Lecture Notes for Math 620
Groups, Rings, and Fields

Michael E. O'Sullivan
mosullivan@sdsu.edu

August 17, 2023

Contents

Chapter 1

Getting Started

1.1 The Integers

What are the integers? This is not a simple question, if you want to be rigorous about defining the integers. Formally doing so would distract from developing our core topics, so we will take as our foundation the following. The ring of integers $\mathbb{Z}$ is

- the set of natural numbers $\mathbb{N} = \{1, 2, 3, \dots\}$ along with the number 0 and the additive inverses of the natural numbers $\{-1, -2, -3. \dots\}$;
- the operation of addition (and the properties of addition we know from elementary school);
- the operation of multiplication (and the properties of multiplication we know from elementary school);
- the ordering defined by positive numbers being greater than 0 and $a < b$ if and only if $b - a > 0$;
- the well-ordering principle—any non-empty subset of the natural numbers has a least element.

We may think of subtraction as either $a - b = a + (-b)$ or equivalently (after some argument) $a - b$ is the the number s (which we should show is unique) such that $s + b = a$. Division will be dealt with below. While the integers are familiar from elementary school, the well-ordering principle is not (unless you attended a very special elementary school!). It is actually key to the formal definition of the integers (see the Peano axioms and [Men15]) and to mathematical induction stripped to its essentials:

The Principle of Mathematical Induction
Let K be a subset of $\mathbb{N}$ satisfying these two properties

- $1 \in K$;
- whenever $a \in K$ it is also true that $a + 1 \in K$.

Then $K = \mathbb{N}$.

Beyond the basic properties above there are five main results for the integers that are fundamental. For the purposes of easy reference I will call them the Quotient-Remainder (QR) Theorem, the Greatest Common Divisor (GCD) Theorem, the Euclidean Algorithm, the Prime-Irreducible Theorem (Euclid's lemma), and the Unique Factorization Theorem (the Fundamental Theorem of Arithmetic). The Quotient-Remainder Theorem and Unique Factorization will be familiar; the other results, perhaps less so. The proofs here will be concise, and just a few exercises are included because this material is treated very well in other resources [Hun12].

In addition to these key results about the integers we introduce modular arithmetic in this section. This is a new algebraic structure know as the integers modulo n (for some $n > 1$), which we write $\mathbb{Z}/n$.

The Quotient-Remainder Theorem and Divisibility

We have implicitly used the Quotient-Remainder Theorem since elementary school, when we computed (the unique!) quotient and remainder of two integers. The proof relies on something sophisticated: the well-ordering principle.

Theorem 1.1.1 (Quotient-Remainder)**.** *Let a and b be integers with $b \neq 0$. There exist unique integers q, r such that*

(1) $a = bq + r$, *and*

(2) $0 \leq r < |b|$.

If the remainder of a divided by b is 0, we say b **divides** *a and a is a* **multiple of** *b.*

Proof. We prove this for $b > 0$ and leave the case $b < 0$ as an exercise. Consider the set $S = \{a - bc : c \in \mathbb{Z}\} \cap \mathbb{N}_{\geq 0}$. As a nonempty subset of the nonnegative integers, it has a least element. Let r be the minimal element of S, and let q be the integer such that $r = a - bq$. If $r \geq b$ we would have a contradiction because then $r - b \geq 0$ and $r - b = a - b(q + 1)$ would put $r - b \in S$. Thus we must have $0 \leq r < b$. This establishes existence of q, r as claimed.

To prove uniqueness, suppose another r', q' satisfy (1) and (2) and suppose without loss of generality that $r \geq r'$. (We want that $r' = r$ and $q' = q$.) Then $a = bq + r = bq' + r'$ so $r - r' = b(q' - q)$. Now $b > r - r' \geq 0$ but $b(q' - q)$ is a multiple of b. The only multiple of b in the interval $[0, b)$ is 0, so the only possibility is $r - r' = 0 = b(q - q)'$, and therefore $r = r'$ and $q = q'$. □

It is fairly common in programming languages (in particular in Python and Sage) to write the integer quotient as $a//b$ and the remainder as $a\%b$. We will use this in the exercises and the discussion of the Euclidean algorithm in this chapter.

Exercises 1.1.2. More on the Quotient-Remainder Theorem.

(a) Prove the QR Theorem for negative integers: Only minor changes are needed.

(b) For $b > 0$, show that $a//(-b) = -(a//b)$ and $a\%(-b) = a\%b$. [Don't let the notation make this hard!]

(c) Prove this alternative version of the QR Theorem. Let a and b be integers with $b \neq 0$. There exist unique integers q, r such that

(1) $a = bq + r$, and

(2) $|b|/2 < r \leq |b|/2$

[There are two approaches: use the existing QR Theorem to prove the alternative, or prove it from scratch by redefining S and modifying the proof of the QR Theorem.]

Let a and b be integers, at least one of which is not 0. The **common divisors** of a and b are the integers that divide both a and b. The **greatest common divisor** (gcd) is the largest positive integer dividing both a and b. The **common multiples** of a and b are the integers that are multiples of both a and b. The **least common multiple** (lcm) is the smallest positive integer that is a multiple of both a and b.

A **linear combination** of a and b is an integer that can be expressed as $au+bv$ for some integers u and v.

Exercises 1.1.3. Properties of divisibility.

(a) Show that if b divides a and d divides b that d also divides a.

(b) Show that if d divides a and d divides b that d also divides any linear combination of a and b. (In particular, this proves Lemma 1.1.5 below.)

The following result is an important property of the integers, and not an obvious one. It is an important tool in the study of groups. We will see echos of this result and the proof when we study ideals in rings (Section 4.4.

Theorem 1.1.4 (GCD)**.** *Let a and b be integers, at least one of them nonzero. The gcd of a and b is the smallest positive linear combination of a and b. In particular,* $\gcd(a,b) = au + bv$ *for some integers u and v.*

The set of all linear combinations of a and b equals the set of multiples of $\gcd(a,b)$*.*

Proof. Let $S = \{ar + bs : r, s \in \mathbb{Z}\}$ be the set of all linear combinations of a and b. Let d be the smallest *positive* element of S and let u, v be such that $d = au + bv$. I claim d divides a and b.

By the QR Theorem applied to a and d, $a = dq + r$ for some integer q and nonnegative integer $r < d$. Then

$$r = a - dq = a - (au + bv)q = a(1 - uq) - bvq$$

This shows that r is also in S. But, d is the smallest positive element of S, and $0 \leq r < d$. This shows $r = 0$, so d divides a.

Similarly, one shows d divides b, so d is a common divisor of a and b. To show it is the greatest common divisor, let c be any other common divisor of a and b. Then c divides $au + bv = d$ (by divisibility properties). Since d is positive $c \leq d$.

Since d divides a and b the elements of S are all divisible by d by Exercise 1.1.3. On the other hand any multiple of d is a linear combination of a and b since d is. This establishes the last sentence of the theorem. □

We say two integers are **coprime** (or **relatively prime**) when their gcd is 1.

Given a and b, how do we find their gcd? The answer (for arbitrary large integers) is not to factor each and look for common factors. Rather, use the Euclidean algorithm.

Let's assume $a \geq b \geq 0$. Recall that $a//b$ is the integer quotient and $a\%b$ the remainder as determined by the QR Theorem. Set $r_{-1} = a$ and $r_0 = b$, and define inductively (while $r_k \neq 0$)

$$\begin{aligned} q_k &= r_{k-1}//r_k \\ r_{k+1} &= r_{k-1}\%r_k, \quad \text{so that} \\ r_{k-1} &= q_k r_k + r_{k+1}. \end{aligned}$$

Rearranging these equalites by solving for r_{k+1}, we get a sequence

$$\begin{aligned} r_1 &= r_{-1} - r_0 q_0 = a - bq_0 \\ r_2 &= r_0 - r_1 q_1 \\ r_3 &= r_1 - r_2 q_2 \\ &\vdots \\ r_{k+1} &= r_{k-1} - q_k r_k \\ &\vdots \\ r_{n+1} &= r_{n-1} - q_n r_n = 0. \end{aligned}$$

The r_k are a strictly decreasing sequence of nonnegative integers, so the process must terminate: as we showed above for some n, $r_{n+1} = 0$ so r_n divides r_{n-1}. Now we make use of the following lemma, proved using basic divisibility properties (see Exercise 1.1.3).

Lemma 1.1.5. *For integers a, b, c, s such that $a = bs + c$, we have* $\gcd(a, b) = \gcd(b, c)$.

Let's apply this to the sequence r_k, letting n be minimal such that $r_{n+1} = 0$. We have (since $r_{n+1} = 0$)

$$\gcd(a, b) = \gcd(b, r_1) = \cdots = \gcd(r_n, r_{n+1}) = \gcd(r_n, 0) = r_n$$

This argument shows that the Euclidean algorithm produces the gcd of a and b. In the following Sage code we only keep two of the remainders at any time, not the whole sequence: after the kth pass through the while loop, r in the algorithm is r_{k-1} and s is r_k.

```
def euclid_alg(a,b):
   if b == 0:
      print ("division by zero")
      return false
   else:
      r = a
      s = b
      while s != 0:
         rem =  r %s
         r = s
         s = rem
      return r
```

There is a heftier Euclidean algorithm—often called the Extended Euclidean algorithm—which produces two integers u, v such that $au + bv = \gcd(a, b)$. I like the following matrix version of the algorithm. Let

$$Q_k = \begin{bmatrix} -q_k & 1 \\ 1 & 0 \end{bmatrix} \qquad \text{and} \qquad R_k = \begin{bmatrix} r_k \\ r_{k-1} \end{bmatrix}$$

where the sequence r_k and q_k are the same as used above in the Euclidean algorithm. Verify that $R_{k+1} = Q_k R_k$. Consequently,

$$R_{n+1} = \begin{bmatrix} 0 \\ r_n \end{bmatrix} = Q_n Q_{n-1} \cdots Q_0 R_0 \tag{1.1}$$

where $R_0 = \begin{bmatrix} b \\ a \end{bmatrix}$. Let $M = \begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{bmatrix} = Q_n Q_{n-1} \cdots Q_0$. Then, after the algorithm terminates, we have $M_{2,1} b + M_{2,2} a = r_n = \gcd(a, b)$.

Here is Sage code for the extended Euclidean algorithm. (Note that Sage indexes rows and columns of matrices starting from 0 not 1.) Initially M is the 2×2 identity matrix and R is the matrix $\begin{bmatrix} a \\ b \end{bmatrix}$. The algorithm iteratively computes q (which, at the kth iteration is q_{k-1}), the quotient of $M_{1,0}$ by $M_{0,0}$. It forms the matrix $Q = \begin{bmatrix} -q & 1 \\ 1 & 0 \end{bmatrix}$ and multiplies both R and M by Q. The result after iteration k (for $k = 1, \dots$) is that R is R_k and M is the product $Q_{k-1} Q_{k-2} \cdots Q_1 Q_0$ in (1.1).

```
def ext_euclid_alg(a,b):
   if b == 0:
      print ("division by zero")
      return false
   else:
      M = matrix.identity(2)
      R = matrix(2,1, [b,a])
      while R[0,0] != 0:
         q =  R[1,0]//R[0,0]
         Q = matrix(2,2,[ -q , 1, 1, 0])
         M = Q * M
         R = Q * R
      return M
```

We have proven that the Euclidean algorithm terminates with the greatest common divisor of the input integers. A bit closer analysis of the algorithm reveals a connection with the Fibonacci numbers and allows one to bound the number of steps for the Euclidean algorithm. See [Ros11] for details.

Theorem 1.1.6 (Euclidean Algorithm)**.** *Let a, b be integers with $b \neq 0$. The Euclidean Algorithm outputs* $\gcd(a,b)$ *in at most* $1 + \log_2 b / \log \alpha$ *steps, where* $\alpha = (1+\sqrt{5})/2$ *is the golden ratio. The Extended Euclidean Algorithm outputs integers u, v with $|u| < b$ and $|v| < a$ such that $au + bv = \gcd(a,b)$.*

The greatest common divisor of a finite set of integers (that contains a nonzero integer) is simply the largest integer that divides each element of the set. A simple induction argument shows that the set of common divisors of $\{a_1, \ldots, a_n\}$ is equal to the set of common divisors of $\{\gcd(a_1, \ldots, a_{n-1}), a_n\}$. To compute the greatest common divisor of this set efficiently, one computes iteratively $\gcd(a_1, \ldots, a_k) = \gcd(\gcd(a_1, \ldots, a_{k-1}), a_k)$. (There are more efficient algorithms, but understanding this approach is sufficient here.)

Exercises 1.1.7. Use the Euclidean algorithm to express the greatest common divisor as a linear combination of the given integers.

(a) 89, 24

(b) 24, 40, 30

Primes, Irreducibles, and Unique Factorization

The next main result is Fundamental Theorem of Arithmetic, which says (roughly) that every nonzero integer has a unique factorization as a product of primes. We will now define what it means for an integer to be prime, but it will not be the school definition. We also define the term irreducible, which is what we customarily use for primality. The definitions given here are the accepted ones in more general contexts. Fortunately, the following theorem (Euclid's Lemma), which I will refer to as the Prime-Irreducible Theorem shows that for integers the notions are equivalent.

Definition 1.1.8. Let r be an integer with $|r| > 1$. We say r is **irreducible** when $r = ab$ implies that either $a = \pm 1$ or $b = \pm 1$ (and the other is $\pm r$). We say r is **prime** when $r|ab$ implies $r|a$ or $r|b$.

Theorem 1.1.9 (Prime-Irreducible)**.** *An integer is irreducible if and only if it is prime.*

Proof. We prove this for positive integers; minor adjustments can be made for a negative number.

