbf{t}$ and $M\mathbf{u}$. And so forth.

Remark 11.42.

1. We can obtain N by adding row 1 of M to row 2 of M, then adding the modified row 2 of M to the modified row 3, and so forth. This is the essence of the proof of Lemma 11.40.
2. While M corresponds to our original definition of grevlex ordering, N corresponds to the definition given in Exercise 11.34

Proof of Lemma 11.40. Let $M \in \mathbb{R}^{n \times n}$ satisfy the criteria of Theorem 11.39. Construct N by building matrices $M_0, M_1, \ldots$ in the following way.

Let $M_1 = M$. Suppose that $M_1, M_2, \ldots, M_{i-1}$ all have nonnegative entries in rows $1, 2, \ldots, i-1$ but M has a negative entry α in row i, column j. By hypothesis, the topmost nonzero entry

β of column j in M_{i-1} is positive; say it is in row k. Use the Archimedean property of $\mathbb{R}$ to find $K \in \mathbb{N}^+$ such that $K\beta \geq |\alpha|$, and add K times row k of M_{i-1} to row j. The entry in row i and column j of M_i is now nonnegative. If there were other negative values in row i of M_i, the fact that row k of M_{i-1} contained nonnegative entries implies that the absolute values of these negative entries are no larger than before. There is a finite number of entries in each row, and a finite number of rows in M, so this process terminates after finitely many additions with a matrix N whose entries are all nonnegative.

In addition, we can write the ith row $N_{(i)}$ of N as

$$N_{(i)} = K_1 M_{(1)} + K_2 M_{(2)} + \cdots + K_i M_{(i)}$$

where $M_{(k)}$ indicates the kth row of M. For any $\mathbf{t} \in \mathbb{M}$, the ith entry of $N\mathbf{t}$ is therefore

$$\begin{aligned} N_{(i)}\mathbf{t} &= \left(K_1 M_{(1)} + K_2 M_{(2)} + \cdots + K_i M_{(i)}\right)\mathbf{t} \\ &= K_1\left(M_{(1)}\mathbf{t}\right) + K_2\left(M_{(2)}\mathbf{t}\right) + \cdots + K_i\left(M_{(i)}\mathbf{t}\right). \end{aligned}$$

We see that if $M_{(1)}\mathbf{t} = \cdots = M_{(i-1)}\mathbf{t} = 0$ and $M_{(i)}\mathbf{t} = \alpha \neq 0$, then $N_{(1)}\mathbf{t} = \cdots = N_{(i-1)}\mathbf{t} = 0$ and $N_{(i)}\mathbf{t} = K_i\alpha \neq 0$. Hence $N\mathbf{t} > N\mathbf{u}$ if and only if $N_{(i)}\mathbf{t} > N_{(i)}\mathbf{u}$ if and only if $K_i\alpha > K_i\left(M_{(i)}\mathbf{u}\right)$ if and only if $M\mathbf{t} > M\mathbf{u}$. □

Now we can prove Theorem 11.39.

Proof of Theorem 11.39. *That (A) implies (B):* Assume that M represents an admissible ordering.

The monomial 1 has the exponent vector $\mathbf{t} = (0, \ldots, 0)$, while the monomial x_i has the exponent vector $\mathbf{u}$ with zeroes everywhere except in the ith position. The product $M\mathbf{t} > M\mathbf{u}$ if the ith element of the top row of M is negative, but this contradicts Proposition 11.22(A).

In addition, property of Definition 11.21 implies that no pair of distinct monomials can produce the same weighted vector. Hence the rows of M are linearly independent over $\mathbb{Z}$.

That (B) implies (A): Assume that M satisfies the criteria of the theorem. We need to show that the properties of an admissible order (Definition 11.21) are satisfied.

Linear ordering? Since the rows of M are linearly independent over $\mathbb{Z}$, every pair of monomials t and u produces a pair of distinct weighted vectors $M\mathbf{t}$ and $M\mathbf{u}$ if and only if $t \neq u$. Reading these vectors from top to bottom allows us to decide whether $t > u$, $t < u$, or $t = u$.

Compatible with divisibility? This follows from linear algebra. Let $t, u \in \mathbb{M}$, and assume that $t \mid u$. Then $\deg_{x_i} t \leq \deg_{x_i} u$ for all $i = 1, 2, \ldots, n$. In the exponent vectors $\mathbf{t}$ and $\mathbf{u}$, $t_i \leq u_i$ for each i. Let $\mathbf{v} \in \mathbb{N}^n$ such that $\mathbf{u} = \mathbf{t} + \mathbf{v}$; then

$$M\mathbf{u} = M(\mathbf{t} + \mathbf{v}) = M\mathbf{t} + M\mathbf{v}.$$

From Lemma 11.40 we can assume that the entries of M are all nonnegative. Thus the entries of $M\mathbf{u}$, $M\mathbf{t}$, and $M\mathbf{v}$ are also nonnegative. Thus the topmost nonzero entry of $M\mathbf{v}$ is positive, and $M\mathbf{u} > M\mathbf{t}$.

Compatible with multiplication? This is similar to compatibility with divisibility, so we omit it. □

In the Exercises you will find other matrices that represent term orderings, some of them somewhat exotic.

Exercises

Exercise 11.43. Find a matrix that represents (a) the lexicographic term ordering, and (b) the gralex ordering.

Exercise 11.44. Explain why the matrix

$$M = \begin{pmatrix} 1 & 1 & & & & & & \\ 1 & 1 & 1 & & & & & \\ 1 & 1 & 1 & 1 & & & & \\ -1 & & & & & & & \\ & & & & 1 & 1 & 1 & 1 \\ & & & & & 1 & 1 & 1 \\ & & & & & & 1 & 1 \\ & & & & & & & -1 \end{pmatrix}$$

represents an admissible ordering. Use M to order the monomials

$$x_1x_3^2x_4x_6, \quad x_1x_4^8x_7, \quad x_2x_3^2x_4x_6, \quad x_8, \quad x_8^2, \quad x_7x_8.$$

Exercise 11.45. Suppose you know nothing about an admissible order $<$ on $\mathbb{F}[x,y]$ except that $x > y$ and $x^2 < y^3$. Find a matrix that represents this order.

Exercise 11.46. Define an ordering on $\mathbb{M}_2$ as follows: A monomial's weight is the dot product of a monomial's exponent vector with the "weight vector" $(5,3)$. Monomials with larger weight are considered larger. For example, $x > y$ because $(5,3)\cdot(1,0) = 5 > 3 = (5,3)\cdot(0,1)$.
(a) Why is this ordering not admissible?
(b) Extend the ordering to a matrix ordering that *is* admissible. By "extend", we mean that if this ordering would decide that $t > u$, then so should the matrix ordering.

11.4: The structure of a Gröbner basis

Throughout this section, assume an admissible ordering of monomials.

When we consider the non-linear case, things become a little more complicated. Consider the following system of equations:

$$\begin{aligned} x^2 + y^2 &= 4 \\ xy &= 1. \end{aligned}$$

We can visualize the real solutions to this system; see Figure 11.1 on the next page. The common solutions occur wherever the circle and the hyperbola intersect. We see four intersections in the real plane; one of them is highlighted with a dot.

However, we don't know if complex solutions exist. In addition, plotting equations involving more than two variables is difficult, and more than three is effectively impossible. Finally, while it's relatively easy to solve the system given above, it isn't a "triangular" system in the sense that

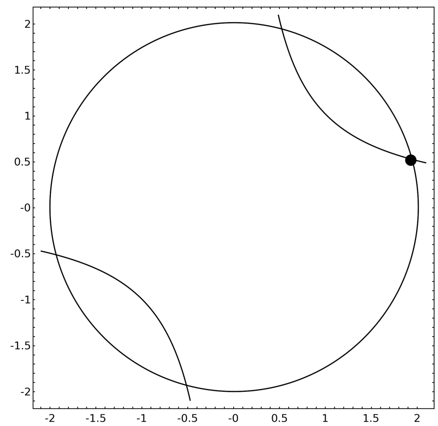


Figure 11.1. Plots of $x^2+y^2=4$ and $xy=1$

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.

It gets worse! Although the system is triangular in a "linear" sense, it is not triangular in a non-linear sense: we can multiply the two polynomials above by monomials and obtain a new polynomial that isn't obviously spanned by either of these two:

$$y\left(x^2+y^2-4\right)-x\,(xy-1)=x+y^3-4y. \tag{39}$$

None of the terms of this new polynomial appears in either of the original polynomials. This sort of thing does *not* happen in the linear case, largely because

- cancellation of *variables* can be resolved using *scalar multiplication*, hence in a vector space; but
- cancellation of *terms* cannot be resolved without *monomial multiplication*, hence it requires an ideal.

So we need to find a "triangular form" for non-linear systems.

Let's rephrase this problem in the language of rings and ideals. The primary issue we would like to resolve is the one that we observed immediately after computing the subtraction polynomial of equation (39): we built a polynomial p whose leading term x was not divisible by the leading term of either the hyperbola (xy) or the circle (x^2). When we built p, we used operations of the polynomial ring that allowed us to remain within the ideal generated by the hyperbola and the circle. That is,

$$p=x+y^3-4y=y\left(x^2+y^2-4\right)-x\,(xy-1)\,;$$

by Theorem 8.7 ideals absorb multiplication and are closed under subtraction, so

$$p\in\left\langle x^2+y^2-4,xy-1\right\rangle.$$

So one problem appears to be that p is in the ideal, but its leading monomial is not divisible by the leading monomials of the ideal's basis. Let's define a special kind of ideal basis that will not give us this problem.

Definition 11.47. 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 how we can decide that any given basis forms a Gröbner basis, because there are infinitely many polynomials that we'd have to check. However, we can certainly determine that the list

$$\left(x^2+y^2-4, xy-1\right)$$

is *not* a Gröbner basis, because we found a polynomial in its ideal that violated the definition of a Gröbner basis: $x+y^3-4y$.

How did we find that polynomial? We built a *subtraction polynomial* that was calculated in such a way as to "raise" the polynomials to the lowest level where their leading monomials would cancel! Let t, u be monomials in the variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. 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 $\gamma_i \geq \alpha_i$ and $\gamma_i \geq \beta_i$ 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 11.48. For any two polynomials $p, q \in \mathbb{F}[x_1, x_2, \ldots, x_n]$, with $\operatorname{lm}(p) = t$ and $\operatorname{lm}(q) = u$, we can build a polynomial in the ideal of p and q that would raise the leading terms to the smallest level where they would cancel by computing

$$S \quad = \quad \operatorname{lc}(q) \cdot \frac{\operatorname{lcm}(t, u)}{t} \cdot p \quad - \quad \operatorname{lc}(p) \cdot \frac{\operatorname{lcm}(t, u)}{u} \cdot q.$$

Moreover, for all other monomials τ, μ and $a, b \in \mathbb{F}$, if $a\tau p - b\mu q$ cancels the leading terms of τp and μq, then it is a multiple of S.

Proof. First we show that the leading monomials of the two polynomials in the subtraction cancel. By Proposition 11.19,

$$\begin{aligned} \operatorname{lm}\left(\frac{\operatorname{lcm}(t, u)}{t} \cdot p\right) &= \frac{\operatorname{lcm}(t, u)}{t} \cdot \operatorname{lm}(p) \\ &= \frac{\operatorname{lcm}(t, u)}{t} \cdot t = \operatorname{lcm}(t, u); \end{aligned}$$

likewise

$$\operatorname{lm}\left(\frac{\operatorname{lcm}(t,u)}{u}\cdot q\right)=\frac{\operatorname{lcm}(t,u)}{u}\cdot \operatorname{lm}(q)$$
$$=\frac{\operatorname{lcm}(t,u)}{u}\cdot u=\operatorname{lcm}(t,u).$$

Thus

$$\operatorname{lc}\left(\operatorname{lc}(q)\cdot\frac{\operatorname{lcm}(t,u)}{t}\cdot p\right)=\operatorname{lc}(q)\cdot\operatorname{lc}(p)$$

and

$$\operatorname{lc}\left(\operatorname{lc}(p)\cdot\frac{\operatorname{lcm}(t,u)}{t}\cdot q\right)=\operatorname{lc}(p)\cdot\operatorname{lc}(q).$$

Hence the leading monomials of the two polynomials in S cancel.

Let τ,μ be monomials over $\mathbf{x}=(x_1,x_2,\ldots,x_n)$ and $a,b\in\mathbb{F}$ such that the leading monomials of the two polynomials in $a\tau p-b\mu q$ cancel. Let $\tau=x_1^{\alpha_1}\cdots x_n^{\alpha_n}$ and $\mu=x_1^{\beta_1}\cdots x_n^{\beta_n}$ for appropriate α_i and β_i in $\mathbb{N}$. Write $\operatorname{lm}(p)=x_1^{\zeta_1}\cdots x_n^{\zeta_n}$ and $\operatorname{lm}(q)=x_1^{\omega_1}\cdots x_n^{\omega_n}$ for appropriate ζ_i and ω_i in $\mathbb{N}$. The leading monomials of $a\tau p-b\mu q$ cancel, so for each $i=1,2,\ldots,n$

$$\alpha_i+\zeta_i=\beta_i+\omega_i.$$

We have

$$\alpha_i=\beta_i+(\omega_i-\zeta_i).$$

Rewrite this as

$$\alpha_i-(\max(\zeta_i,\omega_i)-\zeta_i)=[(\beta_i+(\omega_i-\zeta_i))-(\max(\zeta_i,\omega_i)-\zeta_i)]$$
$$=\beta_i-(\max(\zeta_i,\omega_i)-\omega_i).$$

Let $\eta_i=\alpha_i-(\max(\zeta_i,\omega_i)-\zeta_i)$ and let

$$v=\prod_{i=1}^{n}x_i^{\eta_i}.$$

Then

$$a\tau p-b\mu q=v\left(a\cdot\frac{\operatorname{lcm}(t,u)}{t}\cdot p-b\cdot\frac{\operatorname{lcm}(t,u)}{u}\cdot q\right),$$

as claimed. □

The subtraction polynomial of Lemma 11.48 is important enough that we give it a special name.

Definition 11.49. Let $p, q \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. We define the ***S*-polynomial of p and q** to be

$$\begin{aligned}\operatorname{Spol}(p,q) = \operatorname{lc}(q) \cdot \frac{\operatorname{lcm}(\operatorname{lm}(p), \operatorname{lm}(q))}{\operatorname{lm}(p)} \cdot p \\ - \operatorname{lc}(p) \cdot \frac{\operatorname{lcm}(\operatorname{lm}(p), \operatorname{lm}(q))}{\operatorname{lm}(q)} \cdot q.\end{aligned}$$

Hopefully, you see that S-poly-nomials generalize the cancellation of Gaussian elimination in a natural way.

For some S-polynomials, only one of the leading terms needs to change. This merits its own terminology.

Definition 11.50. Let $p, q \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. If $\operatorname{lm}(p)$ divides $\operatorname{lm}(q)$, then we say that p **top-reduces** q.

If p top-reduces q, let $r = \operatorname{Spol}(p, q)$. We say that p **top-reduces** q **to** r.

Finally, let $F = (f_1, f_2, \ldots, f_m)$ be a list of polynomials in $\mathbb{F}[x_1, x_2, \ldots, x_n]$, and $r_1, r_2, \ldots, r_k \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ such that

- some polynomial of F top-reduces p to r_1,
- some polynomial of F top-reduces r_1 to r_2,
- ...
- some polynomial of F top-reduces r_{k-1} to r_k.

In this case, we say that p **top-reduces to r_k with respect to F**.

Example 11.51. Let $p = x + 1$ and $q = x^2 + 1$. We have $\operatorname{lm}(p) = x$ and $\operatorname{lm}(q) = x^2$. Since $\operatorname{lm}(p)$ divides $\operatorname{lm}(q)$, p top-reduces q. Their S-polynomial is

$$r = q - x \cdot p = -x + 1,$$

so q top-reduces to r with respect to $\{p\}$.

We need the following properties of polynomial operations.

Proposition 11.52. Let $p, q, r \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. Each of the following holds:

(A) $\operatorname{lm}(pq) = \operatorname{lm}(p) \cdot \operatorname{lm}(q)$

(B) $\operatorname{lm}(p \pm q) \leq \max(\operatorname{lm}(p), \operatorname{lm}(q))$

(C) $\operatorname{lm}(\operatorname{Spol}(p,q)) < \operatorname{lcm}(\operatorname{lm}(p), \operatorname{lm}(q))$

(D) If p top-reduces q to r, then $\operatorname{lm}(r) < \operatorname{lm}(q)$.

Proof. For convenience, write $t = \operatorname{lm}(p)$ and $u = \operatorname{lm}(q)$.

(A) Any monomial of pq can be written as the product of two monomials vw, where v is a monomial of p and w is a monomial of q. If $v \neq \operatorname{lm}(p)$, then the definition of a leading monomial implies that $v < t$. Proposition 11.19 implies that

$$vw \leq tw,$$

with equality only if $v = t$. The same reasoning implies that

$$vw \leq tw \leq tu,$$

with equality only if $w = u$. Hence

$$\operatorname{lm}(pq) = tu = \operatorname{lm}(p)\operatorname{lm}(q).$$

(B) Any monomial of $p \pm q$ is a monomial of p or of q. Hence $\operatorname{lm}(p \pm q)$ is a monomial of p or of q. The maximum of these is $\max(\operatorname{lm}(p), \operatorname{lm}(q))$. Hence $\operatorname{lm}(p \pm q) \leq \max(\operatorname{lm}(p), \operatorname{lm}(q))$.

(C) Definition 11.49 and (B) imply $\operatorname{lm}(\operatorname{Spol}(p,q)) < \operatorname{lcm}(\operatorname{lm}(p), \operatorname{lm}(q))$.

(D) By definition, top-reduction is a kind of S-polynomial, so this follows from (C). □

In a triangular linear system, we achieve a triangular form by rewriting all polynomials that share a leading variable. In the *linear* case we can accomplish this using *scalar multiplication*, requiring nothing else. In the non-linear case, we need to check for divisibility of monomials. The following result should, therefore, not surprise you very much.

Theorem 11.53 (Buchberger's characterization)**.** Let $G = \{g_1, g_2, \ldots, g_m\} \subsetneq \mathbb{F}[x_1, x_2, \ldots, x_n]$. It is a Gröbner basis of the ideal $I = \langle g_1, g_2, \ldots, g_m \rangle$ if and only if $\operatorname{Spol}(g_i, g_j)$ top-reduces to zero with respect to G for each pair i, j with $1 \leq i < j \leq m$.

Example 11.54. Recall two systems considered at the beginning of this chapter,

$$F = \left(x^2 + y^2 - 4, xy - 1\right)$$

and

$$G = \left(x^2 + y^2 - 4, xy - 1, x + y^3 - 4y, -y^4 + 4y^2 - 1\right).$$

Is either of these a Gröbner basis?

- We already showed that F is not, as its one S-polynomial is

$$\begin{aligned} S &= \operatorname{Spol}(f_1, f_2) \\ &= y\left(x^2 + y^2 - 4\right) - x(xy - 1) \\ &= x + y^3 - 4y, \end{aligned}$$

  and $\operatorname{lm}(S) = x$, which neither leading term of F divides.
- On the other hand, G is a Gröbner basis. We will not show all six S-polynomials (you will verify this in Exercise 11.57), but

$$\operatorname{Spol}(g_1, g_2) - g_3 = 0,$$

  so the problem with F does not reappear. It is worth noting that

$$\operatorname{Spol}(g_1, g_4) - \left(4y^2 - 1\right) g_1 + \left(y^2 - 4\right) g_4 = 0.$$

If we rewrite $\text{Spol}(g_1, g_4) = y^4 g_1 + x^2 g_4$ and substitute it into the above equation, something very interesting turns up:

$$\begin{aligned}\left(y^4 g_1 + x^2 g_4\right) - \left(4y^2 - 1\right) g_1 + \left(y^2 - 4\right) g_4 &= 0\\ -\left(-y^4 + 4y^2 - 1\right) g_1 + \left(x^2 + y^2 - 4\right) g_4 &= 0\\ -g_4 g_1 + g_1 g_4 &= 0.\end{aligned}$$

Remark 11.55. Buchberger's characterization suggests a method to compute a Gröbner basis of an ideal: given a basis, use S-polynomials to find elements of the ideal that do not satisfy Definition 11.47. Add these to the basis, repeating until all of them reduce to zero.

This approach has two wrinkles we have to iron out:

- We don't know that a Gröbner basis exists for every ideal. For all we know, there may be ideals for which no Gröbner basis exists.
- We don't know that the proposed method will even terminate! It could be that we can go on forever, adding new polynomials to the ideal without ever stopping.

