Notes on Algebra

Donu Arapura

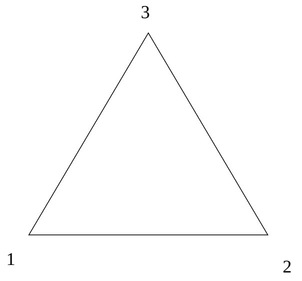
December 5, 2017

Contents

Chapter 1

The idea of a group

One of our goals in this class is to make precise the idea of symmetry, which is important in math, other parts of science, and art. Something like a square has a lot of symmetry, but circle has even more. But what does this mean? One way of expressing this is to a view a *symmetry of a given shape as a motion which takes the shape to itself.* Let us start with the example of an equilateral triangle with vertices labelled by $1, 2, 3$.



We want to describe all the symmetries, which are the motions (both rotations and flips) which takes the triangle to itself. First of all, we can do nothing. We call this I, which stands for identity. In terms of the vertices, I sends $1 \to 1$, $2 \to 2$ and $3 \to 3$. We can rotate once counterclockwise.

$$R_+ : 1 \to 2 \to 3 \to 1.$$

We can rotate once clockwise

$$R_- : 1 \to 3 \to 2 \to 1.$$

We can also flip it in various ways

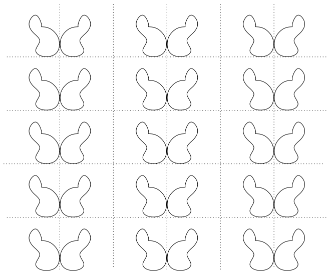
$$F_{12} : 1 \to 2, 2 \to 1, 3 \text{ fixed}$$

$$F_{13} : 1 \to 3, 3 \to 1, 2 \text{ fixed}$$

$$F_{23} : 2 \to 3, 3 \to 2, 1 \text{ fixed}$$

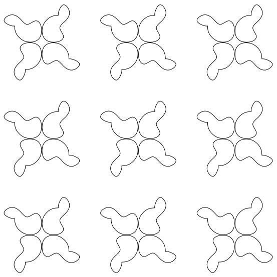
We will say more about this example and generalizations for regular polygons later. In the limit, as the number of vertices go to infinity, we get the circle. This has infinitely many symmetries. We can use any rotation about the center, or a reflection about a line through the center.

Another example which occurs in classical art and design (mosaics, wallpaper....) and two dimensional crystals is a repetetive pattern in the plane such as the one drawn below.



We imagine this covering the entire plane; the grid lines are not part of the pattern. Then there are infinitely many symmetries. We can translate or shift all the "ducks" up or down by one square, or left or right by two squares. We can also flip or reflect the pattern along vertical lines.

Here is another pattern below.



This has translational symmetries as before, but no flipping symmetries. Instead, if the plane is rotated by 90° about any point where four ducks meet, the pattern is preserved. One might ask can we replace four by five, or some arbitrary number of, ducks and still get an infinitely repeating symmetric pattern as above? The answer surprisingly is no. We will prove this later.

The study of symmetry leads to an algebraic structure. To simplify things, let us ignore flips and consider only rotational symmetries of a circle C of radius r. To simplify further, let us start with the limiting case where $r \to \infty$. Then C becomes a line L, and rotations correspond to translations. These can be described precisely as follows. Given a real number $x \in \mathbb{R}$, let $T_x : L \to L$

denote the symmetry which takes a point p on L and moves it by a distance x to the right if $x > 0$, fixes it if $x = 0$, or moves it to the left if $x < 0$. We can see that if we translate by x and then by y, it is the same as translating by $x + y$. So addition emerges naturally with this context. Basic laws of algebra have a natural meaning here: Translating by x and then by 0 is the same as translating by x, or in symbols

$$x + 0 = x \quad (\text{ 0 is the identity})$$

Translating by x and then y is that same translating y and then x, or

$$x + y = y + x \quad (\text{commutative law})$$

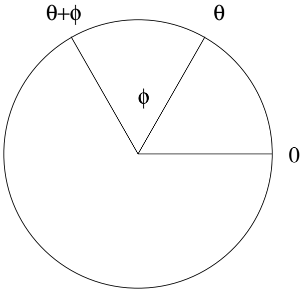
We can always translate back to where we started because

$$\text{Given } x, \text{ we can find } y \text{ with } x + y = 0 \quad (\text{existence of the inverse})$$

Finally, translating by three numbers x, y, z in succesion is the same as translating by $x + y$ and then z, or x then $y + z$. That is

$$(x + y) + z = x + (y + z) \quad (\text{associative law})$$

Now we are ready to consider the rotational symmetries of the circle C of finite radius. Let $R_\theta : C \to C$ be the rotation (counterclockwise) through an angle $\theta \in [0, 2\pi) = \{x \in \mathbb{R} \mid 0 \le x < 2\pi\}$ measured in radians. Note that we can identify C with the set of angles $[0, 2\pi)$ as well. Now we define addition in C as follows: given $\theta, \phi \in C$, let $\theta \oplus \phi$ be given by rotating θ by the additional angle ϕ.



Here are a few simple examples

$$\pi/2 \oplus \pi/2 = \pi$$

$$\pi \oplus \pi = 0$$

In general, we can see that

$$\theta \oplus \phi = \begin{cases} \theta + \phi & \text{if } \theta + \phi < 2\pi \\ \theta + \phi - 2\pi & \text{if } \theta + \phi \ge 2\pi \end{cases}$$

And it will be convenient to adopt this last equation as the official definition.