Let $p > 1$ be an irreducible; let us show it is prime. Suppose that $p|ab$ for some integers a and b. We need to show $p|a$ or $p|b$. If p divides a we are done, so suppose it does not divide a. Since p is irreducible, its only positive divisors are 1

and p, so the GCD of a and p is 1. By the GCD Theorem, there are integers u, v such that

$$1 = au + pv$$

Multiplying by b

$$b = abu + pbv$$

Since $p|ab$ we have that p divides the right hand side, thus p divides b.

Suppose now that $p > 1$ is prime, we will show it is irreducible. Let $p = ab$ be a factorization of p. We must show one of a or b is ± 1. Since p is prime and it divides (in fact equals) the product ab it must divide one of the factors. Without loss of generality, say $p|a$, so $a = px$ for some integer x. Then $p = ab = pxb$, so $p(1 - xb) = 0$. This shows that $1 - xb = 0$, so $x = b = \pm 1$. □

The previous theorem is the key ingredient to establishing unique factorization.

Theorem 1.1.10 (Unique Factorization)**.** *Let a be a positive integer. There is a nonnegative integer t, there are positive prime numbers $p_1 < p_2 < \cdots < p_t$, and there are positive integers $e_1, \ldots, e_t$, such that*

$$a = p_1^{e_1} \cdots p_t^{e_t}$$

Each of t, p_i, e_i and u is uniquely determined.

Proof. TBD □

Modular Arithmetic

Our discussion of the integers culminates with a quick summary of arithmetic modulo an integer n. This is a model for the construction of quotient groups and quotient rings that will be taken up later.

Definition 1.1.11. Let n be a nonzero integer. For integers a and b we say a is **congruent modulo** n to b when n divides $a - b$.

Theorem 1.1.12. *Congruence modulo n is an eqivalence relation. Furthermore, the set $\{0, 1, \ldots, n-1\}$ is a system of representatives for congruence modulo n in the sense that each integer a is congruent modulo n to exactly one element of $\{0, 1, \ldots, n-1\}$.*

Proof. The relation of being congruent modulo n is clearly reflexive, since for any $a \in \mathbb{Z}$, $n|(a - a)$. It is symmetric because if $n|(a - b)$ then also $n|(b - a)$. It is transitive because if a is congruent to b and b is congruent to c modulo n then

$n|(a-b)$ and $n|(b-c)$. This implies that n divides the sum $(a-b)+(b-c) = a-c$, so a is congruent to c modulo n.

From the Quotient-Remainder Theorem, an integer a is congruent to its remainder when divided by n, since there is an integer q such that $a = nq+r$. Theis remainder is one of the elements of $\{0, 1, \ldots, n-1\}$. No two of these numbers differ by a multiple of n so they are distinct modulo n. □

The integers n and $-n$ give the same equivalence relation, so we always use positive integers for the modulus. It is common to write $[a]_n$ for the congruence class of a modulo n, whenever we need to be careful to distinguish between the integer a and the congruence class, or when we have more than one modulus to worry about. If there is a unique modulus the subscript n may be omitted. If it is clear from context that we are working modulo n, we may simply write a.

Finally we have:

Theorem 1.1.13 (Arithmetic modulo n)**.** *Suppose that $a \equiv b \bmod n$ and $r \equiv s \bmod n$. Then $a + r \equiv b + s \bmod n$ and $ar \equiv bs \bmod n$. Thus, arithmetic on congruence classes modulo n is well-defined as follows.*

- $[a] + [r] = [a + r]$
- $[a] * [r] = [ar]$

Proof. Suppose that $a \equiv b \bmod n$ and $r \equiv s \bmod n$. We have $a = b + jn$ for some integer j and $r = s + kn$ for some integer k. Then $a + r = b + s + (j + k)n$ so $a + r \equiv b + s \bmod n$. We also have $ar = bs + (ak + bj + jkn)n$ so $ar \equiv bs \bmod n$.

This shows that no matter what element of an congruence class is used to represent the class, arithmetic operations modulo n will give the same result. □

We will write $\mathbb{Z}/n$ for the set of congruence classes modulo n, with the operations $+$ and $*$. (It is common to use $\mathbb{Z}_n$, but $\mathbb{Z}/n$ is consistent with notation we will use later.) When there is no chance of ambiguity, we write the congruence classes as $0, 1, \ldots, n-1$ (without the brackets and using the least nonnegative representatives for each class). But, sometimes it is handy to be a bit flexible. For example it is good to remember that $n - b$ is equal to $-b$ in $\mathbb{Z}/n$. So (in $\mathbb{Z}/n$)

$$b(n-1) = (-1)b = n - b.$$

Exercises 1.1.14.

(a) Let p be a prime number. Let $[a] \in \mathbb{Z}/p$ with $[a] \neq [0]$ (so a is not divisible by p). Use the GCD Theorem 1.1.4 to show there is some $r \in \mathbb{Z}/p$ such that $[a][r] = [1]$.Consequently, each nonzero element of $\mathbb{Z}/p$ has a multiplicative inverse.

(b) Extend this result, partially, to $\mathbb{Z}/n$ for composite n. If $[a] \in \mathbb{Z}/n$ is such that the integer a is coprime to n, then there is some $[r] \in \mathbb{Z}/n$ such that $[a][r] = [1]$.

(c) Use the Quotient-Remainder Theorem from Exercises 1.1.2 to show alternative sets of representatives for the integers modulo n are:

$$-\frac{n-1}{2}, -\frac{n-3}{2}, \ldots, \frac{n-3}{2}, \frac{n-1}{2} \qquad \text{for } n \text{ odd, and,}$$

$$-\frac{n-2}{2}, -\frac{n-4}{2}, \ldots, \frac{n-4}{2}, \frac{n-2}{2}, \frac{n}{2} \qquad \text{for } n \text{ even.}$$

1.2 Groups, Rings, and Fields

Let us now introduce our three objects of study: groups, rings, and fields. This section will discuss some familiar number systems in the context of groups, rings, and fields. We include also some perspective on the history of the number systems as well as their appearance in our system of education.

Definition 1.2.1. A **binary operation** on a set S is a function from $S \times S$ to S. A **unary operation** on S is a function from S to S. An "operation" on S is usually assumed to be binary if not stated otherwise[1].

A binary operation $*$ on S is **associative** when $(a * b) * c = a * (b * c)$. It is **commutative** when $a * b = b * a$. It has an **identity element** when there is some element $e \in S$ such that $a * e = e * a = a$ for all $a \in S$.

A group has one binary operation, generally denoted $*$, while rings and fields have two binary operations, generally denoted $+$ and $*$.

Definition 1.2.2. A **group** is a set G with a binary operation $*$ and a unary operation, denoted $a \longmapsto a^{-1}$, satisfying the following properties.

(1) Associativity of $*$.

(2) Identity for $*$: There is an element, generally denoted e, such that $e * a = a = a * e$ for all $a \in G$.

(3) Inverses for $*$: For each $a \in G$ the unary operation $a \longmapsto a^{-1}$ gives the inverse for a. That is, $a * a^{-1} = e = a^{-1} * a$.

A group which also satisfies $a * b = b * a$ is called **commutative** or **abelian** (after the mathematician Abel).

Definition 1.2.3. A **ring** is a set R, with two operations $+$ and $*$ that satisfy the following properties.

(1) Associativity for both $+$ and $*$.

(2) Commutativity for both $+$ and $*$.

(3) Identity elements for both $+$ and $*$. There is some element in R, that we call 0, such that $a + 0 = a$ and there is an element, that we call 1, such that $a * 1 = a$.

[1] One can define ternary ($S \times S \times S \longrightarrow S$) and, more generally, n-ary operations, but we will have no use for these.

(4) Inverses for $+$. For each $a \in R$ there is some other element, which we write $-a$, such that $a + (-a) = 0$.

(5) Distributivity of $*$ over $+$. That is, $a * (b + c) = a * b + a * c$.

A **field** is a ring with one additional property,

(6) Inverses for $*$. For each nonzero $a \in R$ there is some other element, that we write a^{-1}, such that $a * a^{-1} = 1$.

Comparing these definitions, one sees that a ring R under the operation $+$ is an abelian (commutative) group with identity element 0 and additive inverse operation $a \longmapsto -a$. Under the operation of multiplication, $*$, a ring may lack the property of inversion. A field F is an abelian group under $+$ and the set of nonzero elements, $F^* = F \setminus \{0\}$, is an abelian group under $*$. The interaction between the two operations of $+$ and $*$ for both rings and fields is given by the distributive property.

Now to the question: what examples do we have of groups, rings, and fields?

Familiar Rings and Fields

The first number system that a child learns in school is the natural numbers $\mathbb{N} = \{1, 2, 3, 4, \dots\}$, and eventually this is expanded to the integers by including 0 and the additive inverse of each positive integer. The integers, denoted $\mathbb{Z}$, are an abelian group under addition. Once the operation of multiplication as repeated addition is introduced, we have the first example of a ring. The integers in fact form the prototypical ring, as we shall see in Theorem 4.2.11.

Students in elementary school—the lucky ones—may also learn "clock arithmetic" in which addition is done on a clock, so 8:00 plus 7 hours is 3:00. This is essentially modular arithmetic with modulus 12 (although we usually use representatives $1{:}00, 2{:}00, \dots, 12{:}00$ rather than using $0{:}00, 1{:}00, \dots, 11{:}00$). We saw in Section 1.1 that multiplication is also well defined modulo n, and one can check that the properties of a ring are satisfied. We will denote this number system $\mathbb{Z}/n$ (although $\mathbb{Z}_n$ is also commonly used).

The next step in mathematics education is to expand this integer number system. The integers do not form a field since the only numbers with a multiplicative inverse in $\mathbb{Z}$ are ± 1. There is a complicated process that enlarges the set of integers by adding fractions to create the rationals, $\mathbb{Q}$. I say the process is complicated because lots of people have trouble understanding fractions well, and a key part of the problem is that a given number has an infinite number of different names: $1/2 = 2/4 = 3/6 = \cdots$. The process of forming fractions can be generalized to other rings, but it has delicate and subtle steps involving equivalence relations.

When you see the construction in Section 4.6 you may appreciate that these subtleties are closely tied to the difficulties people have with fractions.

There are two other fields that are introduced in secondary school education, although they are challenging to understand fully: the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$. Formally defining the real numbers is a sophisticated process, but treating $\mathbb{R}$ as the set of all decimal numbers (including infinite non-repeating ones) and focusing on the number line is a way to work with them effectively enough to do most college level mathematics. We won't have much need for the real numbers, but the relationship between the reals and the complex numbers is something that is key to studying fields in general. A complex number is of the form $a+bi$ in which $a, b \in \mathbb{R}$ and i is the square root of -1. Addition is "componentwise," $(a+bi)+(c+di) = (a+c)+(b+d)i$, and the additive inverse of $a+bi$ is $-a-bi$. Multiplication is based on $i^2 = -1$: $(a+bi)*(c+di) = (ac-bd)+(ad+bc)i$. One can check that $(a-bi)/(a^2+b^2)$ is the multiplicative inverse of $a+bi$. With these operations, $\mathbb{C}$ is a field.

In the above discussion, there was actually no need to use real numbers for a, b, c, d. We could have restricted them to be rational numbers and the statements about addition, multiplication and inverses would still hold true. Thus we can introduce a field derived from $\mathbb{Q}$ that includes i and uses the rules above for addition, multiplication, and the inverses for each. We call this field the Gaussian rationals $\mathbb{Q}(i)$.

There is one other field that is accessible to those who have learned "clock arithmetic." If our clock had a prime number p of positions, we would get $\mathbb{Z}/p$, in which every nonzero element has a multiplicative inverse, as was shown in Exercise 1.1.14. When we focus on this modular ring as a field we will write it as $\mathbb{F}_p$, instead of $\mathbb{Z}/p$. Thus we have our slate of elementary fields, $\mathbb{Q}$, $\mathbb{Q}(i)$ and $\mathbb{F}_p$, supplemented if we want by $\mathbb{R}$ and $\mathbb{C}$.

Each of these fields is of course a ring. Our collection of rings that are not fields includes $\mathbb{Z}$ and $\mathbb{Z}/n$ for n not a prime. We may supplement it by the Gaussian integers

$$\mathbb{Z}[i] = \{a+bi : a, b \in \mathbb{Z}\}$$

It is routine to check that this is a ring. Notice that there are 4 elements that have multiplicative inverses, $\pm 1, \pm i$.

All of our rings are groups under addition and we can study them as groups by "forgetting" the multiplicative structure ("forget" is actually a term used by mathematicians in this context!) We also obtain a few other examples by looking at the multiplicative group of a field. These are the nonzero rationals, denoted $\mathbb{Q}^*$, and the nonzero elements of $\mathbb{Q}(i)$, and the nonzero elements of $\mathbb{F}_p$, denoted $\mathbb{F}_p^*$ (and similarly for $\mathbb{R}$ and $\mathbb{C}$).

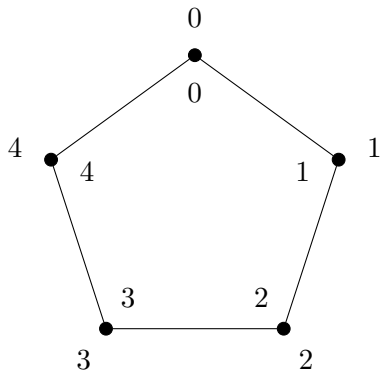


Figure 1.1: The Pentagon

But what about a group that doesn't involve ignoring one of the operations on a ring? In particular, so far, we have no examples of groups in which the operation is not commutative.