We resolve these questions in the following section.

It remains to prove Theorem 11.53, but before we can do that we will need the following useful lemma. While small, it has important repercussions later.

Lemma 11.56. 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 $\text{lm}(q_k)\,\text{lm}(g_k) \leq \text{lm}(p)$.

Proof. You do it! See Exercise 11.63. □

You will see in the following that Lemma 11.56 allows us to replace polynomials that are "too large" with smaller polynomials. This allows us to obtain the desired form.

Proof of Theorem 11.53. Assume first that G is a Gröbner basis, and let i, j be such that $1 \leq i < j \leq m$. Then

$$\text{Spol}\left(g_i, g_j\right) \in \left\langle g_i, g_j\right\rangle \subset \langle g_1, g_2, \ldots, g_m\rangle,$$

and the definition of a Gröbner basis implies that there exists $k_1 \in \{1, 2, \ldots, m\}$ such that g_{k_1} top-reduces $\text{Spol}\left(g_i, g_j\right)$ to a new polynomial, say r_1. The definition further implies that if r_1 is not zero, then there exists $k_2 \in \{1, 2, \ldots, m\}$ such that g_{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 11.52, we see that

$$\text{lm}(r_1) > \text{lm}(r_2) > \cdots.$$

Recall that the monomials are well-ordered under an admissible ordering, so any set of monomials has a least element, including the set $R = \{\text{lm}(r_1), \text{lm}(r_2), \ldots\}$. Thus 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:

  - subtraction by the subring property, and
  - multiplication by the absorption property; hence

- by the definition of a Gröbner basis, a non-zero r_{last} would be top-reducible by some element of G.

The chain of top-reductions must conclude with zero, so $\text{Spol}\left(g_i, g_j\right)$ top-reduces to zero.

Now assume every S-polynomial top-reduces to zero modulo G. We want to show any element of I is top-reducible by an element of G. 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

$$p = h_1 g_1 + \cdots + h_m g_m.$$

For each i, write $t_i = \text{lm}(g_i)$ and $u_i = \text{lm}(h_i)$. 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 $\text{lm}(p) = T$, then we are done, since

$$\text{lm}(p) = T = u_k t_k = \text{lm}(h_k)\,\text{lm}(g_k) \quad \exists k \in \{1, 2, \ldots, m\}.$$

Otherwise, there must be some cancellation among the leading monomials of each polynomial in the sum on the right hand side. That is,

$$T = \text{lm}\left(h_{\ell_1} g_{\ell_1}\right) = \text{lm}\left(h_{\ell_2} g_{\ell_2}\right) = \cdots = \text{lm}\left(h_{\ell_s} g_{\ell_s}\right)$$

for some $\ell_1, \ell_2, \ldots, \ell_s \in \{1, 2, \ldots, m\}$. From Lemma 11.48, we know that we can write the sum of these leading terms as a sum of multiples of a S-polynomials of G. That is,

$$\begin{aligned}\text{lc}\left(h_{\ell_1}\right)\text{lm}\left(h_{\ell_1}\right) g_{\ell_1} + \cdots + \text{lc}\left(h_{\ell_s}\right)\text{lm}\left(h_{\ell_s}\right) g_{\ell_s} &= \\ &= \sum_{1 \le a < b \le s} c_{a,b} u_{a,b} \text{Spol}\left(g_{\ell_a}, g_{\ell_b}\right)\end{aligned}$$

where for each a, b we have $c_{a,b} \in \mathbb{F}$ and $u_{a,b} \in \mathbb{M}$. Let

$$S = \sum_{1 \le a < b \le s} c_{a,b} u_{a,b} \text{Spol}\left(g_{\ell_a}, g_{\ell_b}\right).$$

Observe that

$$\left[\text{lm}\left(h_{\ell_1}\right) g_{\ell_1} + \text{lm}\left(h_{\ell_2}\right) g_{\ell_2} + \cdots + \text{lm}\left(h_{\ell_s}\right) g_{\ell_s}\right] - S = 0. \tag{40}$$

By hypothesis, each S-polynomial of S top-reduces to zero. This fact, Lemma 11.56 and Proposition 11.52, implies that for each a, b we can find $q_\lambda^{(a,b)} \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ such that

$$\text{Spol}\left(g_{\ell_a}, g_{\ell_b}\right) = q_1^{(a,b)} g_1 + \cdots + g_m^{(a,b)} g_m$$

and for each $\lambda = 1, 2, \ldots, m$ we have $q_\lambda^{(a,b)} = 0$ or

$$\begin{aligned}\operatorname{lm}\left(q_\lambda^{(a,b)}\right)\operatorname{lm}(g_\lambda) &\leq \operatorname{lm}\left(\operatorname{Spol}\left(g_{\ell_a}, g_{\ell_b}\right)\right)\\ &< \operatorname{lcm}\left(\operatorname{lm}\left(g_{\ell_a}\right), \operatorname{lm}\left(g_{\ell_b}\right)\right).\end{aligned} \tag{41}$$

Let $Q_1, Q_2, \ldots, Q_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ such that

$$Q_k = \begin{cases} \sum_{1 \leq a < b \leq s} c_{a,b} u_{a,b} q_k^{(a,b)}, & k \in \{\ell_1, \ldots, \ell_s\};\\ 0, & \text{otherwise.}\end{cases}$$

Then

$$S = Q_1 g_1 + Q_2 g_2 + \cdots + Q_m g_m.$$

In other words,

$$S - (Q_1 g_1 + Q_2 g_2 + \cdots + Q_m g_m) = 0.$$

By equation (41) and Proposition 11.52, for each $k = 1, 2, \ldots, m$ we have $Q_k = 0$ or

$$\begin{aligned}\operatorname{lm}(Q_k)\operatorname{lm}(g_k) &\leq \max_{1 \leq a < b \leq s}\left\{\left[u_{a,b}\operatorname{lm}\left(q_k^{(a,b)}\right)\right]\operatorname{lm}(g_k)\right\}\\ &= \max_{1 \leq a < b \leq s}\left\{u_{a,b}\left[\operatorname{lm}\left(q_k^{(a,b)}\right)\operatorname{lm}(g_k)\right]\right\}\\ &\leq \max_{1 \leq a < b \leq s}\left\{u_{a,b}\operatorname{lm}\left(\operatorname{Spol}\left(g_{\ell_a}, g_{\ell_b}\right)\right)\right\}\\ &< u_{a,b}\operatorname{lcm}\left(\operatorname{lm}\left(g_{\ell_a}\right), \operatorname{lm}\left(g_{\ell_b}\right)\right)\\ &= T.\end{aligned} \tag{42}$$

By substitution,

$$\begin{aligned}p &= (h_1 g_1 + h_2 g_2 + \cdots + h_m g_m) - \left(S - \sum_{k \in \{\ell_1, \ldots, \ell_s\}} Q_k g_k\right)\\ &= \left[\sum_{k \notin \{\ell_1, \ldots, \ell_s\}} h_k g_k + \sum_{k \in \{\ell_1, \ldots, \ell_s\}} \left(h_k - \operatorname{lc}(h_k)\operatorname{lm}(h_k)\right) g_k\right]\\ &\quad + \cancelto{0}{\left[\sum_{k \in \{\ell_1, \ldots, \ell_s\}} \operatorname{lc}(h_k)\operatorname{lm}(h_k)\, g_k - S\right]}\\ &\quad + \sum_{k \in \{\ell_1, \ldots, \ell_s\}} Q_k g_k.\end{aligned}$$

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,

$$p = \mathcal{Q}_1 g_1 + \cdots + \mathcal{Q}_m g_m.$$

If $k \notin \{\ell_1, \ldots, \ell_s\}$, then the choice of T as the maximal term of the representation implies that

$$\operatorname{lm}(\mathcal{Q}_k)\operatorname{lm}(g_k) = \operatorname{lm}(h_k)\operatorname{lm}(g_k) < T.$$

Otherwise, Proposition 11.52 and equation (42) imply that

$$\begin{aligned}\operatorname{lm}(\mathcal{Q}_k)\operatorname{lm}(g_k) &\leq \max\left((\operatorname{lm}(h_k - \operatorname{lc}(h_k)\operatorname{lm}(h_k)), \operatorname{lm}(Q_k))\operatorname{lm}(g_k)\right) \\ &< \operatorname{lm}(h_k)\operatorname{lm}(g_k) \\ &= T.\end{aligned}$$

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 $\mathbb{M}$ implies that this cannot continue indefinitely! Hence there must be a representation

$$p = H_1 g_1 + \cdots + H_m g_m$$

such that for each $k = 1, 2, \ldots, m$ $H_k = 0$ or $\operatorname{lm}(H_k)\operatorname{lm}(g_k) \leq \operatorname{lm}(p)$. Both sides of the equation must simplify to the same polynomial, with the same leading variable, so at least one k has $\operatorname{lm}(H_k)\operatorname{lm}(g_k) = \operatorname{lm}(p)$; that is, $\operatorname{lm}(g_k) \mid \operatorname{lm}(p)$. Since p was arbitrary, G satisfies the definition of a Gröbner basis. □

Exercises.

Exercise 11.57. Show that

$$G = \left(xy - 1, \quad x + y^3 - 4y, \quad y^4 - 4y^2 + 1 \right)$$

is a Gröbner basis with respect to the lexicographic ordering.

Exercise 11.58. Show that G of Exercise 11.57 is *not* a Gröbner basis with respect to the grevlex ordering. The Gröbner basis property depends on the choice of term ordering!

Exercise 11.59. Show that any Gröbner basis G of an ideal I is a basis of the same ideal; that is, any $p \in I$ can be written as $p = \sum_{i=1}^{m} h_i g_i$ for appropriate $h_i \in \mathbb{F}[x_1, \ldots, x_n]$.

Exercise 11.60. Show that for any non-constant polynomial f, $F = (f, f+1)$ is not a Gröbner basis.

Exercise 11.61. Show that every list of monomials is a Gröbner basis.

Exercise 11.62. We call a basis G of an ideal a **minimal basis** if no monomial of any $g_1 \in G$ is divisible by the leading monomial of any $g_2 \in G$.

(a) Suppose that a Gröbner basis G is not minimal. Show that we obtain a minimal basis by repeatedly replacing each $g \in G$ by $g - t g'$ where $t \operatorname{lm}(g')$ is a monomial of g.

(b) Explain why the minimal basis obtained in part (a) is also a Gröbner basis of the same ideal.

Exercise 11.63. 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 11.56.

Exercise 11.64. For G to be a Gröbner basis, Definition 11.47 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 Exercise 11.57 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 11.47. 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*.

11.5: Buchberger's algorithm

Algorithm 7 on page 351 shows how to triangularize a linear system. Essentially, it looks for parts of the system that are not triangular (equations with the same leading variable) then adds a new polynomial (an S-polynomial!) to move it closer to the triangular form. The new polynomial replaces one of the older polynomials in the pair.

For non-linear systems, we will try an approach that is similar, not but identical. We *will* look for polynomials in the ideal that do not satisfy the Gröbner basis property, we *will* add a new polynomial to repair this defect. We will not, however, replace the older polynomials, because we may still need their leading monomials, and the S-polynomial may have a very different one. Worse, removing this polynomial could even change the ideal!

Example 11.65. Let $F = (xy + xz + z^2, yz + z^2)$, and use grevlex with $x > y > z$. The S-polynomial of f_1 and f_2 is

$$S = z(xy + xz + z^2) - x(yz + z^2) = z^3.$$

Let $G = (xy + xz + z^2, z^3)$; that is, G is F with f_2 replaced by S. It turns out that $yz + z^2 \notin \langle G \rangle$. If it were, then

$$yz + z^2 = h_1(xy + xz + z^2) + h_2 \cdot z^3.$$

Algorithm 8. Buchberger's algorithm to compute a Gröbner basis

1: **inputs**
2: $F = (f_1, f_2, \ldots, f_m)$, where each $f_i \in \mathbb{F}[x_1, \ldots, x_n]$.
3: $<$, an admissible ordering.
4: **outputs**
5: G, a Gröbner basis of $\langle F \rangle$ with respect to $<$.
6: **do**
7: Let $G := F$
8: Let $P = \{(f, g) : \forall f, g \in G \text{ such that } f \neq g\}$
9: **repeat while** $P \neq \emptyset$
10: Choose $(f, g) \in P$
11: Remove (f, g) from P
12: Let S be the S-polynomial of f, g
13: Let r be the top-reduction of S with respect to G
14: **if** $r \neq 0$
15: Replace P by $P \cup \{(h, r) : h \in G\}$
16: Append r to G
17: **return** G

Every term of the right hand side will be divisible either by x or by z^2, but yz is divisible by neither. Hence $yz + z^2 \in \langle G \rangle$.

We will have to adapt Algorithm 7 without replacing or discarding any polynomials. With non-linear polynomials, Buchberger's characterization (Theorem 11.53) suggests that we compute the S-polynomials, and top-reduce them. If they all top-reduce to zero, then Buchberger's characterization implies that we have a Gröbner basis already, so there is nothing to do. Otherwise, at least one S-polynomial does *not* top-reduce to zero, so we add its reduced form to the basis and test the new S-polynomials as well. This suggests Algorithm 8.

Theorem 11.66. For any list of polynomials F over a field, Buchberger's algorithm terminates with a Gröbner basis of $\langle F \rangle$.

Correctness isn't hard *if* Buchberger's algorithm terminates, because it discards nothing, adds only polynomials that are already in $\langle F \rangle$, and terminates only if all the S-polynomials of G top-reduce to zero. The problem is termination, which relies on the Ascending Chain Condition.

Proof. For *termination*, let $\mathbb{F}$ be a field, and F a list of polynomials over $\mathbb{F}$. Designate

$$
\begin{aligned}
I_0 &= \langle \operatorname{lm}(g_1), \operatorname{lm}(g_2), \ldots, \operatorname{lm}(g_m) \rangle \\
I_1 &= \left\langle \operatorname{lm}(g_1), \operatorname{lm}(g_2), \ldots, \operatorname{lm}(g_m), \operatorname{lm}(g_{m+1}) \right\rangle \\
I_2 &= \left\langle \operatorname{lm}(g_1), \ldots, \operatorname{lm}(g_m), \operatorname{lm}(g_{m+1}), \operatorname{lm}(g_{m+2}) \right\rangle \\
&\vdots \\
I_i &= \left\langle \operatorname{lm}(g_1), \ldots, \operatorname{lm}(g_{m+i}) \right\rangle
\end{aligned}
$$

where g_{m+i} is the ith polynomial added to G by line 16 of Algorithm 8.

We claim that $I_0 \subseteq I_1 \subseteq I_2 \subseteq \cdots$ is a strictly ascending chain of ideals. After all, a polynomial r is added to the basis only when it is non-zero (line 14); since it has not top-reduced to zero, $\text{lm}(r)$ is not top-reducible by

$$G_{i-1} = (g_1, g_2, \ldots, g_{m+i-1}).$$

Thus for any $p \in G_{i-1}$, $\text{lm}(p)$ does not divide $\text{lm}(r)$. We further claim that this implies that $\text{lm}(r) \notin I_{i-1}$. By way of contradiction, suppose that it is. By Exercise 11.61 on page 374, any list of monomials is a Gröbner basis; hence

$$T = (\text{lm}(g_1), \text{lm}(g_2), \ldots, \text{lm}(g_{m+i-1}))$$

is a Gröbner basis, and by Definition 11.47 every polynomial in I_{i-1} is top-reducible by T. Since r is not top-reducible by T, $\text{lm}(r) \notin I_{i-1}$.

Thus $I_{i-1} \subsetneq I_i$, and $I_0 \subseteq I_1 \subseteq I_2 \subseteq \cdots$ is a strictly ascending chain of ideals in $\mathbb{F}[x_1 x_2, \ldots, x_n]$. By Proposition 8.33 and Definition 8.31, there exists $M \in \mathbb{N}$ such that $I_M = I_{M+1} = \cdots$. This implies that the algorithm can add at most $M - m$ polynomials to G; after having done so, any remaining elements of P generate S-polynomials that top-reduce to zero! Line 11 removes each pair (i, j) from P, so P decreases after we have added these $M - m$ polynomials. Eventually P decreases to $\emptyset$, and the algorithm terminates.

For *correctness,* we have to show two things: first, that G is a basis of the same ideal as F, and second, that G satisfies the Gröbner basis property. For the first, observe that every polynomial added to G is by construction an element of $\langle F \rangle$, and we removed no elements from the basis, so the ideal does not change. For the second, observe that the very construction of G ensures that Buchberger's characterization of a Gröbner basis is satisfied. □

Exercises

Exercise 11.67. Using G of Exercise 11.57, compute a Gröbner basis with respect to the grevlex ordering.

Exercise 11.68. Following up on Exercises 11.58 and 11.67, a simple diagram will help show that it is usually "faster" to compute a Gröbner basis in any total degree ordering than it is in the lexicographic ordering. We can diagram the monomials in x and y on the x-y plane by plotting $x^\alpha y^\beta$ at the point (α, β).

(a) Shade the region of monomials that are smaller than $x^2 y^3$ with respect to the lexicographic ordering.

(b) Shade the region of monomials that are smaller than $x^2 y^3$ with respect to the graded reverse lexicographic ordering.

(c) Explain why the diagram implies that top-reduction of a polynomial with leading monomial $x^2 y^3$ will *probably* take less effort in grevlex than in the lexicographic ordering.

Exercise 11.69. However, it is not *always* faster to use the grevlex ordering. To see this, consider the system

$$C_4 = (\quad x_1 + x_2 + x_3 + x_4, \\ x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1, \\ x_1x_2x_3 + x_2x_3x_4 + x_3x_4x_1 + x_4x_1x_2, \\ x_1x_2x_3x_4 - 1 \quad).$$

Compute the size of the Gröbner basis of C_4 over the field $\mathbb{Z}_2$ with respect to grevlex ordering, then with respect to lex ordering.

Exercise 11.70. Let $g_1, g_2, \ldots, g_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. We say that a non-linear polynomial is *homogeneous* if every term is of the same total degree. For example, $xy - 1$ is not homogeneous, but $xy - h^2$ is. As you may have guessed, we can homogenize any polynomial by multiplying every term by an appropriate power of a *homogenizing variable* h. When $h = 1$, we have the original polynomial.

(a) Homogenize the following polynomials.
 (i) $x^2 + y^2 - 4$
 (ii) $x^3 - y^5 + 1$
 (iii) $xz + z^3 - 4x^5y - xyz^2 + 3x$

(b) Explain the relationship between solutions to a system of nonlinear polynomials G and solutions to the system of homogenized polynomials H.

(c) With homogenized polynomials, we usually use a variant of the lexicographic ordering. Although h comes *first* in the dictionary, we pretend that it comes last. So $x > yh^2$ and $y > h^{10}$. Use this modified lexicographic ordering to determine the leading monomials of your solutions for part (a).

(d) Does homogenization preserve leading monomials?

Exercise 11.71. Assume that the $g_1, g_2, \ldots, g_m$ are homogeneous; in this case, we can build the *ordered Macaulay matrix of G of degree D* in the following way.

- Each row of the matrix represents a monomial multiple of some g_i. If g_i is of degree $d \leq D$, then we compute all the monomial multiples of g_i that have degree D.
- Each column represents a monomial. Column 1 corresponds to the largest monomial with respect to the lexicographic ordering; column 2 corresponds to the next-largest polynomial; etc.
- Each entry of the matrix is the coefficient of a monomial for a monomial multiple of some g_i.

(a) The homogenization of the circle and the hyperbola gives us the system

$$F = \left(x^2 + y^2 - 4h^2, xy - h^2\right).$$

Verify that its ordered Macaulay matrix of degree 3 is

$$\begin{pmatrix}
x^3 & x^2y & xy^2 & y^3 & x^2h & xyh & y^2h & xh^2 & yh^2 & h^3 & \\
1 & & 1 & & & & & -4 & & & xf_1 \\
& 1 & & 1 & & & & & -4 & & yf_1 \\
& & & & 1 & & 1 & & & -4 & hf_1 \\
& 1 & & & & & & -1 & & & xf_2 \\
& & 1 & & & & & & -1 & & yf_2 \\
& & & & & 1 & & & & -1 & hf_2
\end{pmatrix}.$$

Show that if you triangularize this matrix *without swapping columns*, the row corresponding to xf_2 now contains coefficients that correspond to the homogenization of $x + y^3 - 4y$.