At first it may seem like a strange operation, but notice that many familiar rules apply:

Lemma 1.1. *If $\theta \in C$, then $\theta \oplus 0 = \theta$.*

Proof. Since $\theta < 2\pi$, $\theta \oplus 0 = \theta + 0 = \theta$ □

Lemma 1.2. *If $\theta, \phi \in C$, then $\theta \oplus \phi = \phi \oplus \theta$.*

Proof. If we compare

$$\phi \oplus \theta = \begin{cases} \phi + \theta & \text{if } \phi + \theta < 2\pi \\ \phi + \theta - 2\pi & \text{if } \phi + \theta \geq 2\pi \end{cases}$$

we see that it is identical to $\theta \oplus \phi$. □

Lemma 1.3. *Given $\theta \in C$, we have $\phi \in C$ such that $\theta \oplus \phi = 0$.*

Proof. We can take $\phi = \ominus\theta = 2\pi - \theta$. □

We omit the proof for now, but the associative law

$$\theta \oplus (\phi \oplus \psi) = (\theta \oplus \phi) \oplus \psi$$

also holds.

So in summary, the set C with the operation $\oplus$ shares the same 4 laws as $\mathbb{R}$ with usual addition: namely the associative and commutative laws, and the existence of identity and inverse. We have a name for such a thing. It is called an *abelian group*, and it will be one of the key concepts in this class. To appreciate the power of this simple set of rules, let us extend a standard result from highschool algebra.

Theorem 1.4. *Suppose that A is any abelian group with operation $+$ and identity 0. For any $a, b \in A$, there is exactly one solution to $x + a = b$.*

Proof. By the axioms, there exists an element that we denote by $-a$ such that $a + (-a) = 0$. Add b to both sides, and use the laws to obtain

$$(b + (-a)) + a = b + (-a + a) = b + 0 = b$$

Therefore $x = b + (-a)$ gives a solution. Suppose that x is any solution to $x + a = b$. Then adding $-a$ to both sides and use the associative law

$$x = x + (a + (-a)) = (x + a) + (-a) = b + (-a)$$

□

We are being a bit pedantic in our notation, since this was the first abstract proof. In the future, we will just write $b - a$ instead of $b + (-a)$.

We want to return to the first example of the triangle, but first we should clarify what kind of mathematical objects we are dealing with. Given a set X, a permutation of X is a *one to one onto function* $f : X \to X$. Recall that function, or map, mapping or transformation $f : X \to Y$ is a rule for am taking element x of one set X to an element $f(x) \in Y$; it is one to one and onto if every element of Y equals $f(x)$ for exactly one $x \in X$. The symmetries R_+ etc. are just permutations of $\{1, 2, 3\}$. Here are some abstractly given permutations of the set $\{1, 2, 3, 4\}$.

$$f(1) = 2,\ f(2) = 3,\ f(3) = 1,\ f(4) = 4$$

$$g(1) = 1,\ g(2) = 1,\ g(3) = 4,\ g(4) = 3$$

The function h defined by

$$h(1) = h(2) = 1,\ h(3) = h(4) = 2$$

is not a permutation. It may be helpful to visualize these

$$f = \left\{ \begin{array}{ccc} 1 & \to & 2 \\ 2 & \to & 3 \\ 3 & \to & 1 \\ 4 & \to & 4 \end{array} \right. = \left\{ \begin{array}{ccc} 2 & \leftarrow & 1 \\ 3 & \leftarrow & 2 \\ 1 & \leftarrow & 3 \\ 4 & \leftarrow & 4 \end{array} \right. ,$$

$$g = \left\{ \begin{array}{ccc} 1 & \to & 2 \\ 2 & \to & 1 \\ 3 & \to & 4 \\ 4 & \to & 3 \end{array} \right. = \left\{ \begin{array}{ccc} 2 & \leftarrow & 1 \\ 1 & \leftarrow & 2 \\ 4 & \leftarrow & 3 \\ 3 & \leftarrow & 4 \end{array} \right.$$

Since the above notations are a bit cumbersome, we often write this in *permutation notation* as

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}, g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

Note these are **not** matrices. There is yet another notation, which a bit more compact. A cycle of a permutation is a sequence of elements $a \to f(a) \to f(f(a)) \ldots \to a$ For f, the cycles are $1 \to 2 \to 3 \to 1$ and $4 \to 4$; for g, $1 \to 2 \to 1$ and $3 \to 4 \to 3$. To specify a permutation it is just enough to list the cycles as in

$$f = (123)(4),\ g = (12)(34)$$

Cycles consisting of just one element are usually omitted, so we would write $f = (123)$. Note that (312) would also represent f.

Given two permutations $f : X \to X$ and $g : X \to X$. We can *multiply* them by *composing* them as functions. In the examples above,

$$f \circ g(1) = f(g(1)) = f(2) = 3, \text{ etc.}$$

We usually omit the $\circ$ symbol. More visually

$$fg = \left\{ \begin{array}{ccccc} 3 & \leftarrow & 2 & \leftarrow & 1 \\ 2 & \leftarrow & 1 & \leftarrow & 2 \\ 4 & \leftarrow & 4 & \leftarrow & 3 \\ 1 & \leftarrow & 3 & \leftarrow & 4 \end{array} \right. = \left\{ \begin{array}{ccc} 3 & \leftarrow & 1 \\ 2 & \leftarrow & 2 \\ 4 & \leftarrow & 3 \\ 1 & \leftarrow & 4 \end{array} \right.$$

Note that we use backword arrows because this is consistent with function composition. Some people (and software) use forward arrows, which is easier to work with, but confusing in other ways.

With a bit of practice, this can this read off directly from the permutation symbols

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

We now return to our triangle example.

$$R_+R_+ = \begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix}\begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix} = \begin{pmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{pmatrix} = R_-$$

Let's do two flips, F_{12} followed by F_{13}

$$F_{12}F_{13} = \begin{pmatrix}1 & 2 & 3\\ 2 & 1 & 3\end{pmatrix}\begin{pmatrix}1 & 2 & 3\\ 3 & 2 & 1\end{pmatrix} = \begin{pmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{pmatrix} = R_-$$

Doing this the other way gives

$$F_{13}F_{12} = R_+$$

Therefore this multiplication is not commutative.