The Dihedral Groups, D_n

Group theory actually arose from the study of transformations that preserved structure in other mathematical areas. The symmetry of physical objects is perhaps the easiest entry point.

Consider a regular pentagon, as in Figure 1.1. Imagine a table with a pentagon carved into it and a clear pentagon that fits neatly into the enclosure. Enumerate the "base points" on the table and the vertices of the pentagon as shown.

Rotation counter-clockwise about the center by 72° takes the pentagon to itself. Only the enumeration of the vertices would indicate that a change occured. Calling this rotation r, we can see there are 5 rotational symmetries, which are rotation by $72°, 144°, 216°, 288°$ and $360°$. The latter has the same result as not moving the pentagon at all. This set of rotations is a group where the operation is just doing one rotation followed by another. Thinking of rotation as a function, we are composing functions. We may write these rotations as r, r^2, r^3, r^4 and r^5; the latter having the same effect as not moving pentagon at all, so $r^5 = r^0$ is the identity element. It should be clear that $r^i * r^j = r^{i+j}$. This group has 5 elements and "looks a lot like" $\mathbb{Z}/5$ under addition. (It is *isomorphic* to $\mathbb{Z}/5$, see Section 2.2.)

There is another type of symmetry indicated by the dashed line. For each vertex of the pentagon $0, \ldots, 4$ there is a line through that vertex and the midpoint of the opposite side that is an axis of symmetry for reflection. Let t_i be the

reflection across the line at vertex i. We now have 10 symmetries of the pentagon: the identity, 4 non-trivial rotations, and 5 reflections. I claim these are the only symmetries. There are 5 possible places to put vertex 1; but then vertex 2 must be one notch away, either clockwise or counterclockwise. The positions of the other vertices are then determined by the rigidity of the pentagon. So there are only 10 possibilities. Notice also that after a rotation, the ordering of the pentagon vertices increases clockwise, but after a reflection the numbers of the pentagon vertices increase counterclockwise (and the numbers on the pentagon would be reversed as in a mirror).

Now let's consider the group operation: what happens when we follow one symmetry by another? First, there are some ambiguities to clear up, so we adopt the following conventions:

- The rotation r is rotation of the vertex i to the position of vertex $i-1$, so it is counter-clockwise when the enumeration on the pentagon increases clockwise and clockwise when the enumeration on the pentagon increases counter-clockwise (as it is after a reflection).
- A reflection t_i is reflection across the line through the ith vertex of the pentagon, not the label i on the table.
- The product rt_i means reflect then rotate as is customary using functional notation; we apply the function on the right first.

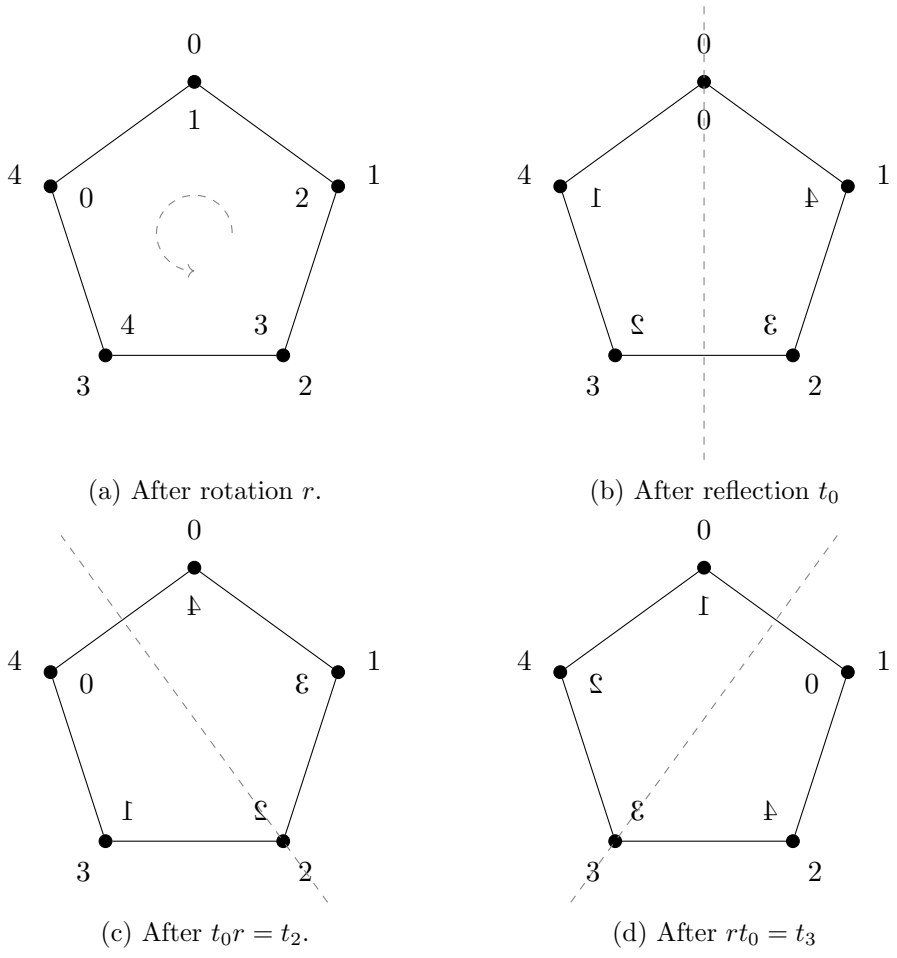
Figure 1.2 shows the two computations, $t_0 r = t_2$ and $rt_0 = t_3$. These are unequal, so these computations show that we have our first example of a nonabelian group. It is called the dihedral group of order 5, and written D_5.

Definition 1.2.4. For $n \geq 3$, the **dihedral group** of order n, written D_n, is the group of symmetries of a regular n-gon. In addition to the identity there are $n-1$ rotations and n reflections. The group operation is composition of the transformations.

One can verify, in a similar manner to that above, that D_n is nonabelian for all n.

We can consider each of the symmetries of the pentagon as a function on $\mathbb{Z}/5$ that assigns to the base point a on the table the index of the pentagon vertex at position a after applying the symmetry. So $r(a) = a+1$, and $r^i(a) = a+i$ (computing modulo 5). The following exercise develops this example in more detail.

Exercises 1.2.5. Continuing with the D_5 example above, the goal of this problem is to find formulas for the product of two arbitrary elements of D_5. We will use arithmetic in $\mathbb{Z}/5$ with the system of representatives $0, 1, 2, 3, 4$.



(a) After rotation r.

(b) After reflection t_0

(c) After $t_0 r = t_2$.

(d) After $r t_0 = t_3$

Figure 1.2: The pentagon after various transformations

(a) Observe that the reflection t_i applied to the original position of the pentagon (in Figure 1.1) switches $i+1$ with $i-1$ and $i+2$ with $i-2$ where computations are modulo 5. Show that when t_i is applied to the original position of the pentagon the vertex at base point a is $2i - a$. We can write this as $t_i(a) = 2i - a$.

(b) Show that $rt_i = t_{i+3}$ by arguing that rt_i is a reflection and that, applied to the pentagon in the original position, it takes $i + 3$ to itself.

(c) Find a formula for $r^j t_i$; that is, give a function of $\mathbb{Z}/5$ [Hint: linear] for $r^j t_i(a)$. Do the same for $t_i r^j$.

(d) Explain why the product of two reflections is a rotation, and find a formula for $t_i t_j(a)$.

The Symmetric Groups, S_n

Recall that a function from one set to another is a bijection when it is both injective (one-to-one) and surjective (onto). If $f : A \longrightarrow B$ is a bijection, then there is a well-defined inverse function, f^{-1}, since each element of B has exactly one preimage. If $f : A \longrightarrow A$ is a bijection from A to itself then we say f **permutes** the elements of A; it rearranges them in a sense. We are particularly interested in the case when A is a finite set, even more specifically the set $\{1, 2, 3, \ldots, n\}$. In this case it is convenient to write a permutation as a table with the columns $i, f(i)$. Here are two examples (it is common to use Greek letters to denote permutations).

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

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

This tabular format makes it evident that there are $n!$ permutations of a set with n elements: There are n choices for the image of 1, call it $a_1 \in \{1, \ldots, n\}$, then there are $n-1$ possible images for 2, since it must be in $\{1, \ldots, n\} \setminus \{a_1\}$ and so on. The table would then be

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ a_1 & a_2 & a_3 & \ldots & a_n \end{pmatrix}$$

The tabular form indicates the sense in which a permutation is a rearrangement, with a_1 now being in the first position, a_2 in the second, and so on.

Since the composition of two bijections from A to itself is itself a bijection from A to itself, composition is an operation on the set of permutations. The inverse of a permutation is also a permutation. Thus we can make the following definition.

Definition 1.2.6. Let n be a positive integer. The set of all permutations of $\{1, 2, 3, \ldots, n\}$ along with composition and the unary operation that takes a permutation to its inverse function is called the **symmetric group** on $\{1, \ldots, n\}$ and is denoted S_n. We will write composition of permutations using $\circ$ when emphasizing that permutations are functions, but generally we use $*$, which is the usual notation for a product in groups.

For $n = 1$ the symmetric group has just 1 element, and for $n = 2$ it has two. The group S_3 has 6 elements. Three of these elements fix exactly one element, as σ does, and are called transpositions. The other two are π and $\pi^{-1} = \pi \circ \pi$. These are called 3-cycles. The symmetric groups S_n for $n \geq 3$ are nonabelian. One can compute $\pi \circ \sigma$ and $\sigma \circ \pi$ and see that they are unequal (applying the right hand function first is our convention).

$$\pi \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \qquad \sigma \circ \pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

The symmetric groups will be discussed in detail in Section 2.4.

An astute reader has perhaps noticed that the discussion of D_5 yielded a permutation of $\{0, 1, 2, 3, 4\}$ for each element of D_5. Allowing ourselves to let S_5 be the permutation group of $\{0, 1, \ldots, 4\}$ for the moment, we have essentially given a function of D_5 into S_5. Using a bit of Exercise 1.2.5 (and computing in $\mathbb{Z}/5$) we have;

$$r^j \longmapsto \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ j & 1+j & 2+j & 3+j & 4+j \end{pmatrix}$$
$$t_i \longmapsto \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 2i & 2i-1 & 2i-2 & 2i-3 & 2i-4 \end{pmatrix}.$$

One can check that composition of the linear functions from Exercise 1.2.5 agree with the composition of the permutations; we are just composing functions.

In the terminology of the next chapter, we have given a homomorphism (Section 2.2) from D_5 to S_5 and the image is a subgroup of S_5 (Section 2.1).

Symmetry of Other Objects

The aesthetic appeal of symmetrical objects seems to be universal in human culture. The following images, some purely geometric, and some from artwork of various civilizations, show how rich the notion of symmetry can be. How does one describe the symmetry group of each of these objects? How does one understand

the structure of these groups? These are questions that I hope will motivate the next two chapters.

NEED SOME IMAGES: CUBE, CRYSTAL, FRIEZE DIAGRAMS, TILINGS,

1.3 The Univariate Polynomial Ring over a Field

Let F be a field, in particular, we may consider F to be $\mathbb{Q}$, $\mathbb{F}_p$, or the real numbers. By $F[x]$ we mean the polynomial ring in the indeterminate x. The key fact about polynomial rings is that all the theorems we discussed for integers in Section 1.1 also hold—with appropriate modifications—for polynomial rings over F. These are the Quotient-Remainder (QR) Theorem, the Greatest Common Divisor (GCD) Theorem, the Euclidean Algorithm, the Prime-Irreducible Theorem, and the Unique Factorization Theorem. This close relationship between $\mathbb{Z}$ and $F[x]$ is such an important theme in algebra and number theory that I want to lay out the fundamentals in detail in this section, which parallels substantially Section 1.1.

We can think of the polynomial ring as

- a vector space over F with an infinite basis $1, x, x^2, \ldots$, and componentwise addition;
- a multiplicative structure defined by $x^i * x^j = x^{i+j}$ and the properties of commutativity, associativity and distributivity of multiplication over addition.

The result is thus a ring.

Sometimes it is useful to write a polynomial $b(x) \in F[x]$ as a sum $b(x) = \sum_{i=0}^{\infty} b_i x^i$ with the understanding that only a finite number of the b_i are nonzero. When all of the $b_i = 0$ we get the **zero polynomial**. Suppose $b(x) \neq 0$ and let δ be the largest integer such that $b_\delta \neq 0$. We call δ the **degree** of $b(x)$; $b_\delta x^\delta$ is the **leading term** of $b(x)$; x^δ is the **leading monomial** of $b(x)$; and b_δ is the **leading coefficient** of $b(x)$. If $b_\delta = 1$ we say $b(x)$ is **monic**. When $\delta = 0$ we say $b(x)$ is a **constant** polynomial. The zero polynomial is also considered a constant polynomial and the degree is sometimes defined to be $-\infty$.

The product of a polynomial $a(x)$ of degree γ and $b(x)$ of degree δ has degree $\gamma + \delta$. Rules for divisibility of polynomials are similar to those for the integers. In particular if $a(x)$ divides $b(x)$ and $b(x)$ divides $c(x)$ then $a(x)$ divides $c(x)$. Furthermore if $d(x)$ divides both $a(x)$ and $b(x)$ then it divides their sum (and also any multiple of either $a(x)$ or $b(x)$).

The Quotient-Remainder Theorem and Divisibility

The following lemma simplifies the proof of the Quotient-Remainder Theorem. It is worth remarking that we are using the properties of a field when we divide by b_γ.