(b) Compute the ordered Macaulay matrix of F of degree 4, then triangularize it. Be sure *not to swap columns*, nor to destroy rows that provide new information. Show that
   - the entries of at least one row correspond to the coefficients of a multiple of the homogenization of $x + y^3 - 4y$, and
   - the entries of at least one other row are the coefficients of the homogenization of $\pm(y^4 - 4y^2 + 1)$.

(c) Explain the relationship between triangularizing the ordered Macaulay matrix and Buchberger's algorithm.

Sage programs

The following programs can be used in Sage to help make the amount of computation involved in the exercises less burdensome. Use

- `M, mons = sylvester_matrix(F,d)` to make an ordered Macaulay matrix of degree d for the list of polynomials F,
- `N = triangularize_matrix(M)` to triangularize M in a way that respects the monomial order, and
- `extract_polys(N,mons)` to obtain the polynomials of N.

```
def make_monomials(xvars,d,p=0,order="lex"):
  result = set([1])
  for each in range(d):
    new_result = set()
    for each in result:
      for x in xvars:
        new_result.add(each*x)
    result = new_result
  result = list(result)
  result.sort(lambda t,u:  monomial_cmp(t,u))
  n = sage.rings.integer.Integer(len(xvars))
  return result

def monomial_cmp(t,u):
  xvars = t.parent().gens()
  for x in xvars:
    if t.degree(x) != u.degree(x):
      return u.degree(x) - t.degree(x)
  return 0

def homogenize_all(polys):
  for i in range(len(polys)):
    if not polys[i].is_homogeneous():
      polys[i] = polys[i].homogenize()

def sylvester_matrix(polys,D,order="lex"):
  L = [ ]
  homogenize_all(polys)
  xvars = polys[0].parent().gens()
  for p in polys:
    d = D - p.degree()
    R = polys[0].parent()
    mons = make_monomials(R.gens(),d,order=order)
    for t in mons:
      L.append(t*p)
  mons = make_monomials(R.gens(),D,order=order)
```

```
  mons_dict = {}
  for each in range(len(mons)):
    mons_dict.update({mons[each]:each})
  M = matrix(len(L),len(mons))
  for i in range(len(L)):
    p = L[i]
    pmons = p.monomials()
    pcoeffs = p.coefficients()
    for j in range(len(pmons)):
      M[i,mons_dict[pmons[j]]] = pcoeffs[j]
  return M, mons

def triangularize_matrix(M):
  N = M.copy()
  m = N.nrows()
  n = N.ncols()
  for i in range(m):
    pivot = 0
    while pivot < n and N[i,pivot] == 0:
      pivot = pivot + 1
    if pivot < n:
      a = N[i,pivot]
      for j in range(i+1,m):
        if N[j,pivot] != 0:
          b = N[j,pivot]
          for k in range(pivot,n):
            N[j,k] = a * N[j,k] - b * N [i,k]
  return N

def extract_polys(M, mons):
  L = [ ]
  for i in range(M.nrows()):
    p = 0 for j in range(M.ncols()):
    if M[i,j] != 0:
      p = p + M[i,j]*mons[j]
    L.append(p)
  return L
```

11.6: Nullstellensatz

The German word *Nullstellensatz* means "Theorem (*satz*) on the locations (*stellen*) of zero (*null*)." There are two different theorems; a *weak* Nullstellensatz, and a "*not-so-weak*" Nullstellensatz. In this section, we consider only the weak version. Throughout this section,

- $\mathbb{F}$ is an *algebraically closed* field—that is, all nonconstant polynomials over $\mathbb{F}$ have all their roots 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;[27] and
- $I = \langle F \rangle$.

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 $\mathbb{R}$.

An interesting and useful consequence of algebraic closure is the following.

Lemma 11.72. $\mathbb{F}$ is infinite.

Proof. Let $n \in \mathbb{N}^+$, and $a_1, \ldots, a_n \in \mathbb{F}$. Obviously, $f = (x - a_1) \cdots (x - a_n)$ satisfies $f(x) = 0$ for all $x = a_1, \ldots, a_n$. Let $g = f + 1$; then $g(x) \neq 0$ for all $x = a_1, \ldots, a_n$. Since $\mathbb{F}$ is closed, g has a root $b \in \mathbb{F} \setminus \{a_1, \ldots, a_n\}$. Thus, no finite list of elements enumerates $\mathbb{F}$, which means $\mathbb{F}$ must be infinite. □

Theorem 11.73 (Hilbert's Weak Nullstellensatz)**.** If $V_F = \emptyset$, then $I = \mathcal{R}$.

Proof. We proceed by induction on n, the number of variables.

Inductive base: Let $n = 1$. Recall that in this case, $\mathcal{R} = \mathbb{F}[x]$ is a Euclidean domain, and hence a principal ideal domain. Thus $I = \langle f \rangle$ for some $f \in \mathcal{R}$. If $V_F = \emptyset$, then f has no roots in $\mathbb{F}$. Theorem 10.18 tells us that every principal ideal domain is a unique factorization domain, so if f is non-constant, it has a unique factorization into irreducible polynomials. Theorem 10.42 tells us that any irreducible p extends $\mathcal{R}$ to a field $\mathbb{E} = R/\langle p \rangle$ containing both $\mathbb{F}$ and a root α of p. Since $\mathbb{F}$ is algebraically closed, $\alpha \in \mathbb{F}$ itself; that is, $\mathbb{E} = \mathbb{F}$. But then $x - \alpha \in \mathcal{R}$ is a factor of p, contradicting the assumption that p is irreducible. Since p was an arbitrary factor, f itself 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 $n = k$ variables, $V_F = \emptyset$ implies $I = \mathcal{R}$.

Inductive step: Let $n = k + 1$. Assume $V_F = \emptyset$. 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

$$\begin{aligned} x_1 &= y_1, \\ x_2 &= y_2 + a_2 y_1, \\ &\vdots \\ x_n &= y_n + a_n y_1, \end{aligned}$$

[27] The notation V_F comes from the term **variety** in algebraic geometry.

for some $a_1, \dots, a_n \in \mathbb{F}$. (We make the choice of which $a_1, \dots, a_n$ specific below.) This can be a little confusing, so let's take an example.

Example 11.74. Suppose $f = x_1 + x_2^2 x_3$. We rewrite f as

$$y_1 + (y_2 + a_2 y_1)^2 (y_3 + a_3 y_1)^3 = \\ y_1 + \left(y_2^2 + 2a_2 y_1 y_2 + a_2^2 y_1^2\right)\left(y_3^4 + 3a_3 y_1 y_3^3 + 3a_3^2 y_1^2 y_3^2 + a_3^3 y_1^3 y_3\right).$$

Take note of the forms within the parentheses.

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 terms containing y_1^d as

$$f = c y_1^d + g\left(y_1, \dots, y_n\right)$$

where $c \in \mathbb{F}$ and $\deg_y g < d$. Since $\mathbb{F}$ is infinite, we can find $a_2, \dots, a_n$ such that $c \neq 0$.

Let $\varphi : \mathcal{R} \longrightarrow \mathbb{F}[y_1, \dots, y_n]$ by

$$\varphi\left(f\left(x_1, \dots, x_n\right)\right) = f\left(y_1, y_2 + a_2 y_1, \dots, y_n + a_n y_1\right);$$

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 isomomorphism (Exercise 11.77), so $J = \varphi(I)$ is an ideal of $\mathbb{F}[y_1, \dots, y_n]$. If $V_J \neq \emptyset$, then any $b \in V_J$ can be transformed into an element of V_F (see Exercise 11.78); hence $V_J = \emptyset$ as well.

Now let $\eta : \mathbb{F}[y_1, \dots, y_n] \longrightarrow \mathbb{F}[y_2, \dots, y_n]$ by $\eta(g) = g(0, y_2, \dots, y_n)$.

Example 11.75. For instance, $\eta\left(x_1^3 + x_1 x_3^2 + x_2^2 x_3 + x_4\right) = x_2^2 x_3 + x_4$.

Again, $K = \eta(J)$ is an ideal, though the proof is different (Exercise 11.80). We claim that if $V_K \neq \emptyset$, then likewise $V_J \neq \emptyset$. To see why, let $h \in \eta(\mathbb{F}[y_1, \dots, y_n])$, and suppose $b \in \mathbb{F}^{n-1}$ satisfies $h(b) = 0$. Let g be any element of $\mathbb{F}[y_1, \dots, y_n]$ such that $\eta(g) = h$; then

$$g\left(0, b_1, \dots, b_{n-1}\right) = h\left(b_1, \dots, b_{n-1}\right) = 0,$$

so that we can prepend 0 to any element of V_K and obtain an element of V_J. Since $V_J = \emptyset$, this is impossible, so $V_K = \emptyset$.

Since $V_K = \emptyset$ and $K \subseteq \mathbb{F}[y_2, \dots, y_n]$, the inductive hypothesis finally helps us see that $K = \mathbb{F}[y_2, \dots, y_n]$. In other words, $1 \in K$. Since $K \subset J$ (see Exercise 11.80), $1 \in J$. Since $\varphi(f) \in \mathbb{F}$ if and only if $f \in \mathbb{F}$ (Exercise 11.79), there exists some $f \in \langle F \rangle$ such that $f \in \mathbb{F}$. □

Exercises

Exercise 11.76. Show that the intersection of two radical ideals is also radical.

Exercise 11.77. Show that φ in the proof of Theorem 11.73 is a ring isomomorphism.

Exercise 11.78. Show that in the proof of Theorem 11.73, any $b \in V_{\varphi(F)}$ can be rewritten to obtain an element of V_F. *Hint:* Reverse the translation that defines φ.

Exercise 11.79. Show that in the proof of Theorem 11.73, $\varphi(f) \in \mathbb{F}$ if and only if $f \in \mathbb{F}$.

Exercise 11.80. Show that η in the proof of Theorem 11.73, if J is an ideal of $\mathbb{F}[y_1,\ldots,y_n]$, then $\eta(J)$ is an ideal of $\mathbb{F}[y_2,\ldots,y_n]$. *Hint:* $\mathbb{F}[y_2,\ldots,y_n] \subsetneq \mathbb{F}[y_1,\ldots,y_n]$ and $\eta(J) = J \cap \mathbb{F}[x_2,\ldots,y_n]$ is an ideal of $\mathbb{F}[y_2,\ldots,y_n]$.

11.7: Elementary applications

We now turn our attention to posing, and answering, questions that make Gröbner bases interesting. As in Section 11.6,

- $\mathbb{F}$ is an *algebraically closed* field—that is, all polynomials over $\mathbb{F}$ have their roots in $\mathbb{F}$;
- $\mathcal{R} = \mathbb{F}[x_1, x_2, \ldots, x_n]$ is a polynomial ring;
- $F \subset \mathcal{R}$;
- $V_F \subset \mathbb{F}^n$ 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 $\mathbb{R}$.

Our first question regards membership in an ideal.

Theorem 11.81 (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) $\Longrightarrow$ (B): Assume that $p \in I$. If $p = 0$, then we are done. Otherwise, the definition of a Gröbner basis implies that $\operatorname{lm}(p)$ is top-reducible by some element of G; let r be the result of this top-reduction. By Proposition 11.52, $\operatorname{lm}(r_1) < \operatorname{lm}(p)$. 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 $\operatorname{lm}(p)$ 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

$$\operatorname{lm}(p) > \operatorname{lm}(r_1) > \operatorname{lm}(r_2) > \cdots.$$

By the well-ordering of $\mathbb{M}$, 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) $\Longrightarrow$ (A): Assume that p top-reduces to zero with respect to G. By Lemma 11.56, $p \in I$. □

Now that we have ideal membership, let us return to a topic we considered briefly in Chapter 7. In Exercise 8.24 you showed that

> ...the common roots of $f_1, f_2, \ldots, f_m$ are common roots of all polynomials in the ideal I.

Since $I = \langle G \rangle$, the common roots of $g_1, g_2, \ldots, g_m$ are common roots of all polynomials in I. Thus if we start with a system F, and we want to analyze its polynomials, we can do so 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 11.82. 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 Exercise 11.59, so a similar argument shows that $V_F \supseteq V_G$. We conclude that $V_F = V_G$.

(B) Let g be a nonzero constant polynomial, and observe that $g(\alpha_1, \ldots, \alpha_n) \neq 0$ for any $\alpha \in \mathbb{F}^n$. Thus, if $g \in G$, then $V_G = \emptyset$. By (A), $V_F = V_G = \emptyset$, so F has no common roots if G contains a nonzero constant polynomial.

For the converse, we need the Weak Nullstellensatz, Theorem 11.73 on page 382. If F has no common roots, then $V_F = \emptyset$, and by the Weak Nullstellensatz, $I = \mathcal{R}$. In this case, $1_{\mathcal{R}} \in I$. By definition of a Gröbner basis, there is some $g \in G$ such that $\text{lm}(g) \mid \text{lm}(1_{\mathcal{R}})$. This requires g to be a constant. □

Once we know common solutions exist, we want to know how many there are.

Theorem 11.83. 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 $\text{lm}(g) = x_i^a$.

Remark 11.84. Theorem 11.83 is related to the strong Nullstellensatz.

Proof. We can find $g \in G$ and $\alpha \in \mathbb{N}$ such that $\text{lm}(g) = x_i^a$ for each $i = 1, 2, \ldots, n$ if and only if $\mathcal{R}/I$ is finite; see Figure 11.2. The definition $\mathcal{R}/I$ is independent of any monomial ordering, so

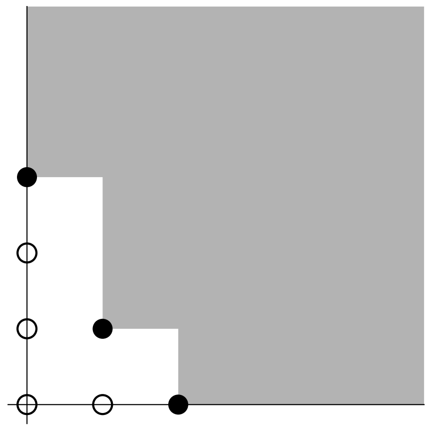


Figure 11.2. This monomial diagram shades the monomials divisible by the leading monomials of a Gröbner basis of I. If $\mathcal{R}/I$ is finite, then we cannot find infinitely many polynomials in $\mathcal{R}$ and outside I. This includes the axes of the monomial diagram, which consist of the monomials 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.

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.

Conversely, assume G has finitely many solutions; call them $\alpha^{(1)}, \ldots, \alpha^{(\ell)} \in \mathbb{F}^n$. Let

$$J = \left\langle x_1 - \alpha_1^{(1)}, \ldots, x_n - \alpha_n^{(1)} \right\rangle \bigcap \cdots \bigcap \left\langle x_1 - \alpha_1^{(\ell)}, \ldots, x_n - \alpha_n^{(\ell)} \right\rangle.$$

Recall that J is an ideal. You will show in the exercises that I and J have the same common solutions; that is, $V_I = V_J$.

For any $f \in \sqrt{I}$, the fact that $\mathcal{R}$ is an integral domain implies that

$$f(\alpha) = 0 \qquad \Longleftrightarrow \qquad f^a(\alpha) = 0 \; \exists a \in \mathbb{N}^+,$$

so $V_I = V_{\sqrt{I}}$. Let K be the ideal of polynomials that vanish on V_I. Notice that $I \subseteq \sqrt{I} \subseteq K$ by definition. We claim that $\sqrt{I} \supseteq K$ as well. Why? Let $p \in K$ be nonzero. Consider the polynomial ring $\mathbb{F}[x_1, \ldots, x_n, y]$ where y is a new variable. Let $A = \langle f_1, \ldots, f_m, 1 - yp \rangle$. Notice that $V_A = \emptyset$, since $f_i = 0$ for each i implies that $p = 0$, but then $1 - yp \neq 0$. By Theorem 11.82, 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 $x_1, \ldots, x_n$ 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 abitrary in K, we have $\sqrt{I} \supseteq K$, as claimed.

We have shown that $K = \sqrt{I}$. Since K is the ideal of polynomials that vanish on V_I, and by construction, $V_{\sqrt{I}} = V_I = V_J$. You will show in the exercises that $J = \sqrt{J}$, so $V_{\sqrt{I}} = V_{\sqrt{J}}$. Hence $\sqrt{I} = \sqrt{J}$. By definition of J,

$$q_j = \prod_{i=1}^{\ell} \left(x_j - a_j^{(i)} \right) \in J$$

for each $j = 1, \ldots, n$. 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 $\text{lm}(q_i) = x_i^{a_i}$ 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 $\text{lm}(g) \mid \text{lm}(q_i)$. In other words, for each i there exists $g \in G$ and $a \in \mathbb{N}$ such that $\text{lm}(g) = x_i^a$. □

Example 11.85. Recall the system from Example 11.54,

$$F = \left(x^2 + y^2 - 4, xy - 1 \right).$$

In Exercise 11.57 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 $x = \text{lm}(g_1)$ and $y^4 = \text{lm}(g_2)$, we know that there are finitely many common roots.

We conclude by pointing in the direction of how to find the common roots of a system.

Theorem 11.86 (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 $h \in \mathbb{F}\left[x_i, x_{i+1}, \ldots, x_n\right]$. 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 $f - g \in \mathbb{F}\left[x_i, x_{i+1}, \ldots, x_n\right]$ 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 $fh \in \mathbb{F}\left[x_i, x_{i+1}, \ldots, x_n\right]$ 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 $g \in \mathbb{F}\left[x_i, x_{i+1}, \ldots, x_n\right]$. 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 11.87. For $i = 1, 2, \ldots, n$ the ideal $\widehat{I} = I \cap \mathbb{F}\left[x_i, x_{i+1}, \ldots, x_n\right]$ is called the i**th elimination ideal of** I.

Theorem 11.86 suggests that to find the common roots of F, we use a lexicographic ordering, then:

- find common roots of $G \cap \mathbb{F}[x_n]$;
- back-substitute to find common roots of $G \cap \mathbb{F}\left[x_{n-1}, x_n\right]$;
- ...
- 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 11.88. We can find the common solutions of the circle and the hyperbola in Figure 11.1 on page 366 using the Gröbner basis computed in Example 387 on page 11.85. 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.

Exercises.

Exercise 11.89. Determine whether $x^6 + x^4 + 5y - 2x + 3xy^2 + xy + 1$ is an element of the ideal $\langle x^2 + 1, xy + 1 \rangle$.

Exercise 11.90. Refer back to Exercise 11.69. How many solutions does this system have? If infinitely many, what is the dimension?

Exercise 11.91. Consider the system

$$\begin{aligned} F = (& \; xyz + xz + 3y + 3, \\ & \; x^2yz^2 + x^2z^2 - y - 1 \;). \end{aligned}$$

Exercise 11.92. Suppose A, B are ideals of $\mathcal{R}$.

(a) Show that $V_{A \cap B} = V(A) \cup V(B)$.

(b) Explain why this shows that for the ideals I and J defined in the proof of Theorem 11.83, $V_I = V_J$.

Chapter 12: Advanced methods of computing Gröbner bases

12.1: The Gebauer-Möller algorithm

Buchberger's algorithm (Algorithm 8 on page 376) allows us to compute Gröbner bases, but it turns out that, without any optimizations, the algorithm is quite inefficient. To explain why this is the case, we make the following observations:

1. The goal of the algorithm is to add polynomials until we have a Gröbner basis. That is, the algorithm is looking for new information.
2. We obtain this new information whenever an S-polynomial does *not* reduce to zero.
3. When an S-polynomial does reduce to zero, we do not add anything. In other words, we have no new information.
4. Thus, reducing an S-polynomial to zero is a wasted computation.

With these observations, we begin to see why the basic Buchberger algorithm is inefficient: it computes every S-polynomial, including those that reduce to zero. Once we have added the last polynomial necessary to satisfy the Gröbner basis property, there is no need to continue. However, at the very least, line 15 of the algorithm generates a larger number of new pairs for P that will create S-polynomials that will reduce to zero. It is also possible that a large number of other pairs will not yet have been considered, and so will also need to be reduced to zero! This prompts us to look for criteria that detect useless computations, and to apply these criteria in such a way as to maximize their usage. Buchberger discovered two additional criteria that do this; this section explores these criteria, then presents a revised Buchberger algorithm that attempts to maximize their effect.

The first criterion arises from an observation that you might have noticed already.

Example 12.1. Let $p = x^2 + 2xy + 3x$ and $q = y^2 + 2x + 1$. Consider any ordering such that $\operatorname{lm}(p) = x^2$ and $\operatorname{lm}(q) = y^2$. Notice that the leading monomials of p and q are *relatively prime*; that is, they have no variables in common.