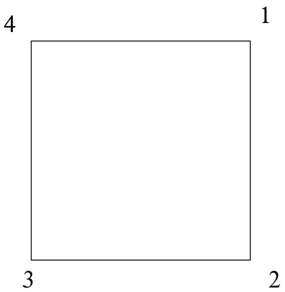
The full multiplication table can be worked out with enough patience as

$\circ$	I	F_{12}	F_{13}	F_{23}	R_+	R_-
I	I	F_{12}	F_{13}	F_{23}	R_+	R_-
F_{12}	F_{12}	I	R_-	R_+	F_{23}	F_{13}
F_{13}	F_{13}	R_+	I	R_-	F_{13}	F_{23}
F_{23}	F_{23}	R_-	R_+	I	F_{12}	F_{13}
R_+	R_+	F_{23}	F_{12}	F_{13}	R_-	I
R_-	R_-	F_{13}	F_{23}	F_{12}	I	R_+

One thing that can be observed from the table is that every element has an inverse, i.e. an element which multiplies with it to give the identity. It is not obvious from the table that the associative law holds, but this is something we will prove later. A *group* is a set with an multiplication, which is associative, has an identity and such that every element has an inverse. We will be clarify the meaning of the axioms later. Suffice it to say that we now have two new examples of groups. One which is abelian and one which isn't.

1.5 Exercises

In the next few exercises, you will study the symmetries of a square with vertices labelled by $1, 2, 3, 4$ as shown



Let

$$I = i = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

R be the clockwise rotation

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

and F be the flip

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

1. Show that all the rotations preserving the square are given by $I, R, R^2 = RR$ and R^3. Write these out explicitly in cycle notation.

2. Show that all the flips (including diagonal flips) preserving the square are given by F, FR, FR^2, FR^3. Write these out explicitly in cycle notation.

3. The above 8 rotations and flips is a complete list of all the symmetries of the square. Describe RF in terms of this list. Give an example of a permutation of $\{1, 2, 3, 4\}$ which is not a symmetry of the square.

4. Determine the inverses of the rotations $R, R^2 = RR$ and R^3.

5. Determine the group of symmetries (rotations and flips) of a rectangle which is not a square. Is this abelian?

6. Determine all the symmetries of a regular pentagon. Regular means that all the sides have the same length.

7. (If you forgot what complex numbers are, now is the time to remind yourself.)

   (a) Given $z = a + bi \in \mathbb{C}$, recall that $\bar{z} = a - bi$. Check that $z\bar{z} = a^2 + b^2$, and also that $\bar{z}\bar{w} = \overline{zw}$ for $w = c + di$.

   (b) Let C be the set of complex numbers of the form $a + bi$, where $a^2 + b^2 = 1$. With the help of the previous exercise, prove that if $z \in C$, then $z^{-1} \in C$, and that the product of any two numbers in C is also in C. Conclude that C is a group under multiplication.

(c) Given an angle θ, show that $e^{i\theta} = \cos\theta + i\sin\theta \in C$ and conversely, every element of $z \in C$ is of this form for a unique $\theta \in [0, 2\pi)$. This is another way to turn C into a group which is the *same* as the previous group in an appropriate sense.

Chapter 2

The group of permutations

Recall that a function $f : X \to Y$ is *one to one* if for any pair of distinct elements $x_1, x_2 \in X$, $f(x_1) \neq f(x_2)$. Equivalently, if $f(x_1) = f(x_2)$ then $x_1 = x_2$. f is *onto* if for every $y \in Y$, we can find an $x \in X$ such that $f(x) = y$. An important example of a function is the identity function $id_X : X \to X$ defined by $id_X(x) = x$. This is clearly one to one and onto. If X is understood, we write this as id.

Lemma 2.1. *Suppose that $f : X \to Y$ and $g : Y \to Z$ are functions.*

1. *If f and g are one to one, then so is $g \circ f$.*

2. *If f and g are onto, then so is $g \circ f$.*

Proof. Suppose that f and g are one to one. If $g \circ f(x_1) = g \circ f(x_2)$, then $g(f(x_1)) = g(f(x_2))$. This implies $f(x_1) = f(x_2)$ because g is one to one. Therefore $x_1 = x_2$ because f is one to one. This proves 1.

Suppose that f and g are onto. Given $z \in Z$, we can find $y \in Y$ such that $g(y) = z$ because g is onto. We can also find $x \in X$ such that $f(x) = y$. Therefore $g \circ f(x) = z$. This proves 2. □

Lemma 2.2. *Suppose that $f : X \to Y$, $g : Y \to Z$ and $h : Z \to W$ are functions, then $h \circ (g \circ f) = (h \circ g) \circ f$*

Proof. To be clear two functions are considered to be equal if they produce equal outputs on the same input. Now observe that

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

□

Lemma 2.3. *If $f : X \to Y$ is one to one and onto, there exists a function $f^{-1} : Y \to X$ called the inverse such that $f \circ f^{-1} = id_Y$ and $f^{-1} \circ f = id_X$.*

Proof. For every $y \in Y$, there exists a unique $x \in X$ such that $f(x) = y$. We define $f^{-1}(y) = x$. Then $f^{-1} \circ f(x) = f^{-1}(y) = x$ and $f \circ f^{-1}(y) = f(x) = y$. □

Lemma 2.4. *Given a function $f : X \to Y$, $f \circ id_X = f$ and $id_Y \circ f = f$.*

Proof. The first equation holds because $f \circ id(x) = f(id(x)) = f(x)$. The proof of the second is similar. □

Now come to the key definition.

Definition 2.5. *A group is a set G with an operation $*$ and a special element e satisfying*

1. *The associative law: $(x * y) * z = x * (y * z)$*
2. *e is the identity: $x * e = e * x = x$*
3. *Existence of inverses: given x, there exists y such that $x * y = y * x = e$*

We sometimes say that $(G, *, e)$ is a group we want to specify the operation and identity. Occasionally, we will omit the operation, and simply write xy for $x*y$. We will see in the exercises that each x has exactly one inverse. We denote this by x^{-1}, or sometimes $-x$, depending on the situation.

It is also worth repeating what we said in the first chapter in this context.