Lemma 1.3.1 (Division)**.** *Let $a(x)$ and $b(x)$ be in $F[x]$ with degrees γ and δ respectively and $\gamma > \delta$. Then the degree of $a(x) - \frac{a_\gamma}{b_\delta} x^{\gamma-\delta} b(x)$ is less than γ.*

Proof. This is a straightforward computation.

$$\begin{aligned}
a(x) &- \frac{a_\gamma}{b_\delta} x^{\gamma-\delta} b(x) \\
&= a_\gamma x^\gamma + a_{\gamma-1} x^{\gamma-1} + \cdots + a_1 x + a_0 \\
&\qquad - \frac{a_\gamma}{b_\delta} x^{\gamma-\delta} \Big(b_\delta x^\delta + b_{\delta-1} x^{\delta-1} + \cdots + b_1 x + b_0 \Big) \\
&= a_\gamma x^\gamma + a_{\gamma-1} x^{\gamma-1} + \cdots + a_1 x + a_0 \\
&\qquad - \Big(a_\gamma x^\gamma + \frac{a_\gamma b_{\delta-1}}{b_\delta} x^{\gamma-1} + \cdots + \frac{a_\gamma b_1}{b_\delta} x^{\gamma-\delta+1} + \frac{a_\gamma b_0}{b_\delta} x^{\gamma-\delta} \Big) \\
&= \Big(a_{\gamma-1} - \frac{a_\gamma b_{\delta-1}}{b_\delta} \Big) x^{\gamma-1} + \Big(a_{\gamma-2} - \frac{a_\gamma b_{\delta-2}}{b_\delta} \Big) x^{\gamma-2} + \cdots
\end{aligned}$$

The leading terms of $a(x)$ and $\frac{a_\gamma}{b_\delta} x^{\gamma-\delta} b(x)$ cancel and the result has degree less than γ. □

Theorem 1.3.2 (Quotient-Remainder)**.** *Let $a(x)$ and $b(x)$ be elements of $F[x]$ with $b(x) \neq 0$. There exist unique $q(x)$, $r(x)$ such that*

(1) $a(x) = b(x)q(x) + r(x)$, and

(2) $\deg r(x) < \deg b(x)$.

Proof. Consider the set $S = \{a(x) - b(x)c(x) : c(x) \in F[x]\}$. The set of degrees of the elements of S is a nonempty subset of the nonnegative integers, so it has a least element, δ. There is some polynomial of degree δ in S, call it $r(x)$, and suppose $r(x) = a(x) - b(x)q(x)$. I claim $\deg(r(x)) < \deg(b(x))$. Suppose not. Let $\gamma = \deg(r(x))$ and $\delta = \deg(b(x))$. Apply Lemma 1.3.1 to $r(x)$ and $b(x)$ to get

$$\begin{aligned}
r(x) - \frac{r_\delta}{b_\gamma} x^{\gamma-\delta} b(x) &= a(x) - b(x)q(x) - \frac{r_\delta}{b_\gamma} x^{\gamma-\delta} b(x) \\
&= a(x) - b(x)(q(x) + \frac{r_\delta}{b_\gamma} x^{\gamma-\delta}).
\end{aligned}$$

This is also in S and by Lemma 1.3.1 has lower degree than $r(x)$. This contradicts our choice of δ as the lowest degree of elements in S. Consequently, we must have $\deg(r(x)) < \deg(b(x))$. This establishes existence of $q(x)$ and $r(x)$ as claimed.

To prove uniqueness, suppose another $r'(x), q'(x)$ satisfy (1) and (2). (We want to show they are equal to $r(x)$ and $q(x)$!) Then

$$\begin{aligned}
a(x) = b(x)q(x) + r(x) = b(x)q'(x) + r'(x) \quad &\text{so} \\
r(x) - r'(x) &= b(x)\big(q'(x) - q(x)\big).
\end{aligned}$$

The degree on the left hand side is strictly less than the degree of $b(x)$. Since the right hand side is a multiple of $b(x)$, it must in fact be 0. Thus $r(x) = r'(x)$ and $q(x) = q'(x)$. □

The greatest common divisor of two integers was easy to define since the integers are well ordered. It is not obvious that, among the common divisors of $a(x), b(x) \in F[x]$, there is just one monic divisor of maximal degree. The following theorem shows the gcd can be uniquely defined and extends the GCD Theorem to polynomials. A **polynomial combination** of $a(x)$ and $b(x)$ is a polynomial that can be expressed as $a(x)u(x) + b(x)v(x)$ for some $u(x), v(x) \in F[x]$.

Theorem 1.3.3 (GCD)**.** *Let $a(x), b(x) \in F[x]$ with at least one of them nonzero. There is a unique polynomial $d(x)$ satisfying*

(1) $d(x)$ is a common divisor of $a(x)$ and $b(x)$,

(2) $d(x)$ is monic,

(3) $d(x)$ is divisible by all other common divisors of $a(x)$ and $b(x)$ (so it is the greatest common divisor).

Furthermore, there exist $u(x), v(x) \in F[x]$ such that $d(x) = a(x)u(x) + b(x)v(x)$. The set of all polynomial combinations of $a(x)$ and $b(x)$ equals the set of multiples of $d(x)$.

Proof. Let $S = \{a(x)s(x) + b(x)t(x) : s(x), t(x) \in F[x]\}$ be the set of all polynomial combinations of $a(x)$ and $b(x)$. Let $d(x)$ be a nonzero polynomial of minimal degree in S and let $u(x), v(x)$ be such that $d(x) = a(x)u(x) + b(x)v(x)$. We may assume that $d(x)$ is monic, since any constant multiple of a polynomial in S is also in S. I claim $d(x)$ divides $a(x)$ and $b(x)$.

By the QR Theorem applied to $a(x)$ and $d(x)$, $a(x) = d(x)q(x) + r(x)$ for some $q(x)$ and $r(x)$ in $F[x]$ with $\deg(r(x)) < \deg(d(x))$. Then

$$\begin{aligned} r(x) &= a(x) - d(x)q(x) \\ &= a(x) - \big(a(x)u(x) + b(x)v(x)\big)q(x) \\ &= a(x)\big(1 - u(x)q(x)\big) - b(x)v(x)q(x) \end{aligned}$$

This shows that $r(x)$ is also in S. If it were nonzero, it could be multiplied by a constant to get a monic element of S with lower degree than $d(x)$, which contradicts the choice of $d(x)$. We can thus conclude that $r(x) = 0$, and consequently $d(x)$ divides $a(x)$.

Similarly, one shows $d(x)$ divides $b(x)$. Thus $d(x)$ is a common divisor of $a(x)$ and $b(x)$. To show it is the greatest common divisor, let $c(x)$ be any other

common divisor of $a(x)$ and $b(x)$. Then $c(x)$ divides $a(x)u(x) + b(x)v(x) = d(x)$ (by divisibility properties) as claimed.

Finally, we note that the set S (which we defined to be the set of polynomial combinations of $a(x)$) is also the set of multiples of $d(x)$. Since $d(x)$ divides each of $a(x)$ and $b(x)$ it will divide any polynomial combination of $a(x)$ and $b(x)$ by divisibility properties. On the other hand, since $d(x)$ is a polynomial combination of $a(x)$ and $b(x)$, any multiple of $d(x)$ is also a polynomial combination of $a(x)$ and $b(x)$ and therefore in S. □

The proof of the theorem can be adapted for any set of polynomials $P \subseteq F[x]$. One can show that the smallest degree monic polynomial that can be expressed as a combination of the elements of P actually divides all the elements of P.

The discussion of the Euclidean algorithm for integers carries over almost verbatim to $F[x]$. We use $a(x)//b(x)$ for the polynomial quotient and $a(x)\%b(x)$ for the remainder of division of $a(x)$ by $b(x)$.

Assume $\deg(a(x)) \geq \deg(b(x))$. Set $r_{-1}(x) = a(x)$ and $r_0(x) = b(x)$, and define inductively (while $r_k(x) \neq 0$)

$$\begin{aligned} q_k(x) &= r_{k-1}(x)//r_k(x) \\ r_{k+1}(x) &= r_{k-1}(x)\%r_k(x), \quad \text{so that} \\ r_{k-1}(x) &= q_k(x)r_k(x) + r_{k+1}(x). \end{aligned}$$

Rearranging these equalites by solving for $r_{k+1}(x)$, we get a sequence

$$\begin{aligned} r_1(x) &= r_{-1}(x) - r_0(x)q_0(x) = a(x) - b(x)q_0(x) \\ r_2(x) &= r_0(x) - r_1(x)q_1(x) \\ r_3(x) &= r_1(x) - r_2(x)q_2(x) \\ &\vdots \\ r_{k+1}(x) &= r_{k-1}(x) - q_k(x)r_k(x) \\ &\vdots \\ r_{n+1}(x) &= r_{n-1}(x) - q_n(x)r_n(x) = 0. \end{aligned}$$

We note that $\deg(r_k(x))$ is a strictly decreasing sequence of nonnegative integers. The process must terminate: for some n, $r_{n+1}(x) = 0$. Now we make use of the following lemma, proved using basic divisibility properties.

Lemma 1.3.4. *For polynomials $a(x), b(x), c(x), s(x)$ in $F[x]$ that satisfy $a(x) = b(x)s(x) + c(x)$, we have $\gcd(a(x), b(x)) = \gcd(b(x), c(x))$.*

Let's apply this to the sequence $r_k(x)$, letting n be minimal such that $r_{n+1}(x) = 0$. We have (since $r_{n+1}(x) = 0$)

$$\gcd(a(x), b(x)) = \gcd(b(x), r_1(x)) = \cdots = \gcd(r_n(x), r_{n+1}(x)) = \gcd(r_n(x), 0) = r_n(x)$$

This argument shows that the Euclidean algorithm produces the gcd of $a(x)$ and $b(x)$.

Exercises 1.3.5. Use the Euclidean algorithm to express the greatest common divisor of these two polynomials as a polynomial combination of them.

(a) $f = x^4 + x^2$ and $g = x^3 + 1$.

(b) $f = x^6 + 1$ and $g = x^4 + x^3 + x^2 + 1$ as elements of $\mathbb{F}_2[x]$.

Primes, Irreducibles and Unique Factorization

Let $r(x) \in F[x]$ have degree at least one. As with integers, we say $r(x)$ is **irreducible** when it can't be factored in a nontrivial way: whenever $r(x) = a(x)b(x)$ either $a(x)$ or $b(x)$ is a constant (that is in F). We say $r(x)$ is **prime** when $r(x)|a(x)b(x)$ implies $r(x)|a(x)$ or $r(x)|b(x)$.

As with integers, we have the equivalence of primality and irreducibility, which is a key step towards proving unique factorization.

Theorem 1.3.6 (Prime-Irreducible)**.** *Any nonconstant element of $F[x]$ is irreducible if and only if it is prime.*

Proof. Let $r(x) \in F[x]$ be irreducible; let us show it is prime. Suppose that $p(x)|a(x)b(x)$ for some $a(x)$ and $b(x)$ in $F[x]$. We need to show $p(x)|a(x)$ or $p(x)|b(x)$. If $p(x)$ divides $a(x)$ we are done, so suppose it does not divide $a(x)$. Since $p(x)$ is irreducible, the GCD of $a(x)$ and $p(x)$ is 1. By the GCD Theorem, there are polynomials $u(x), v(x)$ such that

$$1 = a(x)u(x) + p(x)v(x)$$

Multiplying by $b(x)$

$$b(x) = a(x)b(x)u(x) + p(x)b(x)v(x)$$

Since $p(x)|a(x)b(x)$ we have that $p(x)$ divides the right hand side, and consequently $p(x)$ divides $b(x)$. Thus we have shown that $p(x)$ is prime.

Suppose now that $p(x)$ is prime; we will show it is irreducible. Let $p(x) = a(x)b(x)$ be a factorization of $p(x)$. We must show one of $a(x)$ or $b(x)$ is a constant. Since $p(x)$ is prime and it divides (in fact equals) the product $a(x)b(x)$ it must divide one of the factors. Without loss of generality, say $p(x)|a(x)$. We then have

$\deg(a(x)) \geq \deg(p(x))$. On the other hand, since $p(x) = a(x)b(x)$ we have by additivity of degrees for a product of polynomials,

$$\deg(p(x)) = \deg(a(x)) + \deg(b(x)) \geq \deg(a(x))$$

We conclude that $\deg(p(x)) = \deg(a(x))$ and $\deg(b(x)) = 0$. Thus $b(x)$ is a constant.

□

Theorem 1.3.7 (Unique Factorization)**.** *Let $a(x) \in F[x]$ be nonzero. There is a nonnegative integer t, a constant $u \in F$, distinct monic irreducible polynomials $p_1(x), \ldots, p_t(x)$, and positive integers $e_1, \ldots, e_t$ such that*

$$a(x) = u\big(p_1(x)\big)^{e_1} \cdots \big(p_t(x)\big)^{e_t}$$

Each of t, $p_i(x)$, e_i and u is uniquely determined, up to reordering of the $p_i(x)^{e_i}$.

Proof. TBD □

Polynomial modulus

We now extend the technique of modular arithmetic to the polynomial ring over a field.

Let $m(x) \in F[x]$ have degree d. Polynomials $a(x)$ and $b(x)$ are **congruent modulo** $m(x)$ when $m(x)$ divides $a(x) - b(x)$.

Theorem 1.3.8. *Congruence modulo $m(x)$ is an equivalence relation. The set of polynomials of degree less than $\delta = \deg(m(x))$ is a system of representatives for congruence modulo $m(x)$. That is, each polynomial is congruent modulo $m(x)$ to its remainder when divided by $m(x)$, which has degree less than δ.*

Any constant multiple of $m(x)$ will define the same equivalence relation as $m(x)$ so we usually take $m(x)$ to be monic. We will write $[a(x)]_{m(x)}$ for the congruence class of $a(x)$ modulo $m(x)$ whenever we need to be careful to distinguish between $a(x)$, otherwise we will omit the subscript if the modulus is obvious.