Now consider the S-polynomial of p and q (we highlight in each step the leading monomial under the grevlex ordering):

$$\begin{aligned} S &= y^2 p - x^2 q \\ &= \mathbf{2xy^3} - 2x^3 + 3xy^2 - x^2. \end{aligned}$$

This S-polynomial top-reduces to zero:

$$\begin{aligned} S - 2xyq &= \left(3xy^2 - 2x^3 - x^2\right) - \left(4x^2y + 2xy\right) \\ &= -\mathbf{2x^3} - 4x^2y + 3xy^2 - x^2 - 2xy; \end{aligned}$$

then

$$\begin{aligned} (S - 2xyq) + 2xp &= \left(-4x^2y + 3xy^2 - x^2 - 2xy\right) + \left(4x^2y + 6x^2\right) \\ &= \mathbf{3xy^2} + 5x^2 - 2xy; \end{aligned}$$

then

$$(S-2xyq+2xp)-3xq = \left(5x^2-2xy\right)-\left(6x^2+3x\right)$$
$$= -\mathbf{x}^2-2xy-3x;$$

finally

$$(S-2xyq+2xp-3xq)+p = (-2xy-3x)+(2xy+3x)$$
$$= 0.\triangle$$

To generalize this beyond the example, observe that we have shown that

$$S+(2x+1)\,p-(2xy+3x)\,q = 0$$

or

$$S = -(2x+1)\,p+(2xy+3x)\,q.$$

If you study p, q, and the polynomials in that last equation, you might notice that the quotients from top-reduction allow us to write:

$$S = -(q-\mathrm{lc}(q)\,\mathrm{lm}(q))\cdot p+(p-\mathrm{lc}(p)\,\mathrm{lm}(p))\cdot q.$$

This is rather difficult to look at, so we will adopt the notation for the trailing terms of p—that is, all the terms of p except the term containing the leading monomial. Rewriting the above equation, we have

$$S = -\mathrm{tts}(q)\cdot p+\mathrm{tts}(q)\cdot p.$$

If this were true in general, it might—*might*—be helpful.

Lemma 12.2 (Buchberger's gcd criterion)**.** Let p and q be two polynomials whose leading monomials are u and v, respectively. If u and v have no common variables, then the S-polynomial of p and q has the form

$$S = -\mathrm{tts}(q)\cdot p+\mathrm{tts}(p)\cdot q.$$

Proof. Since u and v have no common variables, $\mathrm{lcm}(u,v) = uv$. Thus the S-polynomial of p and q is

$$\begin{aligned}S &= \mathrm{lc}(q)\cdot\frac{uv}{u}\cdot(\mathrm{lc}(p)\cdot u+\mathrm{tts}(p))-\mathrm{lc}(p)\cdot\frac{uv}{v}\cdot(\mathrm{lc}(q)\cdot v+\mathrm{tts}(q))\\ &= \mathrm{lc}(q)\cdot v\cdot\mathrm{tts}(p)-\mathrm{lc}(p)\cdot u\cdot\mathrm{tts}(q)\\ &= \mathrm{lc}(q)\cdot v\cdot\mathrm{tts}(p)-\mathrm{lc}(p)\cdot u\cdot\mathrm{tts}(q)+[\mathrm{tts}(p)\cdot\mathrm{tts}(q)-\mathrm{tts}(p)\cdot\mathrm{tts}(q)]\\ &= \mathrm{tts}(p)\cdot[\mathrm{lc}(q)\cdot v+\mathrm{tts}(q)]-\mathrm{tts}(q)\cdot[\mathrm{lc}(p)\cdot u+\mathrm{tts}(p)]\\ &= \mathrm{tts}(p)\cdot q-\mathrm{tts}(q)\cdot p.\end{aligned}$$

□

Lemma 12.2 is not quite enough. Recall Theorem 11.53 on page 370, the characterization theorem of a Gröbner basis:

Theorem 12.3 (Buchberger's characterization)**.** Let $g_1, g_2, \ldots, g_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. The following are equivalent.

(A) $G = (g_1, g_2, \ldots, g_m)$ is a Gröbner basis of the ideal $I = \langle g_1, g_2, \ldots, g_m \rangle$.

(B) For any pair i, j with $1 \leq i < j \leq m$, $\text{Spol}\left(g_i, g_j\right)$ top-reduces to zero with respect to G.

To satisfy Theorem 11.53, we have to show that the S-polynomials top-reduce to zero. However, the proof of Theorem 11.53 used Lemma 11.56:

Lemma 12.4. Let $p, f_1, f_2, \ldots, f_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. Let $F = (f_1, f_2, \ldots, f_m)$. Then (A) implies (B) where

(A) p top-reduces to zero with respect to F.

(B) 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:

(B1) $p = q_1 f_1 + q_2 f_2 + \cdots + q_m f_m$; and

(B2) For each $k = 1, 2, \ldots, m, q_k = 0$ or $\text{lm}(q_k)\,\text{lm}(g_k) \leq \text{lm}(p)$.

We can describe this in the following way, due to Daniel Lazard:

Theorem 12.5 (Lazard's characterization)**.** Let $g_1, g_2, \ldots, g_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. The following are equivalent.

(A) $G = (g_1, g_2, \ldots, g_m)$ is a Gröbner basis of the ideal $I = \langle g_1, g_2, \ldots, g_m \rangle$.

(B) For any pair i, j with $1 \leq i < j \leq m$, $\text{Spol}\left(g_i, g_j\right)$ top-reduces to zero with respect to G.

(C) For any pair i, j with $1 \leq i < j \leq m$, $\text{Spol}\left(g_i, g_j\right)$ has the form

$$\text{Spol}\left(g_i, g_j\right) = q_1 g_1 + q_2 g_2 + \cdots + q_m g_m$$

and for each $k = 1, 2, \ldots, m$, $q_k = 0$ or $\text{lm}(q_k)\,\text{lm}(g_k) < \text{lcm}(\text{lm}(p), \text{lm}(q))$.

Proof. That (A) is equivalent to (B) was the substance of Buchberger's characterization. That (B) implies (C) is a consequence of Lemma 11.56. That (C) implies (A) is implicit in the proof of Buchberger's characterization: you will extract it in Exercise 12.13. □

The form of an S-polynomial described in (C) of Theorem 12.5 is important enough to identify with a special term.

Definition 12.6. Let $G = (g_1, g_2, \ldots, g_m)$. We say that the S-polynomial of g_i and g_j has an **S-representation** $(q_1, \ldots, q_m)$ **with respect to** G if $q_1, q_2, \ldots, q_m \in \mathbb{F}[x_1, \ldots, x_n]$ and (C) of Theorem 12.5 is satisfied.

Lazard's characterization allows us to show that Buchberger's gcd criterion allows us to avoid top-reducing the S-polynomial of any pair whose leading monomials are relatively prime.

Corollary 12.7. Let $g_1, g_2, \ldots, g_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. The following are equivalent.

(A) $G = (g_1, g_2, \ldots, g_m)$ is a Gröbner basis of the ideal $I = \langle g_1, g_2, \ldots, g_m \rangle$.

(B) For any pair (i, j) with $1 \le i < j \le m$, one of the following holds:

- (B1) The leading monomials of g_i and g_j have no common variables.
- (B2) $\operatorname{Spol}(g_i, g_j)$ top-reduces to zero with respect to G.

Proof. Since (A) implies (B2), (A) also implies (B). For the converse, assume (B). Let $\widehat{P}$ be the set of all pairs of P that have an S-representation with respect to G. If (i, j) satisfies (B1), then Buchberger's gcd criterion (Lemma 12.2) implies that

$$\operatorname{Spol}(g_i, g_j) = q_1 g_1 + \cdots + q_m g_m \tag{43}$$

where $q_i = -\operatorname{tts}(g_j)$, $q_j = \operatorname{tts}(g_i)$, and $q_k = 0$ for $k \neq i, j$. Notice that

$$\operatorname{lm}(q_i)\operatorname{lm}(g_i) = \operatorname{lm}(\operatorname{tts}(g_j)) \cdot \operatorname{lm}(g_i) < \operatorname{lm}(g_j)\operatorname{lm}(g_i) = \operatorname{lcm}(\operatorname{lm}(g_i), \operatorname{lm}(g_j)).$$

Thus 43 is an S-representation of $\operatorname{Spol}(g_i, g_j)$, so $(i, j) \in \widehat{P}$. If (i, j) satisfes (B2), then by Lemma 11.56, $(i, j) \in \widehat{P}$ also. Hence every pair (i, j) is in $\widehat{P}$. Lazard's characterization now implies that G is a Gröbner basis of $\langle G \rangle$; that is, (A). □

Although the gcd criterion is clearly useful, it is rare to encounter in practice a pair of polynomials whose leading monomials have no common variables. That said, you have seen such pairs once already, in Exercises 11.57 and 11.67.

We need, therefore, a stronger criterion. The next one is a little harder to discover, so we present it directly.

Lemma 12.8 (Buchberger's lcm criterion)**.** Let p and q be two polynomials whose leading monomials are u and v, respectively. Let f be a polynomial whose leading monomial is t. If t divides $\operatorname{lcm}(u, v)$, then the S-polynomial of p and q has the form

$$S = \frac{\operatorname{lc}(q) \cdot \operatorname{lcm}(u, v)}{\operatorname{lc}(f) \cdot \operatorname{lcm}(t, u)} \cdot \operatorname{Spol}(p, f) + \frac{\operatorname{lc}(p) \cdot \operatorname{lcm}(u, v)}{\operatorname{lc}(f) \cdot \operatorname{lcm}(t, v)} \cdot \operatorname{Spol}(f, q). \tag{44}$$

Proof. First we show that the fractions in equation (44) reduce to monomials. Let x be any variable. Since t divides $\operatorname{lcm}(u, v)$, we know that

$$\deg_x t \le \deg_x \operatorname{lcm}(u, v) = \max(\deg_x u, \deg_x v).$$

(See Exercise 12.12.) Thus

$$\deg_x \operatorname{lcm}(t,u) = \max(\deg_x t, \deg_x u) \leq \max(\deg_x u, \deg_x v) = \deg_x \operatorname{lcm}(u,v).$$

A similar argument shows that

$$\deg_x \operatorname{lcm}(t,v) \leq \deg_x \operatorname{lcm}(u,v).$$

Thus the fractions in (44) reduce to monomials.

It remains to show that (44) is, in fact, consistent. This is routine; working from the right, and writing $S_{a,b}$ for the S-polynomial of a and b and $L_{a,b}$ for $\operatorname{lcm}(a,b)$, we have

$$\begin{aligned}\frac{\operatorname{lc}(q)\cdot L_{u,v}}{\operatorname{lc}(f)\cdot L_{t,u}}\cdot S_{p,f} + \frac{\operatorname{lc}(p)\cdot L_{u,v}}{\operatorname{lc}(f)\cdot L_{t,v}}\cdot S_{f,q} &= \operatorname{lc}(q)\cdot\frac{L_{u,v}}{u}\cdot p\\ &\quad -\frac{\operatorname{lc}(p)\cdot\operatorname{lc}(q)}{\operatorname{lc}(f)}\cdot\frac{L_{u,v}}{t}\cdot f\\ &\quad +\frac{\operatorname{lc}(p)\cdot\operatorname{lc}(q)}{\operatorname{lc}(f)}\cdot\frac{L_{u,v}}{t}\cdot f\\ &\quad -\operatorname{lc}(p)\cdot\frac{L_{u,v}}{v}\cdot q\\ &= S_{p,q}.\end{aligned}$$

□

How does this help us?

Corollary 12.9. Let $g_1, g_2, \ldots, g_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. The following are equivalent.

(A) $G = (g_1, g_2, \ldots, g_m)$ is a Gröbner basis of the ideal $I = \langle g_1, g_2, \ldots, g_m\rangle$.

(B) For any pair i, j with $1 \leq i < j \leq m$, one of the following holds:

- (B1) The leading monomials of g_i and g_j have no common variables.
- (B2) There exists k such that
  - $\operatorname{lm}(g_k)$ divides $\operatorname{lcm}(\operatorname{lm}(g_i), \operatorname{lm}(g_j))$;
  - $\operatorname{Spol}(g_i, g_k)$ has an S-representation with respect to G; and
  - $\operatorname{Spol}(g_k, g_j)$ has an S-representation with respect to G.
- (B3) $\operatorname{Spol}(g_i, g_j)$ top-reduces to zero with respect to G.

Proof. We need merely show that (B2) implies the existence of an S-representation of $\operatorname{Spol}(g_i, g_j)$ with respect to G; Lazard's characterization and the proof of Corollary 12.7 supply the rest. So assume (B2). Choose $h_1, h_2, \ldots, h_m$ such that

$$\operatorname{Spol}(g_i, g_k) = h_1 g_1 + \cdots + h_m g_m$$

and for each $\ell = 1, 2, \ldots, m$ we have $h_\ell = 0$ or

$$\mathrm{lm}(h_\ell)\,\mathrm{lm}(g_\ell) < \mathrm{lcm}(\mathrm{lm}(g_i), \mathrm{lm}(g_k)).$$

Also choose $q_1, q_2, \ldots, q_m$ such that

$$\mathrm{Spol}\left(g_k, g_j\right) = q_1 g_1 + \cdots + q_m g_m$$

and for each $\ell = 1, 2, \ldots, m$ we have $q_\ell = 0$ or

$$\mathrm{lm}(q_\ell)\,\mathrm{lm}(g_\ell) < \mathrm{lcm}\left(\mathrm{lm}(g_k), \mathrm{lm}\left(g_j\right)\right).$$

Write $L_{a,b} = \mathrm{lcm}(\mathrm{lm}(g_a), \mathrm{lm}(g_b))$. Buchberger's lcm criterion tells us that

$$\mathrm{Spol}\left(g_i, g_j\right) = \frac{\mathrm{lc}\left(g_j\right) \cdot L_{i,j}}{\mathrm{lc}(g_k) \cdot L_{i,k}} \cdot \mathrm{Spol}(g_i, g_k) + \frac{\mathrm{lc}(g_i) \cdot L_{i,j}}{\mathrm{lc}(g_k) \cdot L_{j,k}} \cdot \mathrm{Spol}\left(g_k, g_j\right).$$

For $i = 1, 2, \ldots, m$ let

$$H_i = \frac{\mathrm{lc}\left(g_j\right) \cdot L_{i,j}}{\mathrm{lc}(g_k) \cdot L_{i,k}} \cdot h_i + \frac{\mathrm{lc}(g_i) \cdot L_{i,j}}{\mathrm{lc}(g_k) \cdot L_{j,k}} \cdot q_i.$$

Substitution implies that

$$\mathrm{Spol}\left(g_i, g_j\right) = H_1 g_1 + \cdots + H_m g_m. \tag{45}$$

In addition, for each $i = 1, 2, \ldots, m$ we have $H_i = 0$ or

$$\begin{aligned}
\mathrm{lm}(H_i)\,\mathrm{lm}(g_i) &\leq \max\left(\frac{L_{i,j}}{L_{i,k}} \cdot \mathrm{lm}(h_i), \frac{L_{i,j}}{L_{j,k}} \cdot \mathrm{lm}(q_i)\right) \cdot \mathrm{lm}(g_i) \\
&= \max\left(\frac{L_{i,j}}{L_{i,k}} \cdot \mathrm{lm}(h_i)\,\mathrm{lm}(g_i), \frac{L_{i,j}}{L_{j,k}} \cdot \mathrm{lm}(q_i)\,\mathrm{lm}(g_i)\right) \\
&< \max\left(\frac{L_{i,j}}{\cancel{L_{i,k}}} \cdot \cancel{L_{i,k}}, \frac{L_{i,j}}{\cancel{L_{j,k}}} \cdot \cancel{L_{j,k}}\right) \\
&= L_{i,j} \\
&= \mathrm{lcm}\left(\mathrm{lm}(g_i), \mathrm{lm}\left(g_j\right)\right).
\end{aligned}$$

Thus equation (45) is an S-representation of $\mathrm{Spol}\left(g_i, g_j\right)$.

The remainder of the corollary follows as described. □

It is not hard to exploit Corollary 12.9 and modify Buchberger's algorithm in such a way as to take advantage of these criteria. The result is Algorithm 9. The only changes to Buchberger's algorithm are the addition of lines 8, 19, 12, and 13; they ensure that an S-polynomial is computed only if the corresponding pair does not satisfy one of the gcd or lcm criteria.

It is possible to exploit Buchberger's criteria more efficiently, using the Gebauer-Möller algorithm (Algorithms 10 and 11). This implementation attempts to apply Buchberger's criteria as quickly as possible. Thus the first `while` loop of Algorithm 11 eliminates new pairs that satisfy

Algorithm 9. Buchberger's algorithm with Buchberger's criteria

inputs
  $F = (f_1, f_2, \ldots, f_m)$, a list of polynomials in n variables, whose coefficients are from a field $\mathbb{F}$.
outputs
  $G = (g_1, g_2, \ldots, g_M)$, a Gröbner basis of $\langle F \rangle$. Notice $\#G = M$ which might be different from m.
do
  Let $G := F$
  Let $P = \{(f, g) : \forall f, g \in G$ such that $f \neq g\}$
  Let $Done = \{\}$
  repeat while $P \neq \emptyset$
    Choose $(f, g) \in P$
    Remove (f, g) from P
    if $\mathrm{lm}(f)$ and $\mathrm{lm}(g)$ share at least one variable — check gcd criterion
      if not ($\exists p \neq f, g$ such that $\mathrm{lm}(p)$ divides $\mathrm{lcm}(\mathrm{lm}(f), \mathrm{lm}(g))$ **and** $(p, f), (p, g) \in$ *Done*) — check lcm criterion
        Let S be the S-polynomial of f, g
        Let r be the top-reduction of S with respect to G
        if $r \neq 0$
          Replace P by $P \cup \{(h, r) : \forall h \in G\}$
          Append r to G
    Add (f, g) to *Done*
  return G

Algorithm 10. Gebauer-Möller algorithm

inputs
  $F = (f_1, f_2, \ldots, f_m)$, a list of polynomials in n variables, whose coefficients are from a field $\mathbb{F}$.
outputs
  $G = (g_1, g_2, \ldots, g_M)$, a Gröbner basis of $\langle F \rangle$. Notice $\#G = M$ which might be different from m.
do
  Let $G := \{\}$
  Let $P := \{\}$
  repeat while $F \neq \emptyset$
    Let $f \in F$
    Remove f from F
    — See Algorithm 11 for a description of Update
    $G, P :=$ Update (G, P, f)
  repeat while $P \neq \emptyset$
    Pick any $(f, g) \in P$, and remove it
    Let h be the top-reduction of Spol (f, g) with respect to G
    if $h \neq 0$
      $G, P :=$ Update (G, P, h)
  return G

Buchberger's lcm criterion; the second `while` loop eliminates new pairs that satisfy Buchberger's gcd criterion; the third `while` loop eliminates *some* old pairs that satisfy Buchberger's lcm criterion; and the fourth `while` loop removes redundant elements of the basis in a safe way (see Exercise 11.64).

We will not give here a detailed proof that the Gebauer-Möller algorithm terminates correctly. That said, you should be able to see intuitively that it does so, and to fill in the details as well. Think carefully about why it is true. Notice that unlike Buchberger's algorithm, the pseudocode here builds critical pairs using elements (f, g) of G, rather than indices (i, j) of G.

For some time, the Gebauer-Möller algorithm was considered the benchmark by which other algorithms were measured. Many optimizations of the algorithm to compute a Gröbner basis can be applied to the Gebauer-Möller algorithm without lessening the effectiveness of Buchberger's criteria. Nevertheless, the Gebauer-Möller algorithm continues to reduce a large number of S-polynomials to zero.

Exercises.

Exercise 12.10. In Exercise 11.57 on page 374 you computed the Gröbner basis for the system

$$F = \left(x^2 + y^2 - 4, xy - 1\right)$$

in the lexicographic ordering using Algorithm 8 on page 376. Review your work on that problem, and identify which pairs (i, j) would not generate an S-polynomial if you had used Algorithm 9 on the preceding page instead.