Definition 2.6. *An abelian group is a group G for which the commutative law $x * y = y * x$ holds.*

Given a set X, recall that a permutation of X is a one to one onto function $f : X \to X$. Let S_X denote the set of permutations of X. When $X = \{1, 2, \dots, n\}$, which is the case we will mostly be interested in, we denote this by S_n. Putting the previous lemmas, we get

Theorem 2.7. *S_X becomes a group under composition, with identity given by id.*

S_n is called the *symmetric group on n letters*. Most of you have actually encountered this before, although perhaps not by name, and in particular, you probably already know is that:

Theorem 2.8. *The number of elements of S_n is $n! = 1 \cdot 2 \cdot 3 \cdots n$.*

We will in fact give a proof of this later on. For $n = 3$, we see that S_3 has 6 elements, so it must coincide with the symmetry group of the triangle. For $n = 4$, we have 24 which is much bigger than the symmetries of the square. This a pretty typical. We are often interested not in the whole of S_n, but some interesting piece of it.

Definition 2.9. *Given a group $(G, *, e)$, a subset $S \subseteq G$ is called a subgroup if $e \in S$, and $x, y \in S$ implies $x * y, x^{-1} \in S$ (one says that S is closed under these operations).*

The definition ensures that if these operations can be restricted to S, we don't leave S.

Proposition 2.10. *A subgroup $S \subseteq G$ of a group is also a group.*

There is actually nothing to prove. The same laws of G hold for elements of S.

Coming back to permutation notation, we note see that the identity is simply

$$id = \begin{pmatrix} 1 & 2 & 3 & \dots \\ 1 & 2 & 3 & \dots \end{pmatrix}$$

To find the inverse, we simply turn it upside down and then rearrange columns. For example,

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

$$f^{-1} = \begin{pmatrix} 1 & 4 & 2 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$$

In cycle notation, we simply reverse the cycles

$$f = (243),\ f^{-1} = (342)$$

2.11 Exercises

1. Let X be a nonempty set and let $f : X \to X$ be a function. Prove that f is one to one if and only if there is a function $g : X \to X$ such that $gf = id$; g is called a left inverse. (One direction is easy, and the other will require you to be a bit creative.)

2. Let X be a nonempty set and let $f : X \to X$ be a function. Prove that f is onto if and only if there is a function $g : X \to X$ such that $fg = id$; g is called a right inverse. (People who know some set theory will need to invoke the axiom of choice.)

3. A permutation $f \in S_n$ is a transposition, if it interchanges two numbers, say i and j and fixes everything else, i.e. $f(i) = j, f(j) = i,\ f(x) = x, i \neq x \neq j$, or $f = (ij)$ in cycle notation.

   (a) Check that everything in S_3 is a product of transpositions.

   (b) Check $(12)(34), (123), (1234) \in S_4$ are products of transpositions. Generalizing from these examples, prove that every element of S_4 is a product of transpositions.

4. Given a group $(G, *, e)$, prove that it is has only one identity element. In other words, if $x * e' = e' * x = x$ holds for all x, prove $e' = e$.

5. Given a group $(G, *, e)$,

(a) Prove that every element x as exactly one inverse. We now denote it by x^{-1}.

(b) Prove that $(x * y)^{-1} = y^{-1} * x^{-1}$.

6. Given a group $(G, *, e)$,

   (a) Given $y, z \in G$, prove that there is exactly one $x_1 \in G$ satisfying $x_1 * y = z$ and exactly one $x_2 \in G$ satisfying $y * x_2 = z$.

   (b) Is *always* true that $x_1 = x_2$? If yes, then prove it; if no, then find a *counterexample*, i.e. a group G and elements x_1, x_2, y, z as above with $x_1 \neq x_2$.

7. Let $R = (123)$ and F a transposition in S_3

   (a) Check that $\{I, R, R^2\}$, and $\{I, F\}$, are subgroups of S_3

   (b) Prove that S_3 does not have a subgroup with exactly 4 elements. (If you happen to know Lagrange's theorem, don't use it. Give a direct argument.)

8. Recall that the intersection (respectively union) of two sets $H \cap K$ $(H \cup K)$ is the set elements x such that $x \in H$ *and* $x \in K$ (respectively $x \in H$ *or* $x \in K$ – x is allowed to be in both).

   (a) Prove that if H and K are both subgroups of a group G, then $H \cap K$ is a subgroup.

   (b) What about $H \cup K$?

Chapter 3

Rotations and reflections in the plane

We want another important source of nonabelian groups, which is one that most people should already be familiar. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots \\ a_{21} & a_{22} & \dots \\ \dots \end{bmatrix}$$

be an $n \times n$ matrix with entries in $\mathbb{R}$. If B is another $n \times n$ matrix, we can form their product $C = AB$ which is another $n \times n$ matrix with entries

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots a_{in}b_{nj} = \sum_k a_{ik}b_{kj}$$

The identity matrix

$$I = \begin{bmatrix} 1 & 0 & \dots \\ 0 & 1 & \dots \\ \dots \end{bmatrix}$$

has entries

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Lemma 3.1. *Matrix multiplication is associative and I is the identity for it, i.e. $AI = IA = A$.*

Proof. Given matrices A, B, C, the ijth entries of $A(BC)$ and $(AB)C$ both work out to

$$\sum_k \sum_\ell a_{ik}b_{k\ell}c_{\ell j}$$

Also

$$a_{ij} = \sum_k a_{ik}\delta_{kj} = \sum_k \delta_{ik}a_{kj}$$

□

An $n \times n$ matrix A is *invertible* if there exists an $n \times n$ matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$. It follows that:

Theorem 3.2. *The set of invertible $n \times n$ matrices with entries in $\mathbb{R}$ forms a group called the* general linear group *$GL_n(\mathbb{R})$.*

For 2×2 matrices there is a simple test for invertibility. We recall that the determinant

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

and

$$e \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ea & eb \\ ec & ed \end{bmatrix}$$

Theorem 3.3. *Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be matrix over $\mathbb{R}$, then A is invertible if and only* $\det(A) \neq 0$. *In this case,*

$$A^{-1} = (\det(A))^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Proof. Let $\Delta = \det(A)$, and let $B = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Then an easy calculation gives

$$AB = BA = \Delta I.$$

If $\Delta \neq 0$, then $\Delta^{-1}B$ will give the inverse of A by the above equation.