Theorem 1.3.9 (Arithmetic modulo $m(x)$)**.** *Suppose that $a(x) \equiv b(x) \bmod m(x)$ and $r(x) \equiv s(x) \bmod m(x)$. Then $a(x) + r(x) \equiv b(x) + s(x) \bmod m(x)$ and $a(x)r(x) = b(x)s(x) \bmod m(x)$. Thus, arithmetic on congruence classes modulo $m(x)$ is well-defined.*

- $[a(x)] + [r(x)] = [a(x) + r(x)]$
- $[a(x)] * [r(x)] = [a(x)r(x)]$

Proof. Suppose that $a(x) \equiv b(x) \bmod m(x)$ and $r(x) \equiv s(x) \bmod m(x)$. We have $a(x) = b(x) + u(x)m(x)$ and $r(x) = s(x) + v(x)m(x)$ for some polynomials $u(x)$ and $v(x)$. Then

$$\begin{aligned} &a(x) + r(x) = b(x) + s(x) + \big(u(x) + v(x)\big)m(x), \qquad \text{so} \\ &a(x) + r(x) \equiv b(x) + s(x) \bmod m(x) \end{aligned}$$

We also have

$$a(x)r(x) = b(x)s(x) + \Big(a(x)v(x) + b(x)u(x) + u(x)v(x)m(x)\Big)m(x)$$

so $a(x)r(x) \equiv b(x)s(x) \bmod m(x)$.

This shows that no matter what element of a congruence class is used to represent a class, arithmetic operations modulo $m(x)$ will give the same result. □

We will write $F[x]/m(x)$ for the set of equivalence classes modulo $m(x)$, with the operations $+$ and $*$ as designated above.

Exercises 1.3.10.

(a) Let $m(x)$ be an irreducible monic polynomial in $F[x]$. Let $[a(x)] \in F[x]/m(x)$ with $a(x)$ not divisible by $m(x)$. Use the GCD Theorem 1.3.3 to show there is some $r(x) \in F[x]$ such that $[a(x)][r(x)] = [1]$.

(b) Extend this result, partially, to other $F[x]/m(x)$. If $a(x) \in F[x]$ is such that $a(x)$ is coprime to $m(x)$, then there is some $r(x) \in F[x]$ such that $[a(x)][r(x)] = [1]$.

Chapter 2

Groups

2.1 Groups and Subgroups

The material in this section is a quick summary of the most fundamental properties of groups. I have omitted some proofs because they are are fairly routine, are good exercises for the reader, and are available from many sources. See in particular the book of Hungerford [Hun12]. It is worthwhile reviewing the proofs as you read!

First let us recall the definition of a group.

Definition 2.1.1. A **group** is a set G with a binary operation $*$ and a unary operation denoted $a \longmapsto a^{-1}$ satisfying the following properties.

(1) Associativity of $*$: For all $a, b, c \in G$, $(a * b) * c = a * (b * c)$.

(2) Identity for $*$: There is an element, usually denoted e, such that $e * a = a = a * e$ for all $a \in G$.

(3) Inverses for $*$: For each $a \in G$ the unary operation $a \longmapsto a^{-1}$ gives the multiplicative inverse for a. That is, $a * a^{-1} = e = a^{-1} * a$.

A group which also satisfies $a * b = b * a$ is called **commutative** or **abelian** (after the mathematician Abel).

The operation is usually called a product. So $a * b$ is the product of a and b. One must be careful, because, in a nonabelian group, the products $a * b$ and $b * a$ are not necessarily equal. In abelian groups the operation is often called addition and is written with a + sign, while the identity is written as 0.

The most basic properties are contained in the following proposition. The proofs of all of these are called "card tricks" by a friend of mine. Any algebraist should have these up a sleeve since similar cleverness is used in other contexts.

Proposition 2.1.2 (Basic Properties)**.** *Let $G, *$ be a group. Then*

(1) The identity element is unique.

(2) The inverse of any element is unique.

(3) The cancellation law holds: $a * b = a * c$ implies $b = c$ (and similarly for cancellation on the right).

(4) If $a * g = g$ for some $g \in G$, then $a = e_G$.

(5) $(a * b)^{-1} = b^{-1} * a^{-1}$.

(6) $(a^{-1})^{-1} = a$.

Let G be a group. When there is risk of confusion, with more than one group under consideration, we will use $*_G$ for the operation on the group G and e_G for the identity element. Otherwise we will not subscript with G. Unless there is some reason to be very clear we rarely write the group operation: $g_1 g_2$ means $g_1 *_G g_2$. For a positive integer n, g^n is shorthand for $\underbrace{gg\cdots g}_{n \text{ factors}}$ and g^{-n} is shorthand for $\underbrace{g^{-1}g^{-1}\cdots g^{-1}}_{n \text{ factors}}$. It is straightforward to check that the usual rules for exponents apply.

For a group in which the operation is $+$ and the identity is 0 (in particular, the group must be abelian), the sum $\underbrace{g+g+\cdots+g}_{n \text{ terms}}$ is written ng. Think of this as repeated addition, not as multiplication: the group just has one operation, and n is an integer, not necessarily an element of the group.

The first examples come from a familiar place, the integers, as discussed in Section 1.2.

Example 2.1.3. The integers $\mathbb{Z}$ form a group with operation $+$, identity element 0 and inversion operation $a \longmapsto -a$. The elements 1 and -1 generate the group $\mathbb{Z}$ in the sense that by applying inversion and repeated addition we can get all the other elements of $\mathbb{Z}$. This is not true for other elements.

The set of multiples of n in $\mathbb{Z}$ also is a group under $+$ with identity 0. Adding two multiples of n gives another multiple of n, and the additive inverse of a multiple of n is also a multiple of n. We will denote this group by $n\mathbb{Z}$. Later in this section we introduce the abstract definition of a subgroup. Here $n\mathbb{Z}$ is a subgroup of $\mathbb{Z}$. The elements n and $-n$ generate $n\mathbb{Z}$ just as 1 and -1 generate $\mathbb{Z}$.

The integers modulo n, introduced in Section 1.1, also form a group under $+$. We can write the elements as $0, 1, 2, \ldots, n-1$, but these are really shorthand for

congruence classes. When consider the integers modulo n as a group (ignoring multiplication) we will write it as $\mathbb{Z}_n$. An interesting question is: what elements generate $\mathbb{Z}_n$?

Definition 2.1.4. A single element g of a group G **generates** G when any element of G is equal to g^n for some $n \in \mathbb{Z}$ (or, if the operation of the group is addition, any element is equal to ng for some n). Such a group is said to be **cyclic**.

Let S be a subset of a group G. We say that G **is generated** by S if any element of G is equal to a product (with an arbitrary number of terms) of elements of S and elements of $\left\{s^{-1} : s \in S\right\}$.

A group is $t-$**generated** if there is a subset S of G with t elements that generates G.

Example 2.1.5. The Cartesian product $\mathbb{Z} \times \mathbb{Z}$ is a group under coordinatewise addition with identity element $(0, 0)$ and inverse operator $(a, b) \longmapsto (-a, -b)$. It is not possible to generate all elements by repeated addition of a single element, but this group is $2-$generated.

The Cartesian product $\mathbb{Z}/m \times \mathbb{Z}/n$ under coordinatewise addition (and using coordinatewise identity elements and inversion) is also a group. In certain cases it is possible to have a single element generate all elements by repeated addition.

Exercises 2.1.6.

(a) Consider $\mathbb{Z}_n$ for $n = 2, 3, 4, 5, 6, 7$. Which elements $a \in \mathbb{Z}_n$ generate all of $\mathbb{Z}_n$?

(b) Show that $\mathbb{Z} \times \mathbb{Z}$ is not generated by a single element.

(c) Experiment with some small integers m and n to find cases in which $\mathbb{Z}_m \times \mathbb{Z}_n$ is generated by a single element and other cases in which it is not.

Definition 2.1.7. The **cardinality** of a group G, written $|G|$ is just the cardinality of the underlying set. It is also called the **order** of the group.

Many of our examples will be finite groups and we will be studying some of the properties that go into understanding the structure of finite groups.

Example 2.1.8. Recall the definition of the dihedral group D_n, which is the symmetry group of a regular n-gon (1.2.4). We showed that D_n has $2n$ elements and that it is 2-generated—by r, the rotation by $2\pi/n$, and any reflection, t_i. (Well, we did this for D_5, but the same argument holds).

Section 1.2 also introduced S_n, the group of all possible permutations on $\{1, 2, \ldots, n\}$. This group has cardinality $n!$. Composing them is just composing functions. The inverse permutation involves flipping the two rows of the permutation and then, for convenience, rearranging the columns so that the first row is

increasing order.

$$\text{For} \qquad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 1 & 5 & 2 \end{pmatrix}, \qquad \text{the inverse is } \pi^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 1 & 4 \end{pmatrix}.$$

The Order of an Element

Definition 2.1.9. For $g \in G$ the **order of the element** g is the smallest positive integer n such that $g^n = e$, if such an n exists. If no such n exists then g has infinite order. We use $|g|$ or $\text{ord}(g)$ for the order of g.

The **exponent** of G is the least common multiple of the orders of the elements of G, if such an integer exists, that is $\exp(G) = \text{lcm}\,\{\text{ord}(g) : g \in G\}$. If no such element exists one can say the exponent is infinite.

Only the identity element of a group has order 1. Every nonzero element of $\mathbb{Z}$ has infinite order. In $\mathbb{Z}_n$ some elements have order n, but other non-identity elements may have a different order. For any finite group there is a well defined exponent, but an infinite group may or may not have one.

Theorem 2.1.10 (Order Theorem)**.** *Let g be an element of the group G.*

(1) If g has infinite order, then elements g^t for $t \in \mathbb{Z}$ are all distinct.

(2) If g has order n then

(a) $g^t = g^s$ if and only if $t \equiv s \mod n$. In particular, $g^t = e$ if and only if n divides t.

(b) $\text{ord}(g^t) = \frac{n}{\gcd(t,n)}$.

Proof. Suppose $g^t = g^s$ for integers s, t. Then $g^{t-s} = e_G$. If g has infinite order then $s - t = 0$ so $s = t$. This proves item (1).

Suppose g has order n. For an integer t use the quotient remainder theorem to write $t = nq + r$. Then $g^t = g^{nq+r} = (g^n)^q g^r = e^q g^r = g^r$. This shows claim (2a). Now let $d = \gcd(t, n)$ and write $t = da$ and $n = db$. Then a, b have no common factor (otherwise d would not be the gcd) and we observe that $b = n/\gcd(t, n)$. We now have $(g^t)^b = g^{dab} = (g^n)^a = e$. Furthermore, if $(g^t)^s = e$ then, by (2a), $n = db$ divides $ts = das$. Cancelling d and taking note of b and a being coprime we get b divides s. This establishes (2b). □

Example 2.1.11. The reflections in D_n all have order 2. The rotation by $2\pi/n$ has order n, but some of the other rotations has order less than n. For example in D_6 with r the rotation by $2\pi/6$, r^2 has order 3 and r^3 has order 2.

Exercises 2.1.12. More card tricks.

(a) Suppose that every element of G has order 2. Show that G is abelian.

(b) If G has even order then G has an element of order 2. (Consider the pairing of g with g^{-1}).

Exercises 2.1.13. Order and commutativity.

(a) If $g \in G$ has order m and $h \in H$ has order n, find the order of $(g, h) \in G \times H$.

(b) Suppose that $a, b \in G$ commute (that is $ab = ba$). If $\text{ord}(a)$ and $\text{ord}(b)$ are coprime find the order of ab.

(c) Let A be an abelian group with finite exponent. Show that there is some $a \in A$ such that $\text{ord}(a) = \exp(A)$.

Subgroups

A key area of investigation in many mathematical subjects is the subsets of a given object that have useful structure. In this section we treat subsets of a group that are themselves groups.

Definition 2.1.14. A nonempty subset H of a group G is a **subgroup**, when H is a group using the operation $*_G$ on G. This means that for any $h \in H$ the inverse h^{-1} must also be in H and for $h, h' \in H$, the product $h *_G h'$ must also be an element of H. We will say H is **closed under inversion and multiplication**. We will write $H \leq G$ when H is a subgroup of G (as opposed to $H \subseteq G$ when H is just a subset), and $H < G$ when H is a proper subgroup (that is $H \neq G$).

The following proposition is a sanity check on our definition of subgroup: the identity element and inversion are the same for the subgroup as for the group.

Proposition 2.1.15. *If K is a subgroup of G then $e_K = e_G$ and the inversion operation is the same on K as it is on G.*

Proof. If K is a subgroup of G then it must have an identity element. For any $k \in K$, we have (using $*_K = *_G$) that $e_K *_K k = e_K *_G k = k$ Proposition 2.1.2 (4) shows that it must be the case that $e_K = e_G$. (If something acts like the identity it is the identity!)

Let $k \in K$ and let k^{-1} be its inverse in G. This k^{-1} is also the inverse of k in the subgroup K because $k * k^{-1} = e_G = e_K$. □

(Thank goodness for both of these facts.)

Proposition 2.1.16 (Subgroup Properties)**.** *If K is a nonempty subset of G that is closed under inversion and closed under multiplication in G then K is a subgroup of G (i.e. it also contains e_G).*

If K is a nonempty subset of G such that $h *_G k^{-1} \in K$ for all $h, k \in K$ then K is a subgroup of G.