Algorithm 11. Update the Gebauer-Möller pairs

inputs
  G_{old}, a list of polynomials in n variables, whose coefficients are from a field $\mathbb{F}$.
  P_{old}, a set of critical pairs of elements of G_{old}
  a non-zero polynomial p in $\left\langle G_{\text{old}} \right\rangle$
outputs
  G_{new}, a (possibly different) basis of $\left\langle G_{\text{old}} \right\rangle$.
  P_{old}, a set of critical pairs of G_{new}
do
  Let $C := \left\{ (p, g) : g \in G_{\text{old}} \right\}$
  — C is the set of all pairs of the new polynomial p with an older element of the basis
  Let $D := \{\}$
  — D is formed by pruning pairs of C using Buchberger's lcm criterion
  — We do not yet check Buchberger's gcd criterion because with the original input there may be some cases of the lcm criterion that are eliminated by the gcd criterion
  repeat while $C \neq \emptyset$
    Pick any $(p, g) \in C$, and remove it
    if $\operatorname{lm}(p)$ and $\operatorname{lm}(g)$ share no variables **or** no $(p, h) \in C \cup D$ satisfies $\operatorname{lcm}(\operatorname{lm}(p), \operatorname{lm}(h)) \mid \operatorname{lcm}(\operatorname{lm}(p), \operatorname{lm}(g))$
      Add (p, g) to D
  Let $E := \emptyset$
  — E is the result of pruning pairs of D using Buchberger's gcd criterion
  repeat while $D \neq \emptyset$
    Pick any $(p, g) \in D$, and remove it
    if $\operatorname{lm}(p)$ and $\operatorname{lm}(g)$ share at least one variable
      $E := E \cup (p, g)$
  — P_{int} is the result of pruning pairs of P_{old} using Buchberger's lcm criterion
  Let $P_{\text{int}} := \{\}$
  repeat while $P_{\text{old}} \neq \emptyset$
    Pick $(f, g) \in P_{\text{old}}$, and remove it
    if $\operatorname{lm}(p)$ does not divide $\operatorname{lcm}(\operatorname{lm}(f), \operatorname{lm}(g))$ **or** $\operatorname{lcm}(\operatorname{lm}(p), \operatorname{lm}(h)) = \operatorname{lcm}(\operatorname{lm}(f), \operatorname{lm}(g))$ for $h \in \{f, g\}$
      Add (f, g) to P_{int}
  — Add new pairs to surviving pre-existing pairs
  $P_{\text{new}} := P_{\text{int}} \cup E$
  — Prune redundant elements of the basis, but not their critical pairs
  Let $G_{\text{new}} := \{\}$
  repeat while $G_{\text{old}} \neq \emptyset$
    Pick any $g \in G_{\text{old}}$, and remove it
    if $\operatorname{lm}(p)$ does not divide $\operatorname{lm}(g)$
      Add g to G_{new}
  Add p to G_{new}
  return $G_{\text{new}}, P_{\text{new}}$

Exercise 12.11. Use the Gebauer-Möller algorithm to compute the Gröbner basis for the system

$$F = \left(x^2 + y^2 - 4, xy - 1\right).$$

Indicate clearly the values of the sets C, D, E, G_{new}, and P_{new} after each `while` loop in Algorithm 11 on the previous page.

Exercise 12.12. Let t, u be two monomials, and x any variable. Show that

$$\deg_x \operatorname{lcm}(t,u) = \max\left(\deg_x t, \deg_x u\right).$$

Exercise 12.13. Study the proof of Buchberger's characterization, and extract from it a proof that (C) implies (A) in Theorem 12.5.

12.2: The F4 algorithm

An interesting development of the last ten years in the computation of Gröbner bases has revolved around changing the point of view to that of linear algebra. Recall from Exercise 11.71 that for any polynomial system we can construct a matrix whose triangularization simulates the computation of S-polynomials and top-reduction involved in the computation of a Gröbner basis. However, a naïve implementation of this approach is worse than Buchberger's method:

- every possible multiple of each polynomial appears as a row of a matrix;
- many rows do not correspond to S-polynomials, and so are useless for triangularization;
- as with Buchberger's algorithm, where most of the S-polynomials are not necessary to compute the basis, most of the rows that are not useless for triangularization are useless for computing the Gröbner basis!

Jean-Charles Faugère devised two algorithms that use the ordered Macaulay matrix to compute a Gröbner basis: F4 and F5. We focus on F4, as F5 requires more discussion than, quite frankly, I'm willing to put into these notes at this time.

Remark 12.14. F4 does not strictly require homogeneous polynomials, but for the sake of simplicity we stick with homogeneous polynomials, so as to introduce d-Gröbner bases.

Rather than build the entire ordered Macaulay matrix for any particular degree, Faugère first applied the principle of building only those rows that correspond to S-polynomials. Thus, given the homogeneous input

$$F = \left(x^2 + y^2 - 4h^2, xy - h^2\right),$$

the usual degree-3 ordered Macaulay matrix would be

$$\left(\begin{array}{ccccccccccc}
x^3 & x^2y & xy^2 & y^3 & x^2h & xyh & y^2h & xh^2 & yh^2 & h^3 & \\
1 & & 1 & & & & & -4 & & & xf_1 \\
 & 1 & & 1 & & & & & -4 & & yf_1 \\
 & & & & 1 & & 1 & & & -4 & hf_1 \\
 & 1 & & & & & & -1 & & & xf_2 \\
 & & 1 & & & & & & -1 & & yf_2 \\
 & & & & & 1 & & & & -1 & hf_2
\end{array}\right).$$

However, only two rows of the matrix correspond to an S-polynomial: yf_1 and xf_2. For top-reduction we might need other rows: non-zero entries of rows yf_1 and xf_2 involve the monomials

$$y^3, xh^2, \text{ and } yh^2;$$

but no other row might reduce those monomials: that is, there is no top-reduction possible. We could, therefore, triangularize just as easily if we built the matrix

$$\left(\begin{array}{cccccccccc|c} x^3 & x^2y & xy^2 & y^3 & x^2h & xyh & y^2h & xh^2 & yh^2 & h^3 & \\ & 1 & & 1 & & & & & -4 & & yf_1 \\ & 1 & & & & & & -1 & & & xf_2 \end{array}\right).$$

Triangularizing it results in

$$\left(\begin{array}{cccccccccc|c} x^3 & x^2y & xy^2 & y^3 & x^2h & xyh & y^2h & xh^2 & yh^2 & h^3 & \\ & 1 & & 1 & & & & & -4 & & yf_1 \\ & & & 1 & & & & 1 & 4 & & yf_1 - xf_2 \end{array}\right),$$

whose corresponds to the S-polynomial $yf_1 - xf_2$. We have thus generated a new polynomial,

$$f_3 = y^3 + xh^2 + 4yh^2.$$

Proceeding to degree four, there are two possible S-polynomials: for (f_1, f_3) and for (f_2, f_3). We can discard (f_1, f_3) thanks to Buchberger's gcd criterion, but not (f_2, f_3). Building the S-polynomial for (f_2, f_3) would require us to subtract the polynomials $y^2 f_2$ and xf_3. The non-leading monomial of $y^2 f_2$ is y^2h^2, and no leading monomial divides that, but the non-leading monomials of xf_3 are x^2h^2 and xyh^2, both of which are divisible by $h^2 f_1$ and $h^2 f_2$. The non-leading monomials of $h^2 f_1$ are y^2h^2, for which we have already introduced a row, and h^4, which no leading monomial divides; likewise, the non-leading monomial of $h^2 f_2$ is h^4.

We have now identified all the polynomials that *might* be necessary in the top-reduction of the S-polynomial for (f_2, f_3):

$$y^2 f_2, xf_3, h^2 f_1, \text{ and } h^2 f_2.$$

We build the matrix using rows that correspond to these polynomials, resulting in

$$\left(\begin{array}{ccccc|c} xy^3 & x^2h^2 & xyh^2 & y^2h^2 & h^2 & \\ 1 & & & -1 & & y^2 f_2 \\ 1 & 1 & 4 & & & xf_3 \\ & 1 & & 1 & -4 & h^2 f_1 \\ & & 1 & & -1 & h^2 f_2 \end{array}\right).$$

Triangularizing this matrix results in (step-by-step)

$$\left(\begin{array}{ccccc|c} xy^3 & x^2h^2 & xyh^2 & y^2h^2 & h^2 & \\ 1 & & & -1 & & y^2 f_2 \\ & -1 & -4 & -1 & & y^2 f_2 - xf_3 \\ & 1 & & 1 & -4 & h^2 f_1 \\ & & 1 & & -1 & h^2 f_2 \end{array}\right);$$

$$\begin{pmatrix} xy^3 & x^2h^2 & xyh^2 & y^2h^2 & h^2 & \\ 1 & & & -1 & & y^2f_2 \\ & & -4 & 0 & -4 & y^2f_2-xf_3+h^2f_1 \\ & 1 & & 1 & -4 & h^2f_1 \\ & & 1 & & -1 & h^2f_2 \end{pmatrix};$$

and finally

$$\begin{pmatrix} xy^3 & x^2h^2 & xyh^2 & y^2h^2 & h^2 & \\ 1 & & & -1 & & y^2f_2 \\ & & & & 0 & y^2f_2-xf_3+h^2f_1+4h^2f_2 \\ & 1 & & 1 & -4 & h^2f_1 \\ & & 1 & & -1 & h^2f_2 \end{pmatrix}.$$

This corresponds to the fact that the S-polynomial of f_2 and f_3 reduces to zero: and we can now stop, as there are no more critical pairs to consider.

Aside from building a matrix, the F4 algorithm thus modifies Buchberger's algorithm (with the additional criteria, Algorithm 9 in the two following ways:

- rather than choose *a* critical pair in line 10, one chooses *all* critical pairs of minimal degree; and
- all the S-polynomials of this minimal degree are computed simultaneously, allowing us to reduce them "all at once".

In addition, the move to a matrix means that linear algebra techniques for triangularizing a matrix can be applied, although the need to preserve the monomial ordering implies that column swaps are forbidden. Algorithm 12 describes a simplified F4 algorithm. The approach outlined has an important advantage that we have not yet explained.

Definition 12.15. Let G be a list of homogeneous polynomials, let $d \in \mathbb{N}^+$, and let I be a an ideal of homogeneous polynomials. We say that G is a **d-Gröbner basis of I** if $\langle G \rangle = I$ and for every $a \leq d$, every S-polynomial of degree a top-reduces to zero with respect to G.

Example 12.16. In the example given at the beginning of this section,

$$G = \left(x^2+y^2-4h^2, xy-h^2, y^3+xh^2+4yh^2\right)$$

is a 3-Gröbner basis. ⊿

A Gröbner basis G is always a d-Gröbner basis for all $d \in \mathbb{N}$. However, not every d-Gröbner basis is a Gröbner basis.

Example 12.17. Let $G = \left(x^2+h^2, xy+h^2\right)$. The S-polynomial of g_1 and g_2 is the degree 3 polynomial

$$S_{12} = yh^2 - xh^2,$$

which does not top-reduce. Let

$$G_3 = \left(x^2+h^2, xy+h^2, xh^2-yh^2\right);$$

the critical pairs of G_3 are

Algorithm 12. A simplified F4 that implements Buchberger's algorithm with Buchberger's criteria

inputs
  $F = (f_1, f_2, \ldots, f_m)$, a list of homogeneous polynomials in n variables, whose coefficients are from a field $\mathbb{F}$.
outputs
  $G = (g_1, g_2, \ldots, g_M)$, a Gröbner basis of $\langle F \rangle$. Notice $\#G = M$ which might be different from m.
do
  Let $G := F$
  Let $P := \{(f, g) : \forall f, g \in G \text{ such that } f \neq g\}$
  Let $Done := \{\}$
  Let $d := 1$
  repeat while $P \neq \emptyset$
    Let P_d be the list of all pairs $(i, j) \in P$ that generate S-polynomials of degree d
    Replace P with $P \backslash P_d$
    Denote $L_{p,q} := \text{lcm}(\text{lm}(p), \text{lm}(q))$
    Let Q be the subset of P_d such that $(f, g) \in Q$ implies that:
    – $\text{lm}(f)$ and $\text{lm}(g)$ share at least one variable; and
    – **not** ($\exists p \in G \backslash \{f, g\}$ such that $\text{lm}(p)$ divides $L_{f,g}$ **and** $(f, p), (g, p) \in Done$)
    Let $R := \left\{ tp, uq : (p, q) \in Q \text{ and } t = L_{p,q} / \text{lm}(p), u = L_{p,q} / \text{lm}(q) \right\}$
    Let S be the set of all tp where t is a monomial, $p \in G$, and $t \cdot \text{lm}(p)$ is a non-leading monomial of some $q \in R \cup S$
    Let M be the submatrix of the ordered Macaulay matrix of F corresponding to the elements of $R \cup S$
    Let N be any triangularization of M that does not swap columns
    Let G_{new} be the set of polynomials that correspond to rows of N that changed from M
    for $p \in G_{\text{new}}$
      Replace P by $P \cup \{(h, p) : \forall h \in G\}$
      Add p to G
    Add (f, g) to $Done$
    Increase d by 1
  return G

- (g_1, g_2), whose S-polynomial now reduces to zero;
- (g_1, g_3), which generates an S-polynomial of degree 4 (the lcm of the leading monomials is x^2h^2); and
- (g_2, g_3), which also generates an S-polynomial of degree 4 (the lcm of the leading monomials is xyh^2).

All degree 3 S-polynomials reduce to zero, so G_3 is a 3-Gröbner basis.

However, G_3 is *not* a Gröbner basis, because the pair (g_2, g_3) generates an S-polynomial of degree 4 that does *not* top-reduce to zero:

$$S_{23} = h^4 + y^2h^2.$$

Enlarging the basis to

$$G_4 = \left(x^2 + h^2, xy + h^2, xh^2 - yh^2, y^2h^2 + h^4\right)$$

gives us a 4-Gröbner basis, which is also the Gröbner basis of G. ⊿

One useful property of d-Gröbner bases is that we can answer some question that require Gröbner bases by short-circuiting the computation of a Gröbner basis, settling instead for a d-Gröbner basis of sufficiently high degree. For our concluding theorem, we revisit the Ideal Membership Problem, discussed in Theorem 11.81.

Theorem 12.18. Let $\mathcal{R}$ be a polynomial ring, let $p \in \mathcal{R}$ be a homogeneous polynomial of degree d, and let I be a homogeneous ideal of $\mathcal{R}$. The following are equivalent.

(A) $p \in I$.

(B) p top-reduces to zero with respect to a d-Gröbner G_d of I.

Proof. That (A) implies (B): If $p = 0$, then we are done; otherwise, let $p_0 = p$ and G_d be a d-Gröbner basis of I. Since $p_0 = p \in I$, there exist $h_1, \dots, h_m \in \mathcal{R}$ such that

$$p_0 = h_1 g_1 + \cdots + h_m g_m.$$

Moreover, since p is of degree d, we can say that for every i such that the degree of g_i is larger than d, $h_i = 0$.

If there exists $i \in \{1, 2, \dots, m\}$ such that $\operatorname{lm}(g_i)$ divides $\operatorname{lm}(p_0)$, then we are done. Otherwise, the equality implies that some leading terms on the right hand side cancel; that is, there exists at least one pair (i, j) such that $\operatorname{lm}(h_i)\operatorname{lm}(g_i) = \operatorname{lm}\left(h_j\right)\operatorname{lm}\left(g_j\right) > \operatorname{lm}(p_0)$. This cancellation is a multiple of the S-polynomial of g_i and g_j; by definition of a d-Gröbner basis, this S-polynomial top-reduces to zero, so we can replace

$$\operatorname{lc}(h_i)\operatorname{lm}(h_i)\, g_i + \operatorname{lc}\left(h_j\right)\operatorname{lm}\left(h_j\right) g_j = q_1 g_1 + \cdots + q_m g_m$$

such that each $k = 1, 2, \dots, m$ satisfies

$$\operatorname{lm}(q_k)\operatorname{lm}(g_k) < \operatorname{lm}(h_i)\operatorname{lm}(g_i).$$

We can repeat this process any time that $\operatorname{lm}(h_i)\operatorname{lm}(g_i) > \operatorname{lm}(p_0)$. The well-ordering of the monomials implies that eventually we must arrive at a representation

$$p_0 = h_1 g_1 + \cdots + h_m g_m$$

where at least one k satisfies $\operatorname{lm}(p_0) = \operatorname{lm}(h_k)\operatorname{lm}(g_k)$. This says that $\operatorname{lm}(g_k)$ divides $\operatorname{lm}(p_0)$, so we can top-reduce p_0 by g_k to a polynomial p_1. Note that $\operatorname{lm}(p_1) < \operatorname{lm}(p_0)$.

By construction, $p_1 \in I$ also, and applying the same argument to p_1 as we did to p_0 implies that it also top-reduces by some element of G_d to an element $p_2 \in I$ where $\operatorname{lm}(p_2) < \operatorname{lm}(p_1)$. Iterating this observation, we have

$$\operatorname{lm}(p_0) > \operatorname{lm}(p_1) > \cdots$$

and the well-ordering of the monomials implies that this chain cannot continue indefinitely. Hence it must stop, but since G_d is a d-Gröbner basis, it does not stop with a non-zero polynomial. That is, p top-reduces to zero with respect to G.

That (B) implies (A): Since p top-reduces to zero with respect to G_d, Lemma 11.56 implies that $p \in I$. □

Exercises.

Exercise 12.19. Use the simplified F4 algorithm given here to compute a d-Gröbner bases for $\left\langle x^2y - z^2h, xz^2 - y^2h, yz^3 - x^2h^2 \right\rangle$ for $d \le 6$. Use the grevlex term ordering with $x > y > z > h$.

Exercise 12.20. Given a non-homogeneous polynomial system F, describe how you could use the simplified F4 to compute a non-homogeneous Gröbner basis of $\langle F \rangle$.

12.3: Signature-based algorithms to compute a Gröbner basis

This section is inspired by recent advances in the computation of Gröbner basis, including my own recent work. As with F4, the original algorithm in this area was devised by Faugère, and is named F5 [Fau02]. A few years later, Christian Eder and I published an article that showed how one could improve F5 somewhat [EP10]; the following year, the GGV algorithm was published [GGV10], and Alberto Arri asked me to help him finish an article that sought to generalize some notions of F5 [AP11]. Seeing the similarities between Arri's algorithm and GGV, I teamed up with Christian Eder again to author a paper that lies behind this work [EP11]. The algorithm as presented here is intermediate between Arri's algorithm (which is quite general) and the one we present there (which is specialized).

In its full generality, the idea relies on a generalization of vector spaces.

Definition 12.21. Let R be a ring. A **module** M **over** R satisfies the following properties. Let $r, s \in R$ and $x, y, z \in M$. Then

- M is an additive group;
- $r\,x \in M$;
- $r\,(x + y) = r\,x + r\,y$;
- $(r + s)\,x = r\,x + s\,x$;
- $1_R x = x$.

We will not in fact use modules extensively, but the reader should be aware of the connection. In any case, it is possible to describe it at a level suitable for the intended audience of these notes (namely, me and any of my students whose research might lead in this direction). We adopt the following notation:

- $\mathcal{R} = \mathbb{F}[x_1, \ldots, x_n]$ is a polynomial ring;
- $\mathbb{M}$ the set of monomials of $\mathcal{R}$;
- $<$ a monomial ordering;
- $f_1, \ldots, f_m \in \mathcal{R}$;
- $F = (f_1, \ldots, f_m)$;
- $I = \langle F \rangle$.

Definition 12.22. Let $p \in I$ and $h_1, \ldots, h_m \in \mathcal{R}$. We say that $H = (h_1, \ldots, h_m)$ is an F**-representation of** p if

$$p = h_1 f_1 + \cdots + h_m f_m.$$

If, in addition, $p = 0$, then we say that H is a **syzygy** of F.[a]

[a]It can be shown that the set of all syzygies is a module over $\mathcal{R}$, called the **module of syzygies**.

Example 12.23. Suppose $F = \left(x^2 + y^2 - 4, xy - 1\right)$. Recall that $p = x + y^3 - 4y \in \langle F \rangle$, since

$$x + y^3 - 4y = y f_1 - x f_2.$$

In this case, (y, x) is not an S-representation of p, since $y \operatorname{lm}(f_1) = x^2 y = \operatorname{lcm}\left(x^2, xy\right)$. However, it is an F-representation.

On the other hand,

$$0 = f_2 f_1 - f_1 f_2 = (xy - 1) f_1 - \left(x^2 + y^2 - 4\right) f_2,$$

so $(f_2, -f_1)$ is an F-representation of 0; that is, $(f_2, -f_1)$ is a syzygy. ⊿

Keep in mind that an F-representation is almost never an S-representation (Definition 12.6). However, an F-representation exists for any element of I, even if F is not a Gröbner basis. An S-representation does *not* exist for at least one S-polynomial when F is not a Gröbner basis.

We now generalize the notion of a leading monomial of a *polynomial* to a leading monomial of an F-representation.