Suppose that $\Delta = 0$ and A^{-1} exists. Then multiply both sides of the above equation by A^{-1} to get $B = \Delta A^{-1} = 0$. This implies that $A = 0$, and therefore that $0 = AA^{-1} = I$. This is impossible. □

Let us study an important subgroup of this. A 2×2 rotation matrix is a matrix of the form

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

This sends a column vector v in the plane $\mathbb{R}^2$ to the vector $R(\theta)v$ obtained by rotation through angle θ. We denote the set of these by $SO(2)$ (SO stands for special orthogonal).

Theorem 3.4. *$SO(2)$ forms a subgroup of $GL_2(\mathbb{R})$.*

Proof. It is easy to check that $\det R(\theta) = \cos^2\theta + \sin^2\theta = 1$ and of course, $R(0) = I \in SO(2)$. If we multiply two rotation matrices

$$\begin{aligned} R(\theta)R(\phi) &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & -\cos\theta\sin\phi - \sin\theta\cos\phi \\ \\ \sin\theta\cos\phi + \cos\theta\sin\phi & \cos\theta\cos\phi - \sin\theta\sin\phi \end{bmatrix} \\ &= \begin{bmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix} \\ &= R(\theta+\phi) \end{aligned}$$

Therefore $SO(2)$ is closed under multiplication. The last calculation also shows that $R(\theta)^{-1} = R(-\theta) \in SO(2)$ □

A matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is called *orthogonal* if the columns are unit vectors $a^2 + c^2 = b^2 + d^2 = 1$ which are orthogonal in the sense that the dot product $ab + cd = 0$. Since the first column is on the unit circle, it can be written as $(\cos\theta, \sin\theta)^T$ (the symbol $(-)^T$, read *transpose*, turns a row into a column). The second column is on the intersection of the line perpendicular to the first column and the unit circle. This implies that the second column is $\pm(-\sin\theta, \cos\theta)^T$. So either $A = R(\theta)$ or

$$A = F(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

In the exercises, you will find a pair of nonzero orthogonal vectors v_1, v_2, $F(\theta)v_1 = v_1$ and $F(\theta)v_2 = -v_2$. This means that $F(\theta)$ is a *reflection* about the line spanned by v_1. In the exercises, you will also prove that

Theorem 3.5. *The set of orthogonal matrices $O(2)$ forms a subgroup of $GL_2(\mathbb{R})$.*

Given a unit vector $v \in \mathbb{R}^2$ and $A \in O(2)$, Av is also a unit vector. So we can interpret $O(2)$ as the full symmetry group of the circle, including both rotations and reflections.

3.6 Exercises

1. Let $UT(2)$ be the set of upper triangular matrices

$$\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$$

   Show this forms a subgroup of $GL_2(\mathbb{R})$.

2. Let $UT(3)$ be the set of upper triangular matrices

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$$

   Show this forms a subgroup of $GL_3(\mathbb{R})$.

3. Find a pair of nonzero orthogonal vectors v_1, v_2, $F(\theta)v_1 = v_1$ and $F(\theta)v_2 = -v_2$. (Hint: if $\theta = 0$ this is easy; when $\theta \neq 0$, try $v_1 = (\sin\theta, 1 - \cos\theta)^T$.)

4. Recall that the transpose of an $n \times n$ matrix A is the $n \times n$ matrix with entries a_{ji}. A matrix is called orthogonal if $A^T A = I = AA^T$ (the second equation is redundant but included for convenience).

   (a) Check that this definition of orthogonality agrees with the one we gave for 2×2 matrices.

   (b) Prove that the set of $n \times n$ orthogonal matrices $O(n)$ is a subgroup of $GL_n(\mathbb{R})$. You'll need to know that $(AB)^T = B^T A^T$.

5. Show that $SO(2)$ is abelian, but that $O(2)$ is not.

6. A 3×3 matrix is called a permutation matrix, if it can be obtained from the identity I by permuting the columns. Write $P(\sigma)$ for the permutation matrix corresponding to $\sigma \in S_3$. For example,

$$F = P((12)) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

   Check that $F^2 = I$. What can you conclude about the set $\{I, F\}$?

7. Prove that the set of permutations matrices in $GL_3(\mathbb{R})$ forms a subgroup. Prove the same thing for $GL_n(\mathbb{R})$, where permutations matrices are defined the same way. (The second part is not really harder than the first, depending how you approach it.)

Chapter 4

Cyclic groups and dihedral groups

Consider the group C_n of rotational symmetries of a regular n-gon. If we label the vertices consecutively by $1, 2 \ldots, n$. Then we can view

$$C_n = \{I, R, R^2, \ldots R^{n-1}\} \subset S_n$$

where $I = id$ and

$$R = (123 \ldots n)$$

A bit of thought shows that $R^n = I$. We won't need to multiply permutations explicitly, we just use this rule: $R^j R^k = R^{j+k}$ and if $j + k \geq n$, we "wrap around" to R^{j+k-n}. We will encounter other groups with a similar structure.

Definition 4.1. *A finite group G is called cyclic if there exists an element $g \in G$, called a generator, such that every element of G is a power of g.*

Cyclic groups are really the simplest kinds of groups. In particular:

Lemma 4.2. *A cyclic group is abelian.*

Proof. $g^j g^k = g^{j+k} = g^k g^j$. □

Let us give a second example. Let

$$\mathbb{Z}_n = \{0, 1, 2 \ldots n-1\}$$

We modify addition using the same wrap around rule as before.

$$x \oplus y = \begin{cases} x + y & \text{if } x + y \in \mathbb{Z}_n \\ x + y - n & \text{otherwise} \end{cases}$$

This is usually called modular addition. It is not completely obvious that this is a group but we will show this later. Here is the table for $n = 2$

$\oplus$	0	1
0	0	1
1	1	0

This is the simplest nonzero abelian group. A somewhat more complicated case is $n = 4$

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

$\mathbb{Z}_n$ with this addition rule is also cyclic with generator 1.