Proof. Since K is nonempty, it contains some element k. Since K is closed under inversion, $k^{-1} \in K$. Since K is closed under multiplication, $k *_G k^{-1} = e_G \in K$. Thus K satisfies the definition of a group since it has associativity (immediate since $*_K$ is the restriction of $*_G$), an identity element, and inverses (by assumption).

Suppose K is a nonempty subset of G such that $h *_G k^{-1} \in K$ for all $h, k \in K$. For any $k \in K$, setting $h = k$ gives $k *_G k^{-1} = e_G \in K$. Letting $h = e_G$ gives $e_G * k^{-1} = k^{-1} \in K$, so K is closed under inversion. Now for any $h, k \in K$ we know $k^{-1} \in K$, so $h *_G (k^{-1})^{-1} = h *_G k \in K$. This shows K is closed under multiplication. □

When proving that a particular subset of a group is a subgroup one can either show the set is closed under inversion and under multiplication, or use the second property of the theorem. I like the clarity of proving closure under each operation.

Exercises 2.1.17.

(a) Let G be a group and g an arbitrary element of G. Show that $\left\{g^i : i \in \mathbb{Z}\right\}$ is a subgroup of G. This group is called the **cyclic subgroup generated by** g and is written $\langle g \rangle$.

(b) Let G be a group. Show that the set $Z(G) = \{a \in G : ag = ga \text{ for all } g \in G\}$ is an abelian subgroup of G. It is called the **center** of G.

(c) Let H and K be subgroups of G. Show that their intersection is also a subgroup of G.

Note that there is consistency between the order of an element and the order of the subgroup it generates. If $g \in G$ has order n then the set of powers of g is $\left\{g^0 = e_G, g, g^2, \ldots, g^{n-1}\right\}$ (any other power of g is one of these). This set is a subgroup of G of order n. If $g \in G$ has infinite order then it generates a cyclic subgroup that is infinite. (It is not really cycling in the infinite case, but that's the term used!)

Example 2.1.18. As was discussed in Section 1.2, the integers, $\mathbb{Z}$, the rational numbers, $\mathbb{Q}$, the real numbers, $\mathbb{R}$, and the complex numbers, $\mathbb{C}$, are all abelian groups under addition. We sometimes write $\mathbb{Z}, +$ (for example) to emphasize that we are are ignoring multiplication, and are just considering the additive properties of $\mathbb{Z}$. Clearly we have a sequence of subgroups.

$$\mathbb{Z} < \mathbb{Q} < \mathbb{R} < \mathbb{C}$$

Example 2.1.19. We showed in Section 1.2 that D_5 may be identified with a subgroup of S_5. The discussion can be generalized to show D_n identified as a subgroup S_n for any $n \geq 3$.

The permutation group S_4 may be also seen as contained in S_5; it is just the set of all permutations in S_5 that take 5 to 5 (we say these "fix" 5). There are other subgroups that have the exact same structure as S_4. For example the set of all permutations in S_5 that fix 3. These subgroups are all isomorphic (see Section 2.4).

The Lattice of Subgroups

Consider a fixed group G and let S be the set of all subgroups of G. For small groups, it is often illuminating to draw a diagram showing all the subgroups and the containment relationships among them. There is no simple and efficient process for this in general, but for our small examples one can do the following.

(1) Find all subgroups generated by 1 element.

(2) Find all subgroups generated by 2 elements by adding an new element to the 1-generated subgroups.

(3) See if there are subgroups that require 3 generators by adding an element to the 2-element subgroups.

Note that there are usually several ways to generate a particular subgroup.

Here are some examples.

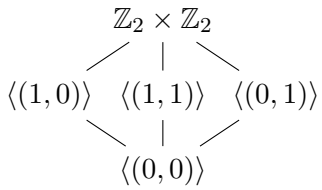


Figure 2.1: The lattice diagram for $\mathbb{Z}_2 \times \mathbb{Z}_2$.

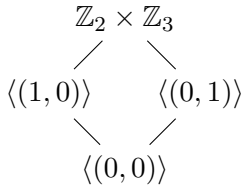


Figure 2.2: The lattice diagram for $\mathbb{Z}_2 \times \mathbb{Z}_3$.

Figure 2.3: The lattice diagram for $\mathbb{Z}_4$.

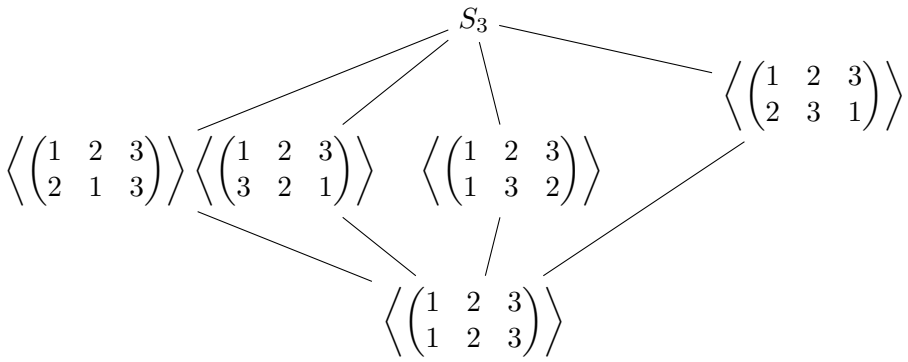


Figure 2.4: The lattice diagram for S_3.

Exercises 2.1.20. Lattice Diagrams for Groups

(a) Draw the subgroup lattice diagram for $\mathbb{Z}_{45}$.

(b) Draw the subgroup lattice diagram for $\mathbb{Z}_{60}$.

(c) Draw the subgroup lattice diagram for $\mathbb{Z}_2 \times \mathbb{Z}_4$.

(d) Draw the subgroup lattice diagram for $\mathbb{Z}_3 \times \mathbb{Z}_4$.

(e) Find all subgroups of $\mathbb{Z}_4 \times \mathbb{Z}_4$. Describe the logic of your process for finding them. Present them in an organized fashion. Draw the lattice if you can.

2.2 Homomorphisms

Definition 2.2.1. Let G and H be groups. A function $\varphi : G \longrightarrow H$ is a homomorphism when

(1) $\varphi(g_1 *_G g_2) = \varphi(g_1) *_H \varphi(g_2)$ for all $g_1, g_2 \in G$, and

(2) $\varphi(e_G) = e_H$, and

(3) $\varphi(g^{-1}) = (\varphi(g))^{-1}$ for all $g \in G$.

I like to speak informally about a homomorphism as a function that **respects structure**: A homomorphism of groups "respects" the property of the identity element, multiplication, and inversion.

It is fairly easy to show that the first item in the definition of homomorphism implies the other two. This result and another important result are contained in the following proposition.

Proposition 2.2.2 (Homomorphisms)**.** *Let G, H, K be groups.*

If $\varphi : G \longrightarrow H$ is a function such that $\varphi(g_1 *_G g_2) = \varphi(g_1) *_H \varphi(g_2)$ then φ is a group homomorphism.

If $\varphi : G \longrightarrow H$ and $\theta : H \longrightarrow K$ are group homomorphisms then the composition $\theta \circ \varphi$ is also a group homomorphism.

Exercises 2.2.3. Prove the proposition.

(a) Assuming that φ respects multiplication, show that it also takes the identity of G to the identity of H (use $e_G * e_G = e_G$) and that it respects inversion (use $gg^{-1} = e_G$).

(b) Prove that the composition of homomorphisms is a homomorphism.

The simplest type of a homomorphism is the **inclusion** of a subgroup H of G into G. That is, when $H < G$, then there is a function $H \longrightarrow G$ that takes elements of H to themselves (now thought of as elements of G).

Proposition 2.2.4. *Let g be an element of a group G. There is homomorphism*

$$\begin{aligned} \varphi : \mathbb{Z} &\longrightarrow G \\ t &\longmapsto g^t \end{aligned}$$

When $g \in G$ has infinite order, this homomorphism is injective.

When $g \in G$ has order n, the function below is an injective homomorphism.

$$\begin{aligned} \varphi : \mathbb{Z}_n &\longrightarrow G \\ t &\longmapsto g^t \end{aligned}$$

Proof. Let g be an element of a group G. Consider the function $\mathbb{Z} \longrightarrow G$ taking t to g^t. The function is a homomorphism because for $s, t \in \mathbb{Z}$, $s + t$ maps to $g^{s+t} = g^s * g^t$ and this is the product of the images of s and t.

Now assume that g has infinite order. Suppose that $g^s = g^t$. Then $g^{s-t} = e_G$. Since g has infinite order $s - t = 0$, so $s = t$. This proves injectivity.

The proof for g of finite order is similar. □

Example 2.2.5. The function $\mathbb{Z} \longrightarrow \mathbb{Z}_n$ that takes k to its equivalence class modulo n is a surjective homomorphism.

Proposition 2.2.6. *Let $G \longrightarrow H$ be a homomorphism. The image of G, which we write $\varphi(G)$, is a subgroup of H.*

Proof. We need only show that the image is closed under inversion and multiplication. Consider an element of $\varphi(G)$. Since it is the image of φ we may write it as $\varphi(g)$ for some $g \in G$. By the properties of a homomorphism

$$\varphi(g^{-1})\varphi(g) = \varphi(g^{-1}g) = \varphi(e_G) = e_H$$

This shows that the inverse of $\varphi(g)$ is $\varphi(g^{-1})$ and therefore that it is in the image of φ.

Consider two elements of $\varphi(G)$, which we may write as $\varphi(g_1)$ and $\varphi(g_2)$. Their product is $\varphi(g_1)\varphi(g_2) = \varphi(g_1 g_2)$, and this is in the image of φ. We have shown that $\varphi(G)$ is closed under inversion and under multiplication so it is a subgroup of G. □

Isomorphisms

A homomorphism φ that is also a bijection (one-to-one and onto) is called an **isomorphism**. When there exists an isomorphism $\varphi : G \longrightarrow H$ we say G and H are **isomorphic** and write $G \cong H$. The following proposition shows that the relation of being isomorphic satisfies symmetry and transitivity so it determines an equivalence relation on any set of groups.

Proposition 2.2.7 (Isomorphisms)**.** *If φ is an isomorphism of groups, then the inverse function φ^{-1} is also an isomorphism of groups.*

If $\varphi : G \longrightarrow H$ and $\theta : H \longrightarrow K$ are group isomomorphisms then the composition $\theta \circ \varphi$ is also a group isomorphism.

On any set of groups $\mathcal{G}$, the relation of being isomorphic is an equivalence relation.

Proof. Let $\varphi : G \longrightarrow H$ be an isomorphism of groups. By definition, φ is both injective and surjective, so there is a well defined inverse function, φ^{-1}. We must show that φ^{-1} is a homomorphism.

Let h_1 and h_2 be two elements of h. Since φ is surjective there are two elements g_1 and g_2 such that $\varphi(g_1) = h_1$ and $\varphi(g_2) = h_2$. Since φ is injective these two elements are uniquely defined. We now show that φ^{-1} respects products, which is enough to show it is a homomorphism.

$$\begin{aligned}
\varphi^{-1}(h_1h_2) &= \varphi^{-1}\left(\varphi(g_1)\varphi(g_2)\right) \\
&= \varphi^{-1}\left(\varphi(g_1g_2)\right) \qquad \text{since } \varphi \text{ is a homomorphism} \\
&= g_1g_2 \qquad \text{since } \varphi \text{ and } \varphi^{-1} \text{ are inverse functions} \\
&= \varphi^{-1}(h_1)\varphi^{-1}(h_2)
\end{aligned}$$

The composition of two bijections is a bijection, and by Proposition 2.2.6 we know that the composition of homomorphisms is a homomorphism. Thus the composition of two isomorphisms is an isomorphism.

Clearly any group is isomorphic to itself under the identity map. The first part of this proposition shows that if G is isomorphic to H then H is also isomorphic to G. The second part establishes transitivity. □

If there is an isomorphism between groups G and H then G and H have the same algebraic structure, so we consider them equivalent.

Exercises 2.2.8.

(a) Show that for each $a \in \mathbb{Z}_n$ there is a unique homomorphism $\varphi_a : \mathbb{Z}_n \longrightarrow \mathbb{Z}_n$ such that $\varphi_a(1) = a$.

(b) Under what conditions on a is φ_a an isomorphism?

(c) Identify all subgroups of $\mathbb{Z}_n$.

Exercises 2.2.9.

(a) Show that if $d \mid n$ then there is a homomorphism $\mathbb{Z}_n$ to $\mathbb{Z}_d$ that takes 1 (in $\mathbb{Z}_n$) to 1 (in $\mathbb{Z}_d$). Show that it is also surjective.

(b) Show that if d does not divide n then there is no homomorphism $\mathbb{Z}_n$ to $\mathbb{Z}_d$ that takes 1 (in $\mathbb{Z}_n$) to 1 (in $\mathbb{Z}_d$).

Exercises 2.2.10.

(a) Show that there is an injective homomorphism from $\mathbb{Z}_n$ into D_n taking 1 to rotation by $2\pi/n$.

(b) How many injective homomorphisms are there from $\mathbb{Z}_n$ into D_n?

(c) Identify all subgroups of D_n for $n = 3, 4, 5, 6$. Draw a lattice diagram showing containment of subgroups.

Exercises 2.2.11. A Perverse Group

(a) Show that $\mathbb{Z}$ is a group under the operation $\square$ defined by $a \square b = a + b - 2$. (What is the identity element? What is the inverse of an element a?)

(b) Find an isomorphism from $\mathbb{Z}, +$ to $\mathbb{Z}, \square$.

Exercises 2.2.12. Prove the following results about the relationship between the order of an element and the order of its image under a homomorphism.