Definition 12.24. Write $\mathbf{F}_i$ for the m-tuple whose entries are all zero *except* for entry i, which is 1.[a] Given an F-representation H of some $p \in I$, whose rightmost nonzero entry occurs in position i, we say that $\operatorname{lm}(h_i)\mathbf{F}_i$ is a **leading monomial of** H, and write $\operatorname{lm}(H) = \operatorname{lm}(h_i)\mathbf{F}_i$. Let

$$\mathbb{S} = \{\operatorname{lm}(H) : h_1 f_1 + \cdots + h_m f_m \in I\};$$

that is, $\mathbb{S}$ is the set of all possible leading monomials of an F-representation.

[a] In the parlance of modules, $\{\mathbf{F}_1, \ldots, \mathbf{F}_m\}$ is the set of **canonical generators** of the free $\mathcal{R}$-module $\mathcal{R}^m$.

Example 12.25. Recall F from Example 12.23. We have $\mathbf{F}_1 = (1, 0)$ and $\mathbf{F}_2 = (0, 1)$. The leading monomial of $(y, 0)$ is $y\mathbf{F}_1$. The leading monomial of (y, x) is $x\mathbf{F}_2 = (0, x)$. The leading monomial of $(f_2, -f_1)$ is $\operatorname{lm}(-f_1)\mathbf{F}_2 = (0, x^2)$. ⊿

Once we have leading monomials of F-representations, it is natural to generalize the ordering of monomials of $\mathbb{M}$ to an ordering of leading monomials.

Definition 12.26. Define a relation $\prec$ on $\mathbb{S}$ as follows: we say that $t\mathbf{F}_i \prec u\mathbf{F}_j$ if

- $i < j$, or
- $i = j$ and $t < u$.

Lemma 12.27. $\prec$ is a well-ordering of $\mathbb{S}$.

Proof. Let $S \subseteq \mathbb{S}$. Since $<$ is a well-ordering of $\mathbb{N}^+$, there exists a minimal $i \in \mathbb{N}^+$ such that $t\mathbf{F}_i \in S$ for any $t \in \mathbb{M}$. Let $T = \{t : t\mathbf{F}_i \in S\}$; notice that $T \subseteq \mathbb{M}$. Since $<$ is a well-ordering of $\mathbb{M}$, T has a least element t. By definition, $t\mathbf{F}_i \preceq u\mathbf{F}_j$ for any $u\mathbf{F}_j \in S$.\ □

Corollary 12.28. Let $p \in I$ and $\mathcal{H}$ the set of all possible F-representations of p. Let

$$S = \{\operatorname{lm}(H) : H \in \mathcal{H}\}.$$

Then S has a smallest element with respect to $\prec$.

Proof. $S \subset \mathbb{S}$, which is well ordered by $\prec$. □

Definition 12.29. We call the smallest element of S the **signature of** p, denoted by $\operatorname{sig}(p)$.

Now let's consider how the ordering behaves on some useful operations with F-representations. First, some notation.

Definition 12.30. If $t \in \mathbb{M}$ and $H, H' \in \mathcal{R}^m$, we define

$$tH = (th_1, \ldots, th_m) \text{ and } H + H' = \left(h_1 + h_1', \ldots, h_m + h_m'\right).$$

In addition, we define $t\operatorname{sig}(p) = t\,(u\mathbf{F}_i) = (tu)\,\mathbf{F}_i$.

Lemma 12.31. Let $p, q \in I$, H an F-representation of f, H' an F-representation of q, and $t, u \in \mathbb{M}$. Suppose $\tau = \operatorname{lm}(H)$ and $\upsilon = \operatorname{lm}(H')$. Each of the following holds.

(A) tH is an F-representation of tp;

(B) $\operatorname{sig}(tp) \preceq t\tau = \operatorname{lm}(tH)$;

(C) if $t\tau \prec u\upsilon$, then $\operatorname{lm}(tH \pm uH') = u\upsilon$;

(D) if $t\tau = u\upsilon$, then there exists $c \in \mathbb{F}$ such that $\operatorname{lm}(ctH + uH') \prec t\tau$.

(E) if $\operatorname{Spol}(p,q) = atp - buq$ for appropriate $a, b \in \mathbb{F}$, then $\operatorname{sig}(\operatorname{Spol}(p,q)) \preceq \max(t\tau, u\upsilon)$;

(F) if H'' is an F-representation of p and $\operatorname{lm}(H'') \prec \operatorname{lm}(H)$, then there exists a syzygy $Z \in \mathcal{R}^m$ such that
  - $H'' + Z = H$ and
  - $\operatorname{lm}(Z) = \operatorname{lm}(H)$;

  and

(G) if H'' is an F-representation of p such that $\operatorname{lm}(H'') = \operatorname{sig}(p)$, then $\operatorname{lm}(H'') < \operatorname{lm}(H)$ if and only if there exists a nonzero syzygy Z such that $H'' + Z = H$ and $\operatorname{lm}(Z) = \operatorname{lm}(H)$.

It is important to note that even if $t\tau = \operatorname{lm}(tH)$, that does not imply that $t\tau = \operatorname{sig}(tp)$ even if $\tau = \operatorname{sig}(p)$.

Proof. (A) Since H is an F-representation of p, we know that $p = \sum h_i f_i$. By the distributive and associative properties, $tp = t\sum h_i f_i = \sum[(th_i) f_i]$. Hence tH is an F-representation of tp.

(B) The definition of a signature implies that $\operatorname{sig}(tp) \preceq t\tau$. That $t\tau = \operatorname{lm}(tH)$ is a consequence of (A).

(C) Assume $t\tau \prec u\upsilon$. Write $\tau = v\mathbf{F}_i$ and $\upsilon = w\mathbf{F}_j$. By definition of the ordering $\prec$, either $i < j$ or $i = j$ and $\operatorname{lm}(h_i) \lessdot \operatorname{lm}(h_j')$. Either way, $\operatorname{lm}(tH \pm uH')$ is $u\operatorname{lm}(h_j')\mathbf{F}_j = u\upsilon$.

(D) Assume $t\tau = u\upsilon$. Let $a = \operatorname{lc}(H)$, $b = \operatorname{lc}(H')$, and $c = b/a$. Then $\operatorname{lm}(tH) = t\tau = u\upsilon = \operatorname{lm}(uH')$, and $c\operatorname{lc}(tH) = \operatorname{lc}(uH')$. Together, these imply that the leading monomials of ctH and uH' cancel in the subtraction $ctH - uH'$. Hence $\operatorname{lm}(ctH - uH') \prec t\tau$.

(E) follows from (B), (C), and (D).

(F) Assume that H'' is an F-representation of p and $\operatorname{lm}(H'') \prec \operatorname{lm}(H)$. Then

$$0 = p - p = \sum h_i f_i - \sum h_i'' f_i = \sum\left(h_i - h_i''\right) f_i.$$

Let $Z = \left(h_1 - h_1'', \ldots, h_m - h_m''\right)$. By definition, Z is a sygyzy. In addition, $\operatorname{lm}(H'') \prec \operatorname{lm}(H)$ and (C) imply that $\operatorname{lm}(Z) = \operatorname{lm}(H)$.

(G) One direction follows from (F); the other is routine. □

We saw in previous sections that if we considered critical pairs by ascending lcm, we were able to take advantage of previous computations to reduce substantially the amount of work needed to compute a Gröbner basis. It turns out that we can likewise reduce the amount of work substantially if we proceed by ascending signature. This will depend on an important fact.

Definition 12.32. Let $p \in I$, and H an S-representation of p. If $\operatorname{lm}(h_k)\operatorname{sig}(g_k) \preceq \operatorname{lm}(p)$ for each k, then we say that H is a **signature-compatible representation** of p, or a **sig-representation** for short.

Lemma 12.33. Let $\tau \in \mathbb{S}$, and suppose that every S-polynomial of $G \subsetneq I$ with signature smaller than τ has a sig-representation. Let $p, q \in I$ and $t, u \in \mathbb{M}$ such that $u\operatorname{sig}(q) \preceq t\operatorname{sig}(p) = \tau$, $\operatorname{Spol}(p,q) = \operatorname{lc}(q)\,t\,p - \operatorname{lc}(p)\,u\,q$. Suppose that one of the following holds:
(A) $\operatorname{sig}(tp) = \operatorname{sig}(uq)$; or
(B) $t\operatorname{sig}(p) \neq \operatorname{sig}(\operatorname{Spol}(p,q))$.
Then $\operatorname{Spol}(p,q)$ has a sig-representation.

Proof. (A) Let H and H' be F-representations of p and q (respectively) such that $\operatorname{lm}(H) = \operatorname{sig}(p)$ and $\operatorname{lm}(H') = \operatorname{sig}(q)$. By Lemma 12.31(D), there exists $c \in \mathbb{F}$ satisfying the property $\operatorname{lm}(ctH + uH') \prec \operatorname{lm}(ctH)$; in other words, $\operatorname{sig}(ctp + uq) \prec \operatorname{sig}(tp)$. Let H'' be an F-representation of $ctp + uq$ such that $\operatorname{lm}(H'') = \operatorname{sig}(ctp + uq)$; by hypothesis, all top-cancellations of the sum

$$h''_1 f_1 + \cdots + h''_m f_m$$

have sig-representations. The fact that the top-cancellations have signature smaller than τ implies that we can rewrite these top-cancellations repeatedly as long as they exist. Each rewriting leads to smaller leading monomials, and signatures no larger than those of the top-cancellations. Since the monomial ordering is a well ordering, we cannot rewrite these top-cancellations indefinitely. Hence this process of rewriting eventually terminates with a sig-representation of $ctp + uq$. If $ctp + uq$ is a scalar multiple of $\operatorname{Spol}(p,q)$, then we are done; notice $\operatorname{sig}(\operatorname{Spol}(p,q)) \prec t\operatorname{sig}(p)$.

If $ctp + uq$ is not a scalar multiple of $\operatorname{Spol}(p,q)$, then $\operatorname{sig}(\operatorname{Spol}(p,q)) = t\operatorname{sig}(p) = \tau$. Consider the fact that $c\operatorname{Spol}(p,q) = \operatorname{lc}(q)(ctp + uq) - (c\operatorname{lc}(p) + \operatorname{lc}(q))\,uq$. One summand on the right hand side is a scalar multiple of q, so it has a sig-representation no larger than $u\operatorname{sig}(q) \prec \tau$. The previous paragraph showed that $ctp + uq$ has a sig-representation smaller than τ. The sum of these sig-representations is also a sig-representation no larger than τ. Hence the left hand side has an F-representation H''' with $\operatorname{lm}(H''') \preceq \tau$.

(B) By part (A), we know that if $u\operatorname{sig}(q) = t\operatorname{sig}(p)$, then $\operatorname{Spol}(p,q)$ has a sig-representation. Assume therefore that $u\operatorname{sig}(q) \prec t\operatorname{sig}(p) = \tau$. Since $t\operatorname{sig}(p) \neq \operatorname{sig}(tp)$, Lemma 12.31 implies that $\operatorname{sig}(tp) \prec t\operatorname{sig}(p) = \tau$. Likewise, $\operatorname{sig}(uq) \preceq u\operatorname{sig}(q) \prec \tau$, so

$$\operatorname{sig}(\operatorname{Spol}(p,q)) \preceq \max(\operatorname{sig}(tp), \operatorname{sig}(uq)) \prec \tau.$$

The hypothesis implies that $\operatorname{Spol}(p,q)$ has a sig-representation. □

To compute a Gröbner basis using signatures, we have to reduce polynomials in such a way that we have a good estimate of the signature. To do this, we cannot allow a reduction $r - tg$ if $\operatorname{sig}(r) \preceq$

Algorithm 13. Signature-based algorithm to compute a Gröbner basis

inputs
  $F \subsetneq \mathcal{R}$
outputs
  $G \subsetneq \mathcal{R}$, a Gröbner basis of $\langle F \rangle$
do
  Let $G = \{(\mathbf{F}_i, f_i)\}_{i=1}^m$
  Let $\mathcal{S} = \left\{\operatorname{lm}\left(f_j\right)\mathbf{F}_i : 1 \le j < i\right\}_{i=1}^m$
  Let $P = \{(\upsilon, p, q) : (\sigma, p), (\tau, q) \in G$ and υ is the expected signature of $\operatorname{Spol}(p,q)\}$
  repeat while $P \neq \emptyset$
    Select any $(\sigma, p, q) \in P$ such that τ is minimal
    Let $S = \operatorname{Spol}(p,q)$
    if $\exists (\tau, g) \in G$, $t \in \mathbb{M}$ such that $t\tau = \sigma$ and $t \operatorname{lm}(g) \le \operatorname{lm}(S)$
      if σ is not a monomial multiuple of any $\tau \in \mathcal{S}$
        Top-reduce S to r over G in such a way that $\operatorname{sig}(r) \preceq \sigma$
        if $r \neq 0$ and r is not sig-redundant to G
          for $(\tau, g) \in G$
            if $g \neq 0$ and $t\sigma \neq u\tau$, where t and u are the monomials needed to construct $\operatorname{Spol}(r,g)$
              Add (υ, r, g) to P, where υ is the expected signature of $\operatorname{Spol}(r,g)$
        else
          Add σ to $\mathcal{S}$
  return $\{g : (\tau, g \in G)$ and $g \neq 0\}$

$t \operatorname{sig}(g)$; otherwise, we have no way to recuperate $\operatorname{sig}(r)$. Thus, a signature-based algorithm to compute a Gröbner basis can sometimes add redundant polynomials to the basis. Recall that termination of the Gröbner basis algorithms studied so far follows from the property of those algorithms that r was not added to a basis if it was redundant. This presents us with a problem. The solution looks like a natural generalization, but it took several years before someone devised it.

Definition 12.34. Let $G = \{(\tau_k, g_k)\}_{k=1}^{\ell}$ for some $\ell \in \mathbb{N}^+$, $g_k \in I$, and $\tau_k \in \mathbb{S}$, satisfying $\tau_k = \operatorname{sig}(g_k)$ for each k. We say that (σ, r) is **signature-redundant**, or **sig-redundant**, if there exists $(\tau, g) \in G$ such that $\tau \mid \sigma$ and $\operatorname{lm}(g) \mid \operatorname{lm}(r)$.

Algorithm 13 uses these ideas to compute a Gröbner basis of an ideal.

Theorem 12.35. Algorithm 13 terminates correctly.

Proof. To see why the algorithm terminates, let $\mathbb{M}'$ be the set of variables in $x_1, \ldots, x_n$ *and* $x_{n+1}, \ldots, x_n$, and define two functions

- $\psi : \mathbb{M} \to \mathbb{M}'$ by $\psi\left(x_1^{\alpha_1} \cdots x_n^{\alpha_n}\right) = x_{n+1}^{\alpha_1} \cdots x_{2n}^{\alpha_n}$, and
- $\varphi : G \to \left(\mathbb{M}'\right)^m$ by $\varphi(u\mathbf{F}_i, g) = (u \cdot \psi(\operatorname{lm}(g)))\,\mathbf{F}_i$.

Notice that the variable shift imposed by ψ implies that $\varphi(u\mathbf{F}_i, g)$ divides $\varphi(u'\mathbf{F}_i, g')$ if and only if $u \mid u'$ and $\mathrm{lm}(g) \mid \mathrm{lm}(g')$. This is true if and only if $(u'\mathbf{F}_i, g')$ is sig-redundant with $(u\mathbf{F}_i, g)$, which contradicts how the algorithm works! Let J be the ideal generated by $\varphi(G)$ in $(\mathbb{M}')^m$. As we just saw, adding elements to G implies that we expand some component of J. However, Proposition 8.33 and Definition 8.31 imply that the components of J can expand only finitely many times. Hence the algorithm can add only finitely many elements to G, which implies that it terminates.

For correctness, we need to show that the output satisfies the criteria of Lemma 12.33. Lines 12, 13, and 17 are the only ones that could cause a problem.

For line 12, suppose $(\tau, g) \in G$ and $t \in \mathbb{M}$ satisfy $t\tau = \sigma$ and $t\,\mathrm{lm}(g) \leq \mathrm{lm}(\mathrm{Spol}(p,q))$. Let $H, H' \in \mathcal{R}^m$ be F-representations of $S = \mathrm{Spol}(p,q)$ and g, respectively. We can choose H and H' such that $\mathrm{lm}(H) = \sigma$ and $\mathrm{lm}(H') = \tau$. By Lemma 12.31, there exists $c \in \mathbb{F}$ such that $\mathrm{sig}(cS + tg) \prec \sigma$. On the other hand, $t\,\mathrm{lm}(g) < \mathrm{lm}(S)$ implies that $\mathrm{lm}(cS + tg) = \mathrm{lm}(S)$. The algorithm proceeds by ascending signature, so $cS + tg$ has a sig-representation H'' (over G, not F). Thus,

$$cS + tg = \sum h''_k g_k \quad \Longrightarrow \quad S = -c^{-1} t g + \sum \left(c^{-1} h''_k\right) g_k$$

Every monomial of H'' is, by definition of a sig-representation, smaller than $\mathrm{lm}(cS + tg) = \mathrm{lm}(S)$. In addition, $\mathrm{sig}(tg) \preceq \sigma$, and $\mathrm{sig}\left(h''_k g_k\right) \prec \sigma$ for each k. Define

$$\widehat{h}_k = \begin{cases} c^{-1}\widehat{h}_k, & g \neq g_k; \\ c^{-1}\left(\widehat{h}_k - t\right), & g = g_k. \end{cases}$$

Then $\widehat{H} = \left(\widehat{h}_1, \ldots, \widehat{h}_{\#G}\right)$ is a sig-representation of $\mathrm{Spol}(p,q)$.

For line 13, inspection of the algorithm shows that either $\tau = \mathrm{lm}\left(f_j\right)\mathbf{F}_i$ for some $j < i$, or $\left(\tau, \widehat{p}, \widehat{q}\right)$ was selected from P, and the algorithm reduced $\mathrm{Spol}\left(\widehat{p}, \widehat{q}\right)$ to zero. In the first case, suppose $\sigma = u\mathbf{F}_i$. Let $H \in \mathcal{R}^m$ an F-representation of $\mathrm{Spol}(p,q)$ such that $\mathrm{lm}(H) = \sigma$, and $t \in \mathbb{M}$ such that $t\,\mathrm{lm}\left(f_j\right)\mathbf{F}_i = \sigma$. Let $Z \in \mathcal{R}^m$ such that

$$z_k = \begin{cases} f_i, & k = j; \\ -f_j, & k = i; \\ 0, & \text{otherwise.} \end{cases}$$

Observe that Z is a syzygy, since $\sum z_\ell f_\ell = f_i f_j + \left(-f_j\right) f_i = 0$. In addition, $j < i$ so $\mathrm{lm}(Z) = \mathrm{lm}\left(f_j\right)\mathbf{F}_i$. Thus

$$\mathrm{Spol}(p,q) = \mathrm{Spol}(p,q) + t\sum z_\ell f_\ell = \mathrm{Spol}(p,q) + \sum (t z_\ell) f_\ell.$$

The right hand side has signature smaller than σ (look at $H + Z$), so the left hand side must, as well. By Lemma 12.33, $\mathrm{Spol}(p,q)$ has a sig-representation.

In the second case, we have some $\left(\tau, \widehat{p}, \widehat{q}\right)$ selected from P whose S-polynomial reduced to zero, and some $t \in \mathbb{M}$ such that $t\tau = \sigma$. Since the reduction respects the signature τ, there exists a sig-representation H of $\mathrm{Spol}(\widehat{p},\widehat{q})$; that is,

$$\mathrm{Spol}(\widehat{p},\widehat{q}) = \sum h_\ell g_\ell$$

and $\mathrm{sig}(h_\ell g_\ell) \prec \tau$ for each $\ell = 1,\ldots,\#G$. Thus $\mathrm{Spol}(\widehat{p},\widehat{q}) - \sum h_\ell g_\ell = 0$. This implies the existence of a syzygy $Z \in \mathcal{R}^m$ such that $\mathrm{lm}(Z) = \mathrm{sig}\left(\mathrm{Spol}(\widehat{p},\widehat{q}) - \sum h_\ell g_\ell\right) = \tau$. Thus

$$\mathrm{Spol}(p,q) = \mathrm{Spol}(p,q) - t\sum z_\ell f_\ell = \mathrm{Spol}(p,q) - \sum (t z_\ell) f_\ell,$$

but the right side clearly has signature smaller than σ, so the left hand side must, as well. By Lemma 12.33, $\mathrm{Spol}(p,q)$ has a sig-representation.[28]

For line 17, let $(\tau, g) \in G$ such that $\tau \mid \sigma$ and $\mathrm{lm}(g) \mid \mathrm{lm}(r)$. Let $t, u \in \mathbb{M}$ such that $t\tau = \sigma$ and $u\,\mathrm{lm}(g) = \mathrm{lm}(r)$. If $u < t$, then $u\tau \prec \sigma$, which contradicts the hypothesis that (σ, r) completed a reduction that respects the signature. Otherwise, $t \le u$, which implies that $t\tau = \sigma$ and $t\,\mathrm{lm}(g) \le \mathrm{lm}(r) \le \mathrm{lm}(\mathrm{Spol}(p,q))$. In this case, an argument similar to the one for line 12 applies. □

[28] Notice that both cases for line 13 use syzygies. This is why $\mathcal{S}$ has that name: $\mathcal{S}$ for syzygy.