We can see that $\mu_2 = \{1, -1\}$ is a cyclic group under multiplication. More generally, the group of nth roots of unity.

$$\mu_n = \left\{ e^{2\pi i k/n} = \cos\left(\frac{2\pi i k}{n}\right) + i \sin\left(\frac{2\pi i k}{n}\right) \mid k = 0, 1, \ldots n-1 \right\}$$

This is a subgroup of the group of nonzero complex numbers $\mathbb{C}^*$ under multiplication. μ_n is generated by $e^{2\pi i/n}$, so it is cyclic.

Although these examples are superficially different, they are the same in some sense. If we associate $k \mapsto R^k$ or $k \mapsto e^{2\pi i k/n}$ and compare addition/multiplication tables, they will match. Here is the precise definition.

Definition 4.3. *If $(G, *, e)$ and $(H, \circ, e')$ are groups. A function $f : G \to H$ is called a homomorphism if $f(e) = e'$ and $f(g_1 * g_2) = f(g_1) \circ f(g_2)$. A one to one onto homomorphism is called an isomorphism. Two groups are isomorphic if there is a homomorphism from one to the other. In symbols, we write $G \cong H$.*

The function $f : \mathbb{Z}_n \to C_n$ defined by $f(k) = R^k$ is an isomorphism. The function $f : \mathbb{Z} \to \mu_n$ defined by $f(k) = e^{2\pi i k/n}$ is a homomorphism which is not an isomorphism because it is not one to one. The *order* of a finite group is the number of elements in it.

Theorem 4.4. *A cyclic group of order n is isomorphic to $\mathbb{Z}_n$.*

Proof. Let G be the cyclic group in question with generator g. Since G is finite, the sequence g^n must repeat itself. That is $g^{n_1} = g^{n_2}$ for $n_1 > n_2$. Taking $n = n_1 - n_2 > 0$ implies that $g^n = e$. Let us assume that n is the smallest such number (this is called the order of g). We claim that $G = \{e, g, \ldots, g^{n-1}\}$ and that all the elements as written are distinct. By distinctness we mean that if $m_1 > m_2$ lie in $\{0, 1, \ldots n-1\}$ then $g^{m_1} \neq g^{m_2}$. If not then $g^{m_1 - m_2} = e$ would contradict the fact that n is the order of g.

So now the function $f(i) = g^i$ is easily seen to given an isomorphism from $\mathbb{Z}_n$ to G. □

We need to come back and check that $\mathbb{Z}_n$ is actually a group. We make use of a result usually called the "division algorithm". Although it's not an algorithm in the technical sense, it is the basis of the algorithm for long division that one learns in school.

Theorem 4.5. *Let x be an integer and n positive integer, then there exists a unique pair of integers q, r satisfying*

$$x = qn + r,\ 0 \leq r < n$$

Proof. Let

$$R = \{x - q'n \mid, q' \in \mathbb{Z} \text{ and } q'n \leq x\}$$

Observe that $R \subseteq \mathbb{N}$, so we can choose a smallest element $r = x - qn \in R$. Suppose $r \geq n$. Then $x = qn + r = (q+1)n + (r - n)$ means that $r - n$ lies in R. This is a contradiction, therefore $r < n$.

Suppose that $x = q'n + r'$ with $r' < n$. Then $r' \in R$ so $r' \geq r$. Then $qn = q'n + (r' - r)$ implies that $n(q - q') = r' - r$. So $r' - r$ is divisible by n. On the other hand $0 \leq r' - r < n$. But 0 is the only integer in this range divisible by n is 0. Therefore $r = r'$ and $qn = q'n$ which implies $q = q'$.

□

We denote the number r given above by $x \, mod \, n$; mod is read "modulo" or simply "mod". When $x \geq 0$, this is just the remainder after long divison by n.

Lemma 4.6. *If x_1, x_2, n are integers with $n > 0$, then*

$$(x_1 + x_2) \, mod \, n = (x_1 \, mod \, n) \oplus (x_2 \, mod \, n)$$

Proof. Set $r_i = x_i \, mod \, n$. Then $x_i = q_i n + r_i$ for appropriate q_i. We have $x_1 + x_2 = (q_1 + q_2)n + (r_1 + r_2)$. We see that

$$(x_1 + x_2) \, mod \, n = \begin{cases} r_1 + r_2 = r_1 \oplus r_2 & \text{if } r_1 + r_2 < n \\ r_1 + r_2 - n = r_1 \oplus r_2 & \text{otherwise} \end{cases}$$

□

This would imply that $f(x) = x \, mod \, n$ gives a homomorphism from $\mathbb{Z} \to \mathbb{Z}_n$ if we already knew that $\mathbb{Z}_n$ were a group. Fortunately, this can be converted into a proof that it is one.

Lemma 4.7. *Suppose that $(G, *, e)$ is a group and $f : G \to H$ is an onto map to another set H with an operation $*$ such that $f(x * y) = f(x) * f(y)$. Then H is a group with identity $f(e)$.*

In the future, we usually just write + for modular addition.