(a) If $\varphi : G \longrightarrow H$ is a homomorphism, then $\text{ord}(\varphi(g))$ divides $\text{ord}(g)$.

(b) If $\varphi : G \longrightarrow H$ is an isomorphism, then $\text{ord}(\varphi(g)) = \text{ord}(g)$.

The previous exercises give important restrictions on homomorphisms. If you want to create a homomorphism from G to H, each element g in G must go to an element of H that has order dividing $\text{ord}(g)$.

Exercises 2.2.13.

(a) Show that there is a nontrivial homomorphism from D_3 to $\mathbb{Z}_2$ but that any homomorphism from D_3 to $\mathbb{Z}_3$ is trivial.

2.3 Some Constructions of Groups

In this section we show three ways to construct new groups from ones that we already have. All three have been touched on briefly; we give more detail here. The first construction really has two aspects, given some subgroups of a group we can form the intersection of the subgroups and also the join of the subgroups. More broadly, given a subset of a group we can define the subgroup generated by the subset. The second construction is the direct product of groups. The idea is simple (and was illustrated in the first examples of groups that we gave). Given two groups, form the Cartesian product as sets, and apply componentwise operations to get a new group. The construction yields in a natural way two types of homomorphisms that are important despite their simplicity. The direct product can also be applied to several groups not just two. Finally, for a given group G the isomorphisms of G to itself have a group structure that is useful in understanding the properties of G.

Intersections and the Join of Subgroups

Proposition 2.3.1. *Let $H_1, \ldots, H_t$ be subgroups of G. The intersection $\bigcap_{i=1}^{t} H_i$ is a subgroup of G.*

More generally, if $\mathcal{H}$ is a set of subgroups of G then $\bigcap_{H\in\mathcal{H}} H$ is a subgroup of G.

Proof. Suppose h and h' are in $\bigcap_{H\in\mathcal{H}} H$. Since each $H \in \mathcal{H}$ is a subgroup of G, $h'h^{-1} \in H$ for each $H \in \mathcal{H}$. Thus $h'h^{-1} \in \bigcap_{H\in\mathcal{H}} H$, and Proposition 2.1.16 shows $\bigcap_{H\in\mathcal{H}} H$ is a subgroup of G. □

Let S be an arbitrary subset of a group G. Let $\mathcal{H}$ be the set of all subgroups of G containing S. Then $\bigcap_{H\in\mathcal{H}} H$ is a subgroup of G, and it contains S, since each $H \in \mathcal{H}$ contains S. Furthermore, any subgroup K of G containing S is in $\mathcal{H}$ so $\bigcap_{H\in\mathcal{H}} H \subseteq K$. This argument justifies the following definition.

Definition 2.3.2. Let G be a group and let S be a subset of G. By $\langle S\rangle$ we mean the **smallest subgroup** of G containing S. It is the intersection of all subgroups of G containing S. We say $\langle S\rangle$ is the subgroup of G **generated** by S.

If H and K are subgroups of G their **join**, written $H \vee K$, is $\langle H \cup K\rangle$.

One interesting problem is to find minimal size sets that generate a group. For example, the elements 1 and -1 both generate $\mathbb{Z}$. The element 1 generates $\mathbb{Z}_n$ as does any $a \in \mathbb{Z}_n$ that is coprime to n. (Exercise 2.1.6.)

The Direct Product

Definition 2.3.3. Let G and H be groups. The Cartesian product $G \times H$, along with the unary operation (of inversion) and the binary operation (of multiplication) below form the **direct product** of G and H.

$$(g, h)^{-1} = (g^{-1}, h^{-1})$$
$$(g_1, h_1) *_{G\times H} (g_2, h_2) = (g_1 *_G g_2, h_1 *_H h_2)$$

The identity element is of course (e_G, e_H).

The following proposition shows that the direct product is in fact a group and gives other important properties. None of the following are surprising and they are routine to prove.

Theorem 2.3.4 (Direct Product)**.** *Let G and H be groups.*

(1) The above definition does, indeed, make $G \times H$ a group.

(2) The associative law for the product of several groups holds: $G_1 \times (G_2 \times G_3) \cong (G_1 \times G_2) \times G_3$.

(3) $G \times H$ *is abelian if and only if* G *and* H *are abelian.*

(4) If G' *is a subgroup of* G *and* H' *is a subgroup of* H *then* $G' \times H'$ *is a subgroup of* $G \times H$. *In particular* $G \times \{e_H\}$ *and* $\{e_G\} \times H$ *are subgroups of* $G \times H$.

(5) There is an injective homomorphism $i_G : G \longrightarrow G \times H$ *taking* g *to* (g, e_H) *(and similarly* $i_H : H \longrightarrow G \times H$*).*

(6) The projection maps $p_G : G \times H \longrightarrow G$ *and* $p_H : G \times H \longrightarrow H$ *are surjective homomorphisms.*

(7) The construction and the observations above can be generalized to the direct product of any set of groups $\{G_i : i \in I\}$ *indexed by a finite set* I. *(It extends to infinite index sets* I *with some modification due to subtle issues.)*

The following result is more subtle and it turns out to be a powerful idea.

Proposition 2.3.5 (Universal Property of the Product)**.** *Let* G, H, *and* T *be groups, and let* $\varphi : T \longrightarrow G$ *and* $\psi : T \longrightarrow H$ *be homomorphisms. The function* $\alpha : T \longrightarrow G \times H$ *defined by* $t \longmapsto (\varphi(t), \psi(t))$ *is a homomorphism. It is the unique homomorphism such that* $p_G \circ \alpha = \varphi$ *and* $p_H \circ \alpha = \psi$.

Proof. Note first that, by construction, $p_G \circ \alpha = \varphi$ and $p_H \circ \alpha = \psi$. Furthermore, there is no other choice for the definition of α that satisfyies these to requirements.

We have to show that α respects inversion and multiplication. Let $t \in T$. We have to show that $\alpha(t^{-1})$ is the inverse of $\alpha(t)$. The subscript on $*$ that we sometimes use to show the group being used is omitted in the following derivation, but it is worthwhile to identify it while reading.

$$\begin{aligned}
\alpha(t) * \alpha(t^{-1}) &= \Big(\varphi(t), \psi(t)\Big) * \Big(\varphi(t^{-1}), \psi(t^{-1})\Big) \\
&= \Big(\varphi(t) * \varphi(t^{-1}), \psi(t) * \psi(t^{-1})\Big) \\
&= \Big(\varphi(t * t^{-1}), \psi(t * t^{-1})\Big) \\
&= (e_H, e_K)
\end{aligned}$$

This proves that $\alpha(t^{-1})$ is the inverse of $\alpha(t)$. Similarly for $t_1, t_2 \in T$,

$$\begin{aligned}\alpha(t_1) * \alpha(t_2) &= \Big(\varphi(t_1), \psi(t_1)\Big) * \Big(\varphi(t_2), \psi(t_2)\Big) \\ &= \Big(\varphi(t_1) * \varphi(t_2), \psi(t_1) * \psi(t_2)\Big) \\ &= \Big(\varphi(t_1 * t_2), \psi(t_1 * t_2)\Big) \\ &= \alpha(t_1 * t_2)\end{aligned}$$

This shows α respects products. □

Exercises 2.3.6.

(a) Not all subgroups of $G \times H$ are direct products of subgroups of G and H. Illustrate with some examples: $\mathbb{Z}_2 \times \mathbb{Z}_2$, $\mathbb{Z}_4 \times \mathbb{Z}_4$.

Exercises 2.3.7. We have shown that there is a homomorphism $\mathbb{Z}_n \longrightarrow \mathbb{Z}_d$ taking 1 in $\mathbb{Z}_n$ to 1 in $\mathbb{Z}_d$ if and only if $d \mid n$. Suppose c and d both divide n.

(a) What is the kernel of the homomorphism $\mathbb{Z}_n \longrightarrow \mathbb{Z}_c \times \mathbb{Z}_d$?

(b) Under what conditions is it an isomorphism?

(c) Illustrate with $n = 8$ and $c = d = 4$. What is the image?

(d) Illustrate with $n = 18$ and $c = 6$ and $d = 9$. What is the image?

Automorphism Groups

Proposition 2.3.8 (Automorphisms)**.** *Let G be a group. The set of all isomorphisms from G to itself is a group. This new group is called* $\mathrm{Aut}(G)$*, the group of* **automorphisms** *of G.*

Proof. The identity map id_G is clearly an automorphism of G, so there is at least one automorphism of G. The composition of id_G with any automorphism $\varphi : G \longrightarrow G$ is φ, since the identity map takes each element to itself. Proposition 2.2.7 shows that the inverse of an isomorphism is an isomorphism and the composition of two isomorphisms is an isomorphism. Thus $\mathrm{Aut}(G)$ is a subgroup of the group of bijections of G. □

Exercises 2.3.9.

(a) Show that $\mathrm{Aut}(\mathbb{Z})$ has two elements and $\mathrm{Aut}(\mathbb{Z}) \cong \mathbb{Z}_2$.

(b) Compute $\mathrm{Aut}(\mathbb{Z}_n)$ for $n = 2, 3, 4, 5, 6, 7$. [In each case the answer is a cyclic group.]

(c) Show that $\mathrm{Aut}(\mathbb{Z}_8)$ is not cyclic.

2.4 Permutation Groups

In this section we delve more deeply into the structure of the **symmetric group** S_n, the group of permutations of $\{1, \ldots, n\}$. The number of elements in S_n is $n!$. Informally, we may justify this claim by noting that there are n possible images for the number 1. Once the image for 1 is chosen, there are $n-1$ choices for the number 2. Continuing in this manner we count $n!$ bijections from $\{1, \ldots, n\}$ to itself. One can give a more formal inductive proof.

We will sometimes write an element π of S_n in **tabular form** with i in the top row and $\pi(i)$ in the bottom row.

Exercises 2.4.1.

(a) Here are two elements of S_5:

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

(b) Compute the inverse of each.

(c) Compute the products $\pi\sigma$ and $\sigma\pi$, using the usual convention for compositions: $(\pi\sigma)(i) = \pi(\sigma(i))$. You should see that the results are not equal.

Example 2.4.2. Let $n = 3$, and enumerate the vertices of a triangle clockwise as 1, 2, 3. Each element of D_3 gives rise to a permutation of $\{1, 2, 3\}$.

Let r be rotation clockwise by $2\pi/3$. Then

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

There are three reflections, each fixes one element of $\{1, 2, 3\}$ and transposes the other two

$$u_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \qquad u_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \qquad u_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

This exhausts all permutations of $\{1, 2, 3\}$ so by enumerating the vertices of the triangle we have established a bijection between D_3 and S_3. This is actually an isomorphism since the operation for D_3 is composition, as it is for S_n.

Exercises 2.4.3.

(a) How many ways are there to embed $\mathbb{Z}_4$ in S_4?

(b) How many ways are there to embed D_4 in S_4?

Cycle Decomposition

Definition 2.4.4. Let $a_1, a_2, \ldots, a_t$ be distinct elements of $\{1, \ldots, n\}$. We use the notation $(a_1, a_2, \ldots, a_t)$ to define an element of S_n called a t-**cycle**. This permutation takes a_i to a_{i+1}, for $i = 1, 2, 3 \ldots, t-1$ and it takes a_t to a_1. Every element of $\{1, \ldots, n\} \setminus \{a_1, \ldots, a_t\}$ is fixed (i.e. taken to itself) by the cycle $(a_1, a_2, \ldots, a_t)$. We will call the set $\{a_1, \ldots, a_t\}$ the **support** of the cycle $(a_1, a_2, \ldots, a_t)$.

A two-cycle is often called a **transposition**.

Two cycles are called **disjoint** when their supports are disjoint sets.

When we use cycle notation we will use id for the identity permutation.

Exercises 2.4.5.

(a) Show that disjoint cycles commute.

(b) Suppose σ is a t-cycle. For which r is σ^r a t-cycle? What can happen for other r?

(c) Show that for a t-cycle σ, there is an injective homomorphism $\mathbb{Z}_r \longrightarrow S_n$ taking 1 to σ.

Definition 2.4.6. Let $\pi \in S_n$. The **orbit** of $a \in \{1, \ldots, n\}$ under π is the set $\{\pi^i(a) : i \in \mathbb{Z}\}$.

Let $\pi \in S_n$. A **cycle decomposition** for π is a product of disjoint cycles that is equal to π.

We want to show every permutation has a unique cycle decomposition. The first step is this lemma.

Lemma 2.4.7. *Let $\pi \in S_n$. Any two orbits of π are either equal or disjoint.*

Proof. Suppose two orbits of $\pi \in S_n$ are not disjoint. We will show they are equal. Let a, b, c be distinct elements of $\{1, 2, \ldots, n\}$. Suppose that b is in the orbit of a and also in the orbit of c. We will show that $\text{orb}(a) = \text{orb}(c)$. We have assumed $\pi^i(a) = b$ and $\pi^j(c) = b$ for some $i, j \in \mathbb{Z}$. Then $\pi^i(a) = \pi^j(c)$ so $\pi^{i-j}(a) = c$ so c is in $\text{orb}(a)$. Moreover, anything in the orbit of c must be in the orbit of a since $\pi^k(c) = \pi^{k+i-j}(a)$. The reverse is also true by the same reasoning, so $\text{orb}(a) = \text{orb}(c)$. □

Proposition 2.4.8. *Every permutation in S_n has a cycle decomposition, and it is unique up to reordering the factors.*

Proof. This is just a sketch that should make sense, and one could formalize it using induction. Take an element $a \in \{1, \ldots, n\}$. Since $\{1, \ldots, n\}$ is finite, there is some pair of distinct positive integers such that $\pi^i(a) = \pi^j(a)$. Notice that

$\pi^i(a) = \pi^j(a)$ implies $\pi^{i-1}(a) = \pi^{j-1}(a)$ and so forth until $\pi^{i-j}(a) = a$. Thus, there is some minimal positive integer, call it d, such that $\pi^d(a) = a$. It should be clear that, for $m \in \mathbb{Z}$, $\pi^m(a) = \pi^r(a)$ for r the remainder when m is divided by d. Now consider the cycle $(a, \pi(a), \ldots, \pi^{r-1}(a))$. The orbit of a is this set of elements, $\text{orb}(a) = \{a, \pi(a), \ldots, \pi^{r-1}(a)\}$. Consequently, π can be written as the product of $(a, \pi(a), \ldots, \pi^{r-1}(a))$ and some other permutation that fixes each element in $\text{orb}(a)$. Now choose an element of $\{1, \ldots, n\} \setminus \text{orb}(a)$ and look at its orbit; continue. □

Definition 2.4.9. We will call the list of cycle lengths, in decreasing order, the **signature** of the permutation.