Appendices

Where can I go from here?

Advanced group theory

Galois theory [Rot98], representation theory, other topics [AF05, Rot06]

Advanced ring theory

Commutative algebra [GPS05], algebraic geometry [CLO97, CLO98], non-commutative algebra

Applications

General [LP98], coding theory, cryptography, computational algebra [vzGG99]

Preface — in the postface

Apologia pro libro suo

Any reader who has wandered to this page is doubtless wondering, *Why on earth is the "preface" in the back?"* That's easy to answer: putting in its "proper" place (the front) distracts from the material. Algebra fascinates me so much that I like to dive right into it; putting these remarks before the material seems like telling a child he can have a snack, then giving him a slice of cake covered in mushrooms. Earlier versions of these notes felt like that.

How did this get started, anyway? A two-semester sequence on modern algebra typically introduces students to the fundamental ideas in group and ring theory. Lots of textbooks do a good job of that, and I always recommend one or more to my classes.

However, many books spend the first chapter (or more!) on material that doesn't feel like algebra. Most every algebra text I've seen spends the first 50–100 pages *on material that is not algebra*. Authors have very, very good reasons for that; for example, the very concrete problems of number theory can motivate certain algebraic ideas. In my experience, however, students' unfamiliarity with number theory means they expend a lot of time and energy learning that, rather than *algebra.* Don't get me wrong: number theory is a wonderful topic; I teach it regularly myself — in a class on, you know, *number theory*.

These notes don't discard that material; they move it to a place that feels more appropriate. For example, the Euclidean algorithm appears in a chapter on number theory, while congruence rears its head in a chapter on cosets. There's still a chapter on "Foundations", but it reviews material students are supposed to have seen in previous classes, so I assign a few exercises from that chapter, and start in Chapter 1 on the first day, telling students to see me if they need help with it. So far, this hasn't been a problem.

In addition, most texbooks seem to target students with a strong mathematical background. Like many instructors these days, I encounter many students with a weaker background in mathematical reasoning; in many cases, the first time they've had to write a proof is in college, and algebra is the first course where they must write a proof without knowing in advance its form ("by induction", "by contradiction", etc.). In some cases, the course attracts *graduate* students who have never written a proof. These students arrive with enthusiasm, and find the material fascinating. Some may possess the great combination of talent, enthusiasm, and preparation, but most lack at least one of those three.

Desiring a mix of simplicity and utility, I decided to set out some notes that would throw the class into algebraic problems and ideas as soon as possible. As it happens, another interest of mine seems to have helped. Typically, an algebra text starts with groups, on account of their simplicity. Another option is to start with rings, on account of the familiarity of their operations. I've tried to marry the best of both worlds by starting with monoids, which are both simple and familiar.[29] An added bonus is that one can introduce notions such as direct products, isomorphism, ideals, the Ascending Chain Condition, and even Hilbert Functions, in a fast, intuitive way that is not at all superficial.

As the notes diverged more and more from the textbooks I was using, they became more or less an unofficial textbook.

[29] To some extent, I owe the idea of starting with monoids to a superb graduate-level text, [KR00].

Overview

These notes have two major parts: in one, we focus on an algebraic structure called a *group*; in the other, we focus on a special kind of group, a *ring*. They correspond roughly to a two-semester course in algebra. I have also thrown in a few things that don't typically appear in undergraduate books, but that fascinate me, and are unquestionably algebraic.

When I teach this material, I try to cover Chapters 1–5 in the first semester. Since a rigorous approach requires *some* sort of introduction, those chapters follow a review of basic ideas students should have seen before — but only to set a foundation for what is to come. That material is really a prerequisite for this material, anyway, so I typically assign a few problems necessary for later, and move directly to Chapter 1.

Monoids are not a popular way to start an algebra course, so much of that chapter is optional. However, a *brief* glance at monoids allows us to introduce prized ideas that we develop in much more depth with groups and rings. With monoids, we preserve a context with which students are far more familiar: natural numbers, matrices, and monomials.

Chapter 6, on number theory, serves as a bridge between the two main parts. Many books on algebra start with this material, I've pushed as much as I felt possible after group theory, so that we can view number theory as an *application* of group theory. Ideally, the chapter on the RSA algorithm would provide a nice "bang" to end the first semester, but I haven't managed that in years, and even then it was too rushed. *Tempus fugit,* and all that.

In the second semester, we definitely cover Chapters 6 through 8, along with at least one of the later chapters. I include Chapter 12 for students who want to pursue a research project, and need an introduction that builds on what came before. As of this writing, some of those chapters still need major debugging, so don't take anything you read there too seriously.

It is not easy to jump around these notes. Not much of the material can be omitted. Within each chapter, many examples are used and reused; this applies to exercises, as well. I do try to concentrate on a few important examples, re-examining them in the light of each new topic. One consequence is that the presentation of groups depends on the introduction to monoids, and the presentation of rings depends on a thorough consideration of groups, which in turn depends on at least some of the material on monoids. On the other hand, most of the material on monoids can be postponed until after groups. In the first semester, I usually omit groups of automorphisms (Section 4.4).

To the student

Most people find advanced algebra quite difficult. There is no shame in that; I find it difficult, too; I sometimes joke that I earned a Ph.D. only because I was too dumb to quit. I'm pretty sure most algebraists find it difficult; the difference is that *we love it.* No other branch of mathematics has quite the same appeal.

I want you to learn algebra, and to see why its ideas have excited not just me, but thousands of others, most of whom are much, much smarter than me. My experiences teaching this class motivate the following remarks.

How to succeed at algebra

You have to keep in mind certain laws of success in algebra, which I'm pretty sure apply not only to me, but to everyone out there.

1. You won't "get it" right away.

One of the big shocks to students who study algebra is that they can't apply the same strategy that they have applied successfully in other mathematical courses. In many undergraduate textbooks, each section introduces some property or technique, *maybe* explains why it works, then illustrates an application of the property, asking you to repeat it on some problems. At most, they ask you to adapt the method used to apply the property.

Algebra isn't like that. The problems almost always require you to use some properties to derive or explain *other properties!* That requires a new style of solving problems, one where you develop the method of solution. Typically, this takes the form of a proof, a short explanation as to why some property is true. You're not really used to that, and you may even have thought that you were studying mathematics precisely to *escape* writing! Sorry!

2. Anything worth doing requires effort and time.

Modern technology can execute in moments tasks that were once impossible, such as speaking across the ocean. Books and films tend to portray the process of discovery and invention as if it were the nearly instantaneous result of a genius' intuition, but the reality is far different. The people who developed these technologies did not do so with a snap of their fingers! They spent years, if not their entire lives, trying to solve difficult and important problems. They tried lots of different methods, and failed with most before finding something that worked.

The same is true with mathematics. For example, the material covered in Chapter 9 is commonly called "Galois Theory". It's entirely possible that the reason it isn't called "Ruffini Theory" is that Paolo Ruffini, who discovered many of its principles, couldn't get anyone to take his notions seriously. None of the leading minds of his day would talk with him about it, which meant that he couldn't see easily the flaws in his work, let alone correct, develop, and deepen them. For that matter, the accomplishments of Evariste Galois were not recognized until decades after he stayed up all night before a duel to to write down ideas that had fermented in his mind. Eventually, they would inebriate the world with understanding.

Algebra is worth spending time on. Don't try to do it on the cheap, devoting only a few spare moments here and there. It will take more than 30 minutes per week to succeed with the homework problems in this class. It may well take more than 30 minutes *per problem*! Don't let that intimidate you.

3. You actually have to know the definitions.

I strongly suggest writing every definition down on a notecard, and creating flashcards to quiz yourself on basic definitions.

Most people no longer seem to think the meanings of words matter. This manifests itself even in mathematics, where students who walk around with A's in high school and college Calculus can't tell you the definition of a limit or a derivative! How do you earn a top score without learning what the fundamental ideas mean?

By its nature, you can't even understand the basic problems in algebra unless you know the meaning of the terms. I can talk myself blue in the face while helping students, but a student who can't state the definition of the technical words used in the problem will not understand the problem, let alone how to find the solution.

4. Don't be afraid to make a fool of yourself.

As I wrote above, I succeeded only because I was too dumb to quit. In a world that glorifies selfesteem just a little too much, many students have to work up courage to ask a question, because they think asking questions makes them look stupid. In my experience, most questions students ask are very good questions. Often enough, I have to correct something stupid I said. Honestly, the only "dumb" questions in this class are the ones where someone asks me what a word means. That's a *definition;* if you can't be bothered to look it up, I can't be bothered to tell you. (Most of the time. Sometimes, I'll tell you anyway — especially if it's been a while since we talked about it.) All other questions pertinent to this material really are fair game.

So, ask away. With any luck, you'll end up embarrassing *me*.

Ways these notes try to help you succeed

I have tried to present a large number of "concrete" examples. Some examples are more important than others, and you will notice that I return frequently to a few. I am not unique in emphasizing these examples; most textbooks in algebra emphasize at least some of them.

Spend time familiarizing yourself with these examples. Students often make the mistake of thinking that the purpose of the examples is to show them how to solve the exercises. While that may be true in a textbook on, say, calculus, linear algebra, or differential equations, it can be a fatal assumption in non-linear algebra. Here, the purpose of the examples is to illustrate the objects and ideas that you have to understand in order to solve to the exercises. I suspect these notes are unusual in dedicating an entire section to the roots of unity, but if not, that only proves how important this example is.

I could say the same about the exercises. Even if an exercise isn't assigned, and you choose not to solve it, familiarize yourself with the statement of the exercise. A significant proportion of the exercises build on examples or even exercises that appear earlier.

An approach I've used that seems uncommon, if not unique, is the presence of fill-in-the-blank exercises. I've designed these with two goals in mind. First, most algebra students are overwhelmed by the rush of ideas and objects — and have very little experience solving theoretical problems, where the "answer" is already given, and the "method of solution" is what they must produce! So, I've taken some of the problems that seem to present students with more difficulty, and sketched a proof where nearly every statement lacks either a phrase or a justification; students need merely fill the hole. Second, even when students have a basic understanding of the proof of a statement, they typically write a very poor proof. The fill-in-the-blank problems are meant to illustrate what a correct proof looks like — although, in my attempt to leave no stone unturned, they may seem pedantic.

A few acknowledgements

These notes are inspired from some of my favorite algebra texts: [AF05, CLO97, PHLA88, KR00, vzGG99, Lau03, LP98, Rot06, Rot98]. (Believe it or not, that is *not* a comprehensive list.) They started out as notes to parallel [Lau03], but have since taken on a life of their own, and are now quite different. I have tried to cite a source when I followed a particular approach.

Thanks to the students who found typos, including (in no particular order, and not including everyone) Jonathan Yarber, Kyle Fortenberry, Lisa Palchak, Ashley Sanders, Sedrick Jefferson, Shaina Barber, Blake Watkins, Kris Katterjohn, Taylor Kilman, Eric Gustaffson, Patrick Lambert, and others. Special thanks go to my graduate student Miao Yu, who endured the first drafts of Chapters 7, 8, and 11.

Rogério Brito of Universidade de São Paolo made several helpful comments, found some nasty errors, and suggested some of the exercises.

I have been lucky to have had great algebra professors; in chronological order:

- Vanessa Job at Marymount University;
- Adrian Riskin at Northern Arizona University;
- and at North Carolina State University:
    - Kwangil Koh,
    - Hoon Hong,
    - Erich Kaltofen,
    - Michael Singer, and
    - Agnes Szanto.

Boneheaded innovations of mine that looked good at the time but turned out bad in practice were entirely my idea. *This is not a peer-reviewed text*, which is why you have a supplementary text in the bookstore.

The following software helped prepare these notes:

- Sage 3.*x* and later [Ste08];
- Lyx [Lyx] (and therefore LaTeX [Lam86, Grä04] (and therefore TeX [Knu84])), along with the packages
    - `cc-beamer` [Pip07],
    - `hyperref` [RO08],
    - AMS-LaTeX[Soc02],
    - `mathdesign` [Pic06],
    - `thmtools`, and
    - `algorithms` (modified slightly from the version released 2006/06/02) [Bri]; and
- Inkscape [Bah08].

I've likely forgotten some other non-trivial resources that I used. Let me know if another citation belongs here.

My wife forebore a number of late nights at the office (or at home) as I worked on these.

Ad maiorem Dei gloriam.

Hints to Exercises

Hints to Chapter 0

Exercise 0.25: Use the definition of an equivalence relation to answer the question, and use Example 0.15 as a guide.

Exercise 0.26: Consider two cases, depending on whether a is positive or negative. In each case, use the definition of $\leq$.

Exercise 0.27: Apply the definitions of these orderings. Seriously, that's all the hint you need.

Exercise 0.29: Apply the definitions of these orderings. Seriously, that's all the hint you need.

Exercise 0.30(b): Try to show that $a - b = 0$.

Exercise 0.31: Use the definition of $<$. Seriously, that's all the hint you need.

Exercise 0.33: To show that a unique element of a set possesses a property, pick two elements of the set that share the property, and show that they are equal. In this case, let m, n be two smallest elements of S. Since m is a smallest element of S, what do you know about m and n? Likewise, since n is a smallest element of S, what do you know about m and n? Proceed from there.

Exercise 0.34: For the linear ordering, it's okay to argue using reasoning similar to our examples of linear and non-linear orderings. For the well ordering, you'll have to take an arbitrary subset of $\mathbb{Z}$ (*why?!?* be sure you understand that) and consider a few cases. (I'd consider these cases: if it has zero, it if it has negative numbers but not zero, and if it has only positive numbers).

Exercise 0.35: For (a), it helps to assume the denominators are positive—after all, we can always rewrite a fraction with positive denominators. For (b), exploit the meaning of "compatible."

Exercise 0.36: This problem asks if a particular function *exists*. Existence problems are usually evil, because to prove that an object exists, you typically have to construct it, and finding the right object out of an infinite cloud of objects can leave one feeling downright nebulous.

Still, we might as well try. After all, that's why they pay us the big bucks, right?[30] We will consider each property of an equivalence relation in turn, and see if we can build a function that satisfies that property. After finding a function that satisfies one property, we'll try to show that the same function satisfies the other properties. If we find that too hard to do, we should derive some insight that will help us explain that no such property exists.

Now, this problem is doubly evil in that it asks you to work rather nebulously with functions, which you're not used to dealing with in the abstract, and to look at them as relations, which you're probably not used to at all. So I'll be nice here and clarify that what we're saying is that the function f can be viewed as an expression $\underset{f}{\sim}$ where $a \underset{f}{\sim} b$ means $f(a) = b$.

To start with, then, a function that is an equivalence relation must satisfy the reflexive property. Guess what? It's pretty easy to find a function f on a set A that satisfies the reflexive property for each $a \in A$. If you find the same function I'm thinking of, then you shouldn't have too much trouble showing that it also satisfies the symmetric and transitive properties.

And then you'd be done.

[30]Wait. What "big bucks"?

Exercise 0.39: For (a), use the definition. Seriously, use it. For (b), use substitution.

Exercise 0.40: The strategy used to show uniqueness in the hint for Exercise 0.33 can be used here, but you have to adapt it to a different property: inverse of a function, rather than minimal element of a set.

Exercise 0.47: Don't forget that remainders cannot be negative.

Exercise 0.49: Use Exercise 0.32(c).

Exercise ??: Pick an example $n, d \in \mathbb{Z}$ and look at the resulting M. Which value of q gives you an element of $\mathbb{N}$ as well? If $n \in \mathbb{N}$ then you can easily identify such q. If $n < 0$ it takes a little more work.

Exercise 0.88: You need to show that for any vectors $\mathbf{x}$ and $\mathbf{y}$, you have $f_A(\mathbf{x}+\mathbf{y}) = f_A(\mathbf{x}) + f_A(\mathbf{y})$. Showing an example won't cut it here; you have to do it for *generic* A, $\mathbf{x}$, and $\mathbf{y}$. Even the dimension n has to be generic. Using summation notation ($\sum$) may look unappealing, but it makes your life incredibly simpler; try getting used to it now, when it's relatively easy.

Exercise 0.89: Use the definition of a kernel. Seriously, use it. For any matrix A, the definition says that $\mathbf{x} \in \ker A$ if and only if $A\mathbf{x} = 0$. Both the matrices in this exercise are 3×3 matrix, so $\mathbf{x}$ is a column vector with three entries. Call these three entries x_1, x_2, and x_3. Then solve the matrix equations $M\mathbf{x} = 0$ and $N\mathbf{x} = 0$.

Exercise 0.90: Use the definitions involved here. Seriously, use them. Keep in mind that you don't know the size of the matrices.

Exercise 0.93: Do *not* analyze the entries of A and A^{-1}. Instead, think about what the determinant of AA^{-1} should be.

Exercise 0.95(b): You have to use a *generic* orthogonal matrix. Be clever in your reliance on properties of determinants.

Hints to Chapter 1

Exercise 1.22: Don't confuse what you have to do here, or what the elements are. You have to work with elements of $P(S)$; these are *subsets of* S. So, if I choose $X \in P(S)$, I know that $X \subseteq S$. Notice that I use capital letters for X, even though it is an element of $P(S)$, precisely because it is a set. This isn't something you *have* to do, strictly speaking, but you might find it helpful to select an element of X to prove at least one of the properties of a monoid, and it looks more natural to select $x \in X$ than to select $a \in x$, even if this latter x is a set.

Exercise 1.27: To show closure, you have to explain how we know that the set specified in the definition of lcm has a minimum.

Exercise 1.39: By Definition 0.6, you have to show that

- for any monoid M, $M \cong M$ (reflexive);
- for any two monoids M and N, if $M \cong N$, then also $N \cong M$ (symmetric); and
- for any three monoids M, N, and P, if $M \cong N$ and $N \cong P$, then $M \cong P$ (transitive).

In the first case, you have to find an isomorphism $f : M \longrightarrow M$. In the second, you have to assume that there exist isomorphisms $f : M \longrightarrow N$, then show that there exists an isomorphism $f : N \longrightarrow M$.

Hints to Chapter 2

Exercise 2.26: Use substitution and the inverse property.

Exercise 2.27: Work with arbitrary elements of $\mathbb{R}^{2\times 2}$. The structure of such elements is

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad \text{where } a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{R}.$$

Exercise 2.30: Be as specific as possible: don't just say, "functions that have an inverse."

Exercise 2.31: At least one such monoid appears in this section's exercises.

Exercise 2.32: Most of this has already been done in Exercise 1.51.

Exercise 2.33: You probably did this in linear algebra, or saw it done. Work with arbitrary elements of $\mathbb{R}^{m\times m}$, which have the structure

$$A = \left(a_{i,j}\right)_{i=1\ldots m, j=1\ldots m}.$$

Exercises 2.35 and 2.36, and similar: It's quite alright to use inheritance from a superset to explain why some of these sets satisfy certain properties of a group.

Exercise 2.20: There's no trickery here; just use the properties of a group.

Exercise 2.37:

(a) Try $m = 2$, and find two invertible matrices A, B such that $(AB)\left(A^{-1}B^{-1}\right) \neq I_2$.

(b) Use the associative property to help simplify the expression $(ab)\left(b^{-1}a^{-1}\right)$.

Exercise 2.38:

(b) Use the fact that $(F \circ F)(P) = F(F(P))$ to show that $(F \circ F)(P) = I(P)$, and repeat with the other functions.

(c) One of closure, identity, or inverse fails. Which? At least one property comes from the fact that F_S is a set for any set S. Do you remember what F_S is? If not, see Example 1.10.

(d) Add elements to G that are lacking, until all the properties are now satisfied.

(e) A clever argument would avoid a brute force computation.