The dihedral group D_n is the full symmetry group of regular n-gon which includes both rotations and flips. There are $2n$ elements in total consisting of n rotations and n flips. Label the vertices consecutively by $1, 2, 3 \ldots$. Let $R = (123 \ldots n)$ be the basic rotation. This generates a cyclic subgroup $C_n \subset D_n$. The reflection around the line through the midpoint of $\overline{1n}$ and opposite side or vertex is

$$F = (1\,n)(2\,n-1)(3\,n-2)\ldots$$

One can calculate that

$$\begin{aligned} FR &= \begin{pmatrix} 1 & 2 & \ldots & n \\ n & n-1 & \ldots & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & \ldots & n \\ 2 & 3 & \ldots & 1 \end{pmatrix} \\ &= (1\,n-1)(2\,n-2)\ldots \end{aligned}$$

is another flip, and furthermore that

$$\begin{aligned} FRF &= \begin{pmatrix} 1 & 2 & \ldots & n \\ n-1 & n-2 & \ldots & n \end{pmatrix} \begin{pmatrix} 1 & 2 & \ldots & n \\ n & n-1 & \ldots & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & \ldots & n \\ n & 1 & \ldots & n-1 \end{pmatrix} \\ &= R^{-1} \end{aligned}$$

Here's the point. We will eventually see that the elements of D_n are given by $I, R, R^2 \ldots, F, FR, FR^2$. So we say that these elements generate the group. (In general, to say that a set elements generates a group, means that we have to take products in every possible way such as FR^2F^3.) We have three basic relations among the generators

$$F^2 = I,\ R^n = I,\ FRF = R^{-1}$$

Everything else about D_n follows from this. In particular, we won't have to multiply any more permutations. For instance, let us check that $(FR)^2 = I$ using only these relations

$$(FR)^2 = (FRF)R = R^{-1}R = I$$

4.8 Exercises

1. Determine all the generators of $\mathbb{Z}_6$ and $\mathbb{Z}_8$. Is there an obvious pattern?

2. Let $\mathbb{Z}_7^* = \{1, 2, 3, 4, 5, 6\}$ with an operation defined by $x \odot y = (x \cdot y)\ mod\ 7$. Assume that it is associative, and check that $\mathbb{Z}_7^*$ is a cyclic group.

3. Given a finite group G and $g \in G$, prove that $\{e, g, g^2, \ldots\}$ is a cyclic subgroup. This called the subgroup generated by G. The order of this group is called the *order of* g. Prove that the order is the smallest positive integer n such that $g^n = e$.

4. Given a function $f : H \to G$ such that $f(x * y) = f(x) * f(y)$, prove that f takes the identity to the identity and is therefore a homomorphism.

5. Complete the proof of lemma 4.7.

6. Let us say that an infinite group is cyclic if it isomorphic to $\mathbb{Z}$. Prove that the set of even integers is cyclic.

7. Let $G \subseteq \mathbb{Z}$ be nonzero subgroup, where $\mathbb{Z}$ is a group under addition. Let $d \in G$ be the smallest positive element. Prove that if $x \in G$, then $x = qd$ for some integer q. Conclude that G is cyclic.

8. Let $F, R \in D_n$ be as above.

   (a) For any $i > 0$, show that $FR^iF = R^{-i}$, where R^{-i} is the inverse of R^i.

   (b) Show that for any $i, j > 0$, $(FR^i)(FR^j)$ is a rotation.

   (c) Show every element of D_n is either R^i or FR^i with $i = 0, 1, \dots, n$.

9. Assuming the previous exercise, show that $f : D_n \to \mathbb{Z}_2$ given by $f(R^i) = 0$ and $f(FR^i) = 1$ is a homomorphism.

10. Let $G \subset O(2)$ be the set of matrices

$$\left\{ \begin{bmatrix} \cos\theta & \pm\sin\theta \\ \sin\theta & \mp\cos\theta \end{bmatrix} \mid \theta = \frac{2\pi k}{n}, k = 0, 1, \dots n-1 \right\}$$

Let

$$R = R\left(\frac{2\pi}{n}\right), F = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Check that G is generated by these two elements, and that they satisfy the same relations as the generators of the D_n. Use these facts to prove that D_n is isomorphic to G.

Chapter 5

Finite sets, counting and group theory

Let $\mathbb{N} = \{0, 1, 2 \ldots\}$ be the set of natural numbers. Given n, let $[n] = \{x \in \mathbb{N} \mid x < n\}$. So that $[0] = \emptyset$ is the empty set, and $[n] = \{0, 1, \ldots, n-1\}$ if $n > 0$. A set X is called *finite* if there is a one to one onto function (also called a one to one correspondence) $f : [n] \to X$ for some $n \in \mathbb{N}$. The choice of n is unique (which we will accept as a fact), and is called the cardinality of X, which we denote by $|X|$.

Lemma 5.1. *If X is finite and $g : X \to Y$ is a one to one correspondence, then Y is finite and $|Y| = |X|$.*

Proof. By definition, we have a one to one correspondence $f : [n] \to X$, where $n = |X|$. Therefore $g \circ f : [n] \to Y$ is a one to one correspondence. □

Proposition 5.2. *If a finite set X can be written as a union of two disjoint subsets $Y \cup Z$, then $|X| = |Y| + |Z|$. (Recall that $Y \cup Z = \{x \mid x \in Y \text{ or } x \in Z\}$, and disjoint means their intersection is empty.)*

Proof. Let $f : [n] \to Y$ and $g : [m] \to Z$ be one to one correspondences. Define $h : [n+m] \to X$ by

$$h(i) = \begin{cases} f(i) & \text{if } i < n \\ g(i-n) & \text{if } i \geq n \end{cases}$$

This is a one to one correspondence. □

A *partition* of X is a decomposition of X as a union of subsets $X = Y_1 \cup Y_2 \cup \ldots Y_n$ such that Y_i and Y_j are disjoint whenever $i \neq j$.