We will include one-cycles in the definition of the cycle decomposition, although we will not write them unless it is needed for clarity. For example, the permutation π in S_5 from Exercise 2.4.1 has cycle decomposition $\pi = (1, 3)(2, 5, 4)$ and signature $3, 2$. If we consider π as an element of S_6, we have $\pi = (1, 3)(2, 5, 4)(6)$ and the signature is $3, 2, 1$.

Exercises 2.4.10. Prove the following results about the signature of a permutation.

(a) For $\pi \in S_n$, the sum of the signature list is n.

(b) If $\pi = \sigma_1\sigma_2 \cdots \sigma_r$ is a cycle decomposition, then $\pi^k = \sigma_1^k\sigma_2^k \cdots \sigma_r^k$. Under what conditions is this also a cycle decomposition in the sense that each σ_i^k is a cycle?

(c) The order of $\pi \in S_n$ is the lcm of the signature list.

Exercises 2.4.11. A sanity check: For $n = 4$ and $n = 5$ do the following to check that all elements of S_n are accounted for.

(a) Identify all possible signatures for elements of S_n and find the order of an element with the given signature.

(b) What is the exponent of S_n?

(c) For each possible signature in S_n, count how many elements have that signature. Then check that you get the correct total number of elements in S_n.

Transpositions and the Alternating Group

There is another factorization that is important.

Proposition 2.4.12. *Every permutation can be written as a product of transpositions.*

Proof. Since every permutation is a product of cycles, it is enough to show that every cycle is a product of transpositions. This is shown by verifying that

$$(a_1, a_2, \ldots, a_t) = (a_1, a_2) * (a_2, a_3) * \cdots * (a_{t-2}, a_{t-1}) * (a_{t-1}, a_t)$$

Recall that we treat permutations are functions and we apply the rightmost permutation first. One can see that a_t gets mapped to a_{t-1} then a_{t-2} and so forth, until the final transposition is applied and takes a_2 (the image of a_t at this point) to a_1. Similar arguments apply to the other a_i. □

We may interpret the previous result as saying that S_n is generated by transpositions. That is somewhat good news: there are $n!$ elements of S_n but only $\binom{n}{2}$ transpositions. Thus $n(n-1)/2$ elements of S_n are enough to generate S_n. In fact we can do much better!

Exercises 2.4.13. Generators for S_n.

(a) Show that S_n is generated by the $n-1$ elements $(1, k)$ for $k = 2, \ldots, n$. [Show that you can get an arbitrary transposition by conjugating $(1, k)$ by some $(1, j)$, see Definition 2.6.9.]

(b) Show that S_n is generated by 2 elements: $(1, 2)$ and $(1, 2, 3, \ldots, n-1, n)$. [Show that you can get all $(1, k)$ from these two using conjugation and then apply the previous exercise.]

Exercises 2.4.14.

(a) Let a, b, c be distinct elements of $\{1, \ldots, n\}$. Write down all possible factorizations of the 3-cycle (a, b, c) as a product of 2 transpositions.

(b) Let a be an element of $\{1, \ldots, n\}$. Let τ_1, τ_2 be transpositions in S_n with $\tau_1 \neq \tau_2$. Show that there exist transpositions $\sigma_1, \sigma_2 \in S_n$ such that $\sigma_1\sigma_2 = \tau_1\tau_2$ and a is not in the support of σ_2. [You will need to consider a few different cases depending on whether a is in the support of τ_1 or τ_2.]

We know from the previous proposition that a permutation can be written as a product of transpositions. This "factorization" is not unique, for example $\mathrm{id} = (1, 2)(1, 2) = (1, 3)(1, 3)$, but the next proposition shows that the parity of the factorization is.

Proposition 2.4.15. *The identity element of S_n cannot be written as the product of an odd number of transpositions.*

Consequently, any permutation can be written as a product of an even number of transpositions, or an odd number of transpositions, but not both.

Proof. We will show that if id is the product of n transpositions then it is the product of $n-2$ transpositions. Consequently, if it is the product of an odd number

of transpositions, inductively we could show that id is a single transposition. This is clearly false.

Suppose that $\mathrm{id} = \tau_1 \cdots \tau_n$ with $\tau_i = (b_{2i-1}, b_{2i})$. The b_i are not necessarily distinct, except $b_{2i-1} \neq b_{2i}$ so that τ_i is indeed a transposition. Let $a = b_1$. Let k be the largest integer such that a is in the support of τ_k (so either b_{2k-1} or b_{2k} is equal to a). Note that $k \neq 1$ because if a was only in the support of τ_1 then $\tau_1 \cdots \tau_n(a) = \tau_1(b_1) = b_2 \neq a$ and the factorization would not be the identity.

Using the previous exercise we can rewrite the factorization of the identity replacing $\tau_{k-1}\tau_k$ with $\sigma_{k-1}\sigma_k$ in which a is not in the support of σ_k (the indexing of $k-1$ and k on σ_{k-1} and σ_k is just for notational convenience). We have a new factorization of id with n terms, but now, only the transpositions $\tau_1, \ldots, \tau_{k-2}$ and σ_{k-1} can have a in the support. If $\tau_{k-2} = \sigma_{k-1}$ we can cancel and get a shorter factorization of the identity using $n-2$ transpositions, as claimed. Otherwise we repeat the process: find the largest index such that the transposition with that index has a in the support; use the exercise to move a into a lower index term; cancel if possible; if not repeat. Eventually we either get a cancelation, or we arrive at a factorization $\tau_1\sigma_2 \cdots \tau_{k+1}\tau_{k+2} \cdots \tau_n$ in which only the first two transpositions τ_1 and σ_2 have a in their support. Then $\tau_1\sigma_2(a) = \mathrm{id}(a) = a$. This is possible only if $\tau_1 = \sigma_2$. Thus we may cancel and get id equal to the product of $n-2$ transpositions as claimed.

For the second part, suppose that π is the product of transpositions in two ways: $\pi = \sigma_1\sigma_2 \ldots \sigma_m = \theta_1\theta_2 \ldots \theta_k$. Then $\mathrm{id} = \sigma_1\sigma_2 \ldots \sigma_m\theta_1^{-1}\theta_2^{-1} \ldots \theta_k^{-1}$. Thus $m+k$ must be even, and this implies that m and k must have the same parity. □

We now have an important and easy consequence.

Proposition 2.4.16. *The set of even parity permutations forms a subgroup of S_n. This is called the* **alternating group** *and is denoted A_n.*

Furthermore, there is a homomorphism from S_n to $\mathbb{Z}_2$ whose kernel is A_n.

Exercises 2.4.17.

(a) Suppose that σ is a k-cycle and τ is an m-cycle and there is exactly one element of $\{1, \ldots, n\}$ that is in the support of both σ and τ. Show that $\sigma\tau$ is a $(k+m-1)$-cycle.

(b) Show that the product of two disjoint transpositions can also be written as the product of two 3-cycles.

(c) Use part (a) (with $k = m = 2$) and part (b) to prove that A_n is generated by 3 cycles.

(d) Compute $(1, 2, a)(1, b, 2)$ for a, b distinct and not equal to 1 or 2. Use the result as motivation to show that the 3-cycles of the form $(1, 2, a)$ generate A_n for $n \geq 4$.

Exercises 2.4.18.

(a) Show that there is a homomorphism from S_n to $\mathbb{Z}_2$ whose kernel is A_n.

(b) Find all subgroups of A_4. Draw a diagram of the subgroup lattice.

(c) What is the intersection of A_4 and D_4?

Cayley's Theorem

For any set T, the set of bijections from T to itself forms a group under composition with the indentity map, $\mathrm{id}(t) = t$, acting as identity element and f mapping to the inverse of a function f as the operation of inversion. In a manner similar to our notation for permutations of $\{1, \ldots, n\}$, we will write the group of bijections of an arbitrary set T as S_T. The elements of S_T will also be called permutations of T.

For a group G, we can forget that G is a group and just look at arbitrary bijections (set maps) from G to itself, that is S_G. The next theorem shows there is an injective homomorphism from G to S_G.

Theorem 2.4.19. *Any group G is isomorphic to a subgroup of S_G, the group of (set) bijections of G to itself. If $|G| = n$ there is an embedding of G in S_n.*

Proof. For each $a \in G$, left multiplication by a maps elements of G to elements of G. Let us call this map $\lambda_a : g \longrightarrow ag$. We can see that λ_a is a permutation of G as follows. For any $g \in G$, $\lambda_a(a^{-1}g) = a(a^{-1}g) = g$, so $a^{-1}g$ is a preimage for g. Since g was arbitrary, λ_a is surjective. We also have λ_a is injective because $\lambda_a(g) = \lambda_a(g')$ implies $ag = ag'$, which by cancellation in G gives $g = g'$.

Define $\lambda : G \longrightarrow S_G$ by $\lambda : a \longrightarrow \lambda_a$. Since $\lambda_a(e) = a$, we have $\lambda_a = \lambda_b$ can only be true if $a = b$. Thus λ is injective. To show it is a homomorphism we have to show that $\lambda_{ab} = \lambda_a \circ \lambda_b$. The following computation does that. We have for all $g \in G$,

$$\lambda_{ab}(g) = (ab)g = a(bg) = \lambda_a(bg) = \lambda_a(\lambda_b(g)) = (\lambda_a \circ \lambda_b)(g)$$

For a finite set T of cardinality n, it should be clear that S_T is isomorphic to S_n. Nevertheless, it is worth explicitly giving a construction of an embedding of G into S_n, for $|G| = n$. Suppose G has order n and enumerate the elements of G so $G = \{g_1, \ldots, g_n\}$. For any $a \in G$, we have shown that λ_a, left multipication by a, permutes the elements of G. Define $\varphi_a \in S_n$ by $\varphi_a(i)$ is the unique j such that $ag_i = g_j$. We may then write $ag_i = g_{\varphi_a(i)}$. Observe that

$$g_{\varphi_{ab}(i)} = abg_i = a(bg_i) = a(g_{\varphi_b(i)}) = g_{\varphi_a(\varphi_b(i))} = g_{\varphi_a \circ \varphi_b(i)}$$

This shows that $\varphi_{ab} = \varphi_a \circ \varphi_b$ so the function φ is a homomorphism.

□

Exercises 2.4.20. Cayley's Theorem

(a) Let $n = 5$ and think of $\mathbb{Z}_n$ in the usual way as $\{0, 1, 2, 3, 4\}$ with addition modulo n. For each $a \in \mathbb{Z}_n$ write down in tabular form the function on $\mathbb{Z}_n$ defined by addition of a.

(b) Show that part (a) defines a function from $\mathbb{Z}_5$ to S_5, provided you think of S_5 as the group of permutations of $\{0, 1, 2, 3, 4\}$. Show that this function is a homomorphism.

(c) Now consider $\mathbb{Z}_2 \times \mathbb{Z}_2$. Enumerate the 4 elements in any way you choose as a_1, a_2, a_3, a_4. For each a_i define a permutation σ_i by $a_i a_1 = a_{\sigma_i(1)}$, $a_i a_2 = a_{\sigma_i(2)}$, $a_i a_3 = a_{\sigma_i(3)}$, $a_i a_4 = a_{\sigma_i(4)}$.

(d) Show in part (c) that this gives a homomorphism from $\mathbb{Z}_2 \times \mathbb{Z}_2$ to S_4.

(e) Similarly, the next steps define a homomorphism from D_3 to S_6. Enumerate the elements of as follows

$$D_3 = \left\{a_1 = r^0, a_2, = r, a_3 = r^2, a_4 = t, a_5 = rt, a_6 = r^2 t\right\}$$

For each a_i define a permutation σ_i in S_6. Since a_1 is the identity in D_3, σ_1 is the identity permutation in S_3. One can see that σ_2 is given by $\sigma_2(i) = k$ whenever $ra_i = a_k$. Verify that each σ_i is indeed a permutation by writing it in permutation notation.

(f) Verify in three examples that for any $a, b \in D_3$, the permutation corresponding to ab equals the product of the permutations corresponding to a and b.

(g) Which elements of D_3 correspond to odd permutations in S_6?

Exercises 2.4.21.

(a) Let A and B be disjoint subsets of $\{1, \ldots, n\}$. Explain how to think of $S_A \times S_B$ as a subgroup of S_n.

(b) Generalize to any partition of $\{1, \ldots, n\}$.

Exercises 2.4.22.

(a) Let n be a positive integer and $k > n/2$. Find a formula for the number of elements of S_n that include a k-cycle.

(b) Use Stirling's formula to approximate the formula you just computed.

(c) Estimate the probability that a random element of S_n has a cycle of length larger than $n/2$.