Exercise ??: To rewrite products so that ρ never precedes φ, use Theorem 2.41. To show that D_3 satisfies the properties of a group, you may use the fact that D_3 is a subset of $\mathrm{GL}(2)$, the multiplicative group of 2×2 invertible matrices. Thus D_3 "inherits" certain properties of $\mathrm{GL}(2)$, but which ones?

Exercise 2.52:

(a) You may use the property that $|P-Q|^2 = |P|^2 + |Q|^2 - 2P \cdot Q$, where $|X|$ indicates the distance of X from the origin, and $|X-Y|$ indicates the distance between X and Y.

(c) Use the hint from part (a), along with the result in part (a), to show that the distance between the vectors is zero. Also use the property of dot products that for any vector X, $X \cdot X = |X|^2$. Don't use part (b).

Exercise 2.53: Let $P = (p_1, p_2)$ be an arbitrary point in $\mathbb{R}^2$, and assume that ρ leaves it stationary. You can represent P by a vector. The equation $\rho \cdot \vec{P} = \vec{P}$ gives you a system of two linear equations in two variables; you can solve this system for p_1 and p_2.

Exercise 2.54: Repeat what you did in Exercise 2.53. This time the system of linear equations will have infinitely many solutions. You know from linear algebra that in $\mathbb{R}^2$ this describes a line. Solve one of the equations for p_2 to obtain the equation of this line.

Exercise 2.66: Use the product notation as we did.

Exercise 2.68: Look back at Exercise 2.18 on page 81.

Exercise 2.70: Use denominators to show that no matter what you choose for $x \in \mathbb{Q}$, there is some $y \in \mathbb{Q}$ such that $y \notin \langle x \rangle$.

Exercise 2.71: One possibility is to exploit $\det(AB) = \det A \cdot \det B$. It helps to know that $\mathbb{R}$ is not cyclic (which may not be obvious, but should make sense from Exercise 2.70).

Exercise 2.81: Just use basic facts of exponents, along with the property that $\omega^{10} = 1$. *Don't* use trigonometry; that would be gross overkill, and would probably discourage you.

Hints to Chapter 3

Exercise 3.15: Start with the smallest possible subgroup, then add elements one at a time. Don't forget the adjective "proper" subgroup.

Exercise 3.18(c): Look at what L has in common with H from Example 3.10.

Exercise 3.21: Use Exercise 2.70 on page 101.

Exercise 3.23: Look at Exercise 3.20 on page 115.

Exercise 3.60: Use Corollary 3.55 on page 125.

Exercise 3.61: Exercises 2.68 on page 101 and 3.60 would lead you in one direction.

Exercise 3.77: If nothing else comes to mind, you should be able to compute the Cayley table of Ω_8/Ω_2, and that should tell you.

Exercise 3.84: Theorem 3.67 tells us that the subgroup of an abelian group is normal. If you can show that $D_m(\mathbb{R})$ is abelian, then you are finished.

Exercise 3.87: First you must show that $H \subseteq N_G(H)$. Then you must show that $H < N_G(H)$. Finally you must show that $H \triangleleft N_G(H)$. Make sure that you separate these steps and justify each one carefully!

Exercise 3.88: List the two left cosets, then the two right cosets. What does a partition mean? Given that, what sets must be equal?

Exercise 3.89(d): The "hard" way is to show that for all $g \in G$, $g[G,G] = [G,G]g$. This requires you to show that two sets are equal. Any element of $[G,G]$ is the product $[x_1,y_1][x_2,y_2]\cdots[x_n,y_n]$ for some $x_i, y_i \in G$. At some point, you will have to show that $g[x,y] = [w,z]g$ for some $w, x \in G$. This is an existence proof, and it suffices to construct w and z that satisfy the equation. To construct them, think about conjugation.

An “easier” way uses the result of Exercise 3.82. Write $G' = [G, G]$, and show that $gG'g^{-1} = G'$ for any $g \in G$; the Exercise shows that G' is normal. Exercise 2.40 should help you see why $gG'g^{-1} \subseteq G'$. To show the reverse direction, let $g \in G$, and show why we can find $x, y \in G$ such that any $g' \in G'$ has the form $g^{-1}\left[x_1^g, y_1^g\right]\left[x_2^g, y_2^g\right]\cdots\left[x_n^g, y_n^g\right]g$, so $gG'g^{-1} \supseteq G'$.

Exercise 3.101: Use Lemma 3.37 on page 121. Remember that the operation here is addition, so you have to write cosets additively, not multiplicatively.

Hints to Chapter 4

Exercise 4.15(b): Generalize the isomorphism of (a).

Exercise 4.22: For a homomorphism function, think about a projection. Or, adapt the equation that describes the points on L.

Exercise 4.23: Since it’s a corollary to Theorem 4.9, you should use that theorem.

Exercise 4.24: Consider what happens to two elements of the kernel, and all the criteria required to be an isomorphism, not just a homomorphism.

Exercise 4.25(a): Use the Subgroup Theorem along with the properties of a homomorphism.

Exercise 4.28: Use induction on the positive powers of g; use a theorem for the nonpositive powers of g.

Exercise 4.29(b): Let $G = \mathbb{Z}_2$ and $H = D_3$; find a homomorphism from G to H.

Exercise 4.30: Recall that

$$f(A) = \{y \in H : f(x) = y \ \exists x \in A\},$$

and use the Subgroup Theorem.

Exercise 4.31(b): See the last part of Exercise 4.29.

Exercise 4.49(a): Consider $f : \mathbb{R}^2 \to \mathbb{R}$ by $f(a) = b$ where the point $a = (a_1, a_2)$ lies on the line $y = x + b$.

Exercise 4.50(b): You already know the answer from Exercise 3.80 on page 132; find a homomorphism f from Q_8 to that group such that $\ker f = \langle -\mathbf{1} \rangle$.

Exercise 4.64: Use some of the ideas from Example 4.54(c), as well as the Subgroup Theorem.

Exercise 4.66: We can think of D_3 as generated by the elements ρ and φ, and each of these generates a non-trivial cyclic subgroup. Any automorphism α is therefore determined by these generators, so you can find all automorphisms α by finding all possible results for $\alpha(\rho)$ and $\alpha(\varphi)$, then examining that carefully.

Hints to Chapter 5

Exercise 5.31: Life will probably be easier if you convert it to cycle notation first.

Exercise 5.34: List the six elements of S_3 as $(\ 1\ )$, α, α^2, β, $\alpha\beta$, $\alpha^2\beta$, using the previous exercises both to justify and to simplify this task.

Exercise 5.35: Show that f is an isomorphism either exhaustively (this requires $6 \times 6 = 36$ checks for each possible value of $f\left(\rho^a \varphi^b\right)$), or by a clever argument, perhaps using using the Isomorphism Theorem (since $D_3 / \{\iota\} \cong D_3$).

Exercise 5.39: Try computing $\alpha \circ \beta$ and $\beta \circ \alpha$.

Exercise 5.71: Lemma 5.60 tells us that any permutation can be written as a product of cycles, so it will suffice to show that any cycle can be written as a product of transpositions. For that, take an arbitrary cycle $\alpha = \left(\begin{array}{cccc} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{array}\right)$ and write it as a product of transpositions, as suggested by Example 5.59. Be sure to explain why this product does in fact equal α.

Exercise 5.72: You can do this by showing that any transposition is its own inverse. Take an arbitrary transposition $\alpha = \left(\begin{array}{cc} \alpha_1 & \alpha_2 \end{array}\right)$ and show that $\alpha^2 = \iota$.

Exercise 5.74: Let α and β be arbitrary cycles. Consider the four possible cases where α and β are even or odd.

Exercise 5.75: See a previous exercise about subgroups or cosets.

Exercise 5.79: Use the same strategy as that of the proof of Theorem 5.78: find the permutation σ^{-1} that corresponds to the current configuration, and decide whether $\sigma^{-1} \in A_{16}$. If not, you know the answer is no. If so, you must still check that it can be written as a product of transpositions that satisfy the rules of the puzzle.

Hints to Chapter 6

Exercise 6.13: From Bézout's identity you know that you can find $x, y \in \mathbb{Z}$ such that $kx + ny = \gcd(k, n)$, *and* that $\gcd(k, n)$ is the *smallest positive integer* that can be written in this fashion. If $\gcd(k, n) = 1$, then any $\ell \in \mathbb{Z}$ can be written in the form $k\widehat{x} + n\widehat{y} = \ell$ (explain how to find $\widehat{x}$ and $\widehat{y}$!) and we can use this to write $\omega^\ell = \cdots = \left(\omega^k\right)^?$ (fill in the dots and the question mark!). Otherwise, if $\gcd(k, n) \neq 1$, then ω^k can only generate powers of $\omega^{\gcd(k,n)}$ (why?) which means we cannot generate ω itself, and thus cannot generate the entire group.

Exercise 6.24: At least you know that $\gcd(16, 33) = 1$, so you can apply the Chinese Remainder Theorem to the first two equations and find a solution in $\mathbb{Z}_{16 \cdot 33}$. Now you have to extend your solution so that it also solves the third equation; use your knowledge of cosets to do that.

Exercise 6.15: Since $d \in S$, we can write $d = am + bn$ for some $a, b \in \mathbb{Z}$. Show first that every common divisor of m, n is also a divisor of d. Then show that d is a divisor of m and n. For this last part, use the Division Theorem to divide m by d, and show that if the remainder is not zero, then d is not the smallest element of M.

Exercise 6.35: Use the properties of prime numbers.

Exercise 6.63: Consider Lemma 6.40 on page 209.

Exercise 6.66(c): Using the Extended Euclidean Algorithm might make this go faster. The proof of the RSA algorithm outlines how to use it.

Exercise 6.67:

(b) That largest number should come from encrypting ZZZZ.

(d) Using the Extended Euclidean Algorithm might make this go faster. The proof of the RSA algorithm outlines how to use it.

Exercise 6.68: There are a couple of ways to argue this. The best way for you is to explain why there exist a, b such that $ap + bq = 1$. Next, explain why there exist integers d_1, d_2 such that $m = d_1 a$ and $m = d_2 b$. Observe that $m = m \cdot 1 = m \cdot (ap + bq)$. Put all these facts together to show that $ab \mid m$.

Hints to Chapter 7

Exercise 7.12: The cases where $n = 0$ and $n = 1$ can be disposed of rather quickly; the case where $n \neq 0, 1$ is similar to (a).

Exercise 7.14:
(a) This is short, but not trivial. You need to show that $(-r)s + rs = 0_R$. Try using the distributive property.
(b) You need to show that $-1_R \cdot r + r = 0$. Try using a proof similar to part (a), but work in the additive identity as well.

Exercise 7.15: Proceed by contradiction. Show that if $r \in R$ and $r \neq 0, 1$, then something goes terribly wrong with multiplication in the ring.

Exercise 7.16: Use the result of Exercise 7.15.

Exercise 7.17: You already know that $(B, \oplus)$ is an additive group, so it remains to decide whether $\wedge$ satisfies the requirements of multiplication in a ring.

Exercise 7.33: Use the definition of equality in this set given in Example 7.23. For the first simplification rule, show the equalities separately; that is, show first that $(ac)/(bc) = a/b$; then show that $(ca)/(cb) = a/b$.

Exercise 7.34: For the latter part, try to find fg such that f and g are not even defined, let alone an element of $\mathrm{Frac}(R)$.

Exercise 7.49: Proceed by induction on $\deg f$. We did *not* say that r was unique.

Exercise 7.57: Showing that φ is multiplicative should be straightforward. To show that φ is additive, use the Binomial Theorem

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^n y^{n-i}$$

along with the fact that p is irreducible.

Exercise 7.69: $\mathbb{Z}[x]$ is a subring of what Euclidean domain? But don't be too careless–if you can find the gcd in that Euclidean domain, how can you go from there back to a gcd in $\mathbb{Z}[x]$?

Exercise 7.70: Since it's a field, you should never encounter a remainder, so finding a valuation function should be easy.

Exercise 7.71: There are two parts to this problem. The first is to find a "good" valuation function. The second is to show that you can actually divide elements of the ring. You should be able to do both if you read the proof of Theorem 7.58 carefully.

Exercise 7.72: For correctness, you will want to show something similar to Theorem 6.8 on page 197.

Exercise ??(a,iii): A system of equations could help with this latter division.

Hints to Chapter 8

Exercise 8.16: Use the Division Theorem for Integers (Theorem 0.41).

Exercise 8.18: The Extended Euclidean Algorithm (Theorem 6.8 on page 197) would be useful.

Exercise 8.23: For part (b), consider ideals of $\mathbb{Z}$.

Exercise 8.38: Show that if there exists $f \in \mathbb{F}[x,y]$ such that $x, y \in \langle f \rangle$, then $f = 1$ and $\langle f \rangle = \mathbb{F}[x,y]$. To show that $f = 1$, consider the degrees of x and y necessary to find $p, q \in \mathbb{F}[x,y]$ such that $x = pf$ and $y = qf$.

Exercise 8.39: Use the Ideal Theorem.

Exercise 8.67: Follow the argument of Example 8.59.

Exercise 8.71:

(c) Let g have the form $cx + d$ where $c, d \in \mathbb{C}$ are unknown. Try to solve for c, d. You will need to reduce the polynomial fg by an appropriate multiple of $x^2 + 1$ (this preserves the representation $(fg) + I$ but lowers the degree) and solve a system of two linear equations in the two unknowns c, d.

(e) Use the fact that $x^2 + 1$ factors in $\mathbb{C}[x]$ to find a zero divisor in $\mathbb{C}[x] / \langle x^2 + 1 \rangle$.

Exercise 8.72: Try the contrapositive. If $\mathbb{F}[x] / \langle f \rangle$ is not a field, what does Theorem 8.63 tell you? By Theorem 7.71, $\mathbb{F}[x]$ is a Euclidean domain, so you can find a greatest common divisor of f and a polynomial g that is not in $\langle f \rangle$ (but where is g located?). From this gcd, we obtain a factorization of f.

Or, follow the strategy of Exercise 8.71 (but this will be very, very ugly).

Exercise 8.73:

(a) Look at the previous problem.

(b) Notice that

$$y\left(x^2 + y^2 - 4\right) + I = I$$

and $x(xy - 1) + I = I$. This is related to the idea of the *subtraction polynomials* in later chapters.

Exercise 8.89: Use strategies similar to those used to prove Theorem 4.9 on page 144.

Exercise 8.92: Follow Example 8.83 on page 278.

Exercise 8.93: Multiply two polynomials of degree at least two, and multiply two matrices of the form given, to see what the polynomial map should be.

Exercise 8.94(d): Think about $i = \sqrt{-1}$.

Hints to Chapter 9

Exercise 9.51: There are one subgroup of order 1, three subgroups of order 2, one subgroup of order 3, and one subgroup of order 6. From Exercise 5.35 on page 176, you know that $S_3 \cong D_3$, and some subgroups of D_3 appear in Example 3.11 on page 113 and Exercise 3.20 on page 115.

Hints to Chapter 10

Exercise 10.13: You could do this by proving that it is a subring of $\mathbb{C}$. Keep in mind that $\left(\sqrt{-5}\right)\left(\sqrt{-5}\right) = -5$.

Exercise 10.37(d): Proceed by induction on n.

Exercise 10.41: Think of a fraction field over an appropriate ring.

Hints to Chapter 11

Exercise 11.63(b): Use part (a).

Exercise 11.64(c): Don't forget to explain why $\langle G \rangle = \langle G_{\text{minimal}} \rangle$! It is essential that the S-polynomials of these redundant elements top-reduce to zero. Lemma 11.56 is also useful.

Index

References

[AF05] Marlow Anderson and Todd Feil. *A First Course in Abstract Algebra*. Chapman and Hall/CRC, second edition, 2005.

[AP11] Alberto Arri and John Perry. The F5 Criterion revised. *Journal of Symbolic Computation*, 46(2):1017–1029, June 2011. Preprint online at `arxiv.org/abs/1012.3664`.

[Bah08] Tavmjong Bah. *Inkscape: Guide to a Vector Drawing Program*. Prentice-Hall, second edition, 2008. Retrieved from `www.inkscape.org`.

[Bri] Rogério Brito. The `algorithms` bundle. Retrieved 16 October 2008 from `http://www.ctan.org/tex-archive/macros/latex/contrib/algorithms/`. Version 0.1.

[Buc65] Bruno Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalem Polynomideal (An Algorithm for Finding the Basis Elements in the Residue Class Ring Modulo a Zero Dimensional Polynomial Ideal)*. PhD thesis, Mathematical Insitute, University of Innsbruck, Austria, 1965. English translation published in the Journal of Symbolic Computation (2006) 475–511.

[CLO97] David Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms*. Springer-Verlag New York, Inc., New York, second edition, 1997.

[CLO98] David Cox, John Little, and Donal O'Shea. *Using Algebraic Geometry*. Springer-Verlag New York, Inc., New York, 1998.

[EP10] Christian Eder and John Perry. F5C: A variant of Faugère's F5 algorithm with reduced Gröbner bases. *Journal of Symbolic Computation*, 45(12):1442–1458, 2010.

[EP11] Christian Eder and John Perry. Signature-based algorithms to compute Gröbner bases. In *Proceedings of the International Symposium on Symbolic and Algebraic Manipulation*, 2011.

[Fau02] Jean-Charles Faugère. A new efficient algorithm for computing Gröbner bases without reduction to zero F5. In *International Symposium on Symbolic and Algebraic Computation Symposium - ISSAC 2002, Villeneuve d'Ascq, France*, pages 75–82, Jul 2002. Revised version downloaded from `http://fgbrs.lip6.fr/jcf/Publications/index.html`.

[FR97] Benjamin Fine and Gerhard Rosenberger. *The Fundamental Theorem of Al-*

gebra. Undergraduate Texts in Mathematics. Springer-Verlag New York, Inc., 1997.

[GGV10] Shuhong Gao, Yinhua Guan, and Frank Volny. A New incremental algorithm for computing Groebner Bases. In *Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation*. ACM Press, 2010.

[GPS05] G.-M. Greuel, G. Pfister, and H. Schönemann. SINGULAR 3.0. A Computer Algebra System for Polynomial Computations, Centre for Computer Algebra, University of Kaiserslautern, 2005. http://www.singular.uni-kl.de.

[Grä04] George Grätzer. *Math into LaTeX*. Birkhäuser, Boston, third edition, 2004.

[Knu84] Donald Knuth. *The TeXbook*. Addison-Wesley Professional, spi edition, 1984.

[KR00] Martin Kreuzer and Lorenzo Robbiano. *Computational Commutative Algebra*, volume 1. Springer-Verlag, Berlin - Heidelberg - New York, 2000.

[Lam86] Leslie Lamport. *LaTeX: a Document Preparation System*. Addison-Wesley Publishing Company, 1986.

[Lau03] Niels Lauritzen. *Concrete Abstract Algebra: from Numbers to Gröbner Bases*. Cambridge University Press, Cambridge, 2003.

[LP98] Rudolf Lidl and Günter Pilz. *Applied Abstract Algebra*. Springer-Verlag, New York, second edition edition, 1998.

[Lyx] Lyx Team. *Lyx*, 2008–. Retrieved too many times to count from http://www.lyx.org.

[PHLA88] Abraham P. Hillman and Gerald L. Alexanderson. *A First Undergraduate Course in Absract Algebra*. Wadsworth Publishing Company, Belmont, California, fourth edition, 1988.

[Pic06] Paul Pichaureau. *The mathdesign package*, 2006. Retrieved from http://www.ctan.org/tex-archive/help/Catalogue/entries/mathdesign.html.

[Pip07] Sebastian Pipping. *ccBeamer*, 2007. Retrieved from http://www.hartwork.org/cgi-bin/download.cgi?file=CCBEAMER.

[RO08] Sebastian Rahtz and Heiko Oberdiek. *Hypertext marks in LaTeX:*

a manual for hyperref, 2008. Retrieved 21 April 2009 from http://www.tug.org/applications/hyperref/manual.html.

[Rot98] Joseph Rotman. *Galois Theory*. Springer-Verlag, New York, second edition, 1998.

[Rot06] Joseph J. Rotman. *A First Course in Abtract Algebra with Applications*. Pearson Education, Inc., New Jersey, third edition, 2006.

[Soc02] Americal Mathematical Society. *User's Guide for the* amsmath *Package*, version 2.0 edition, 2002. Retrieved 21 April 2009 from http://www.ams.org/tex/amslatex.html.

[Ste08] William Stein. *Sage: Open Source Mathematical Software (Version 3.1.1)*. The Sage Group, 2008. http://www.sagemath.org.

[vzGG99] Joachim von zur Gathen and Jurgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, Cambridge, 1999.