Corollary 5.3. *If $X = Y_1 \cup Y_2 \cup \ldots Y_n$ is a partition, then $|X| = |Y_1| + |Y_2| + \ldots |Y_n|$.*

Proof. We have that

$$|X| = |Y_1| + |Y_2 \cup \dots Y_n| = |Y_1| + |Y_2| + |Y_3 \cup \dots Y_n| = \dots = |Y_1| + |Y_2| + \dots |Y_n|$$

□

Given a function $f : X \to Y$ and an element $y \in Y$, the preimage

$$f^{-1}(y) = \{x \in X \mid f(x) = y\}$$

Proposition 5.4. *If $f : X \to Y$ is a function, then*

$$|X| = \sum_{y \in Y} |f^{-1}(y)|$$

Proof. The collection $\{f^{-1}(y)\}$ forms a partition of X. □

The cartesian product of two sets is the set of ordered pairs

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}$$

Theorem 5.5. *If X and Y are finite sets, then $|X \times Y| = |X||Y|$.*

Proof. Let $p : X \times Y \to Y$ be the projection map defined by $p(x, y) = y$. Then

$$p^{-1}(y) = \{(x, y) \mid x \in X\}$$

and $(x, y) \to x$ gives a one to one correspondence to X. Therefore, by the previous corollary,

$$|X \times Y| = \sum_{y \in Y} |p^{-1}(y)| = |Y||X|$$

□

Let us apply these ideas to group theory.

Given a subgroup $H \subset G$ and $g \in G$, let $gH = \{gh \mid h \in H\}$. This is called a (left) coset. For example, when $G = S_3$ and $H = \{I, (123), (321)\}$, the cosets are

$$IH = (123)H = (321)H = H$$

and

$$(12)H = (13)H = (23)H = \{(12), (13), (23)\}$$

Thus the collection of distinct cosets gives a partition of S_3 into rotations and flips, and there are the same number of each. We will prove that is a similar statement in general.

Lemma 5.6. *If two cosets g_1H and g_2H have a nonempty intersection then $g_1H = g_2H$.*

Proof. If $g \in g_1H \cap g_2H$, we can write $g = g_1h_1 = g_2h_2$ with $h_1, h_2 \in H$. Then $g_2 = g_1h_1h_2^{-1}$. If $h \in H$, then $h_1h_2^{-1}h \in H$ because H is a subgroup. Therefore $g_2h = g_1h_1h_2^{-1}h \in g_1H$. This proves that $g_2H \subseteq g_1H$. The same argument, with g_1 and g_2 interchanged, shows that $g_1H \subseteq g_2H$. Therefore these sets are equal. □

Lemma 5.7. *G/H is a partition of G*

Proof. Every element $g \in G$ lies in the coset gH. Therefore G is the union of cosets. By the previous lemma, the cosets are pairwise disjoint. □

Lemma 5.8. *If H is finite, $|gH| = |H|$ for every g.*

Proof. Let $f : H \to gH$ be defined by $f(h) = gh$. Then f is onto. Suppose that $f(h_1) = f(h_2)$. Then $h_1 = g^{-1}gh_1 = g^{-1}gh_2 = h_2$. Therefore f is also one to one. Consequently $|gH| = |H|$. □

Theorem 5.9 (Lagrange)**.** *If $H \subseteq G$ is a subgroup of a finite group, then*

$$|G| = |H| \cdot |G/H|$$

In particular, the order of H divides the order of G.

Proof. By the previous results, G/H is a partition of G into $|G/H|$ sets each of cardinality $|H|$. □

Given $g \in G$, the *order* of g is the smallest positive n such that $g^n = e$. This was shown in a previous exercise to be the order of the subgroup generated by g. Therefore:

Corollary 5.10. *The order of any element $g \in G$ divides the order of G.*

Corollary 5.11. *If the order G is a prime number, then G is cyclic.*

Proof. Let $p = |G|$. By the previous corollary $g \in G$ divides p. If $g \neq e$, then the order must be p. Therefore G is generated by g. □

One can ask whether the converse of the first corollary holds, that is if $|G|$ is divisible by n, does G necessarily have element of order n? The answer is no, it would fail for $n = |G|$ unless G is cyclic. Even if we require $n < |G|$ then it may still fail (exercise 9). However, if n is prime, then it is true.

Theorem 5.12 (Cauchy)**.** *If the order of a finite group G is divisible by a prime number p, then G has an element of order p*

Proof when $p = 2$. Suppose that G is even. We can partition G into $A = \{g \in G \mid g^2 = e\}$ and $B = \{g \in G \mid g^2 \neq e\}$. Therefore $|G| = |A| + |B|$. Every element $g \in B$ satisfies $g \neq g^{-1}$. Therefore $|B|$ is even, because we can write B as a disjoint union of pairs $\{g, g^{-1}\}$. Therefore $|A| = |G| - |B|$ is even. Furthermore $|A| \geq 1$ because $e \in A$. It follows that A contains an element different from e, and this must have order 2. □

Next, we want to develop a method for computing the order of a subgroup of S_n.

Definition 5.13. *Given $i \in \{1, \dots, n\}$, the orbit $\operatorname{Orb}(i) = \{g(i) \mid g \in G\}$. A subgroup $G \subseteq S_n$ is called transitive if for some i, $\operatorname{Orb}(i) = \{1, \dots, n\}$.*

Definition 5.14. *Given subgroup $G \subseteq S_n$ and $i \in \{1, \dots n\}$, the stabilizer of i, is $\operatorname{Stab}(i) = \{f \in G \mid f(i) = i\}$*

Theorem 5.15 (Orbit-Stabilizer theorem)**.** *Given a subgroup $G \subseteq S_n$, and $i \in \{1, \dots, n\}$ then*

$$|G| = |\operatorname{Orb}(i)| \cdot |\operatorname{Stab}(i)|$$

In particular,

$$|G| = n|\operatorname{Stab}(i)|$$

if G is transitive.

Proof. We define a function $f : G \to \operatorname{Orb}(i)$ by $f(g) = g(i)$. The preimage $T = f^{-1}(j) = \{g \in G \mid g(i) = j\}$. By definition if $j \in \operatorname{Orb}(i)$, there exists $g_0 \in T$. We want to show that $T = g_0 \operatorname{Stab}(i)$. In one direction, if $h \in Stab(i)$ then $g_0 h(i) = j$. Therefore $g_0 h \in T$. Suppose $g \in T$. Then $g = g_0 h$ where $h = g_0^{-1} g$. We see that $h(i) = g_0^{-1} g(i) = g_0^{-1}(j) = i$. Therefore, we have established that $T = g_0 \operatorname{Stab}(i)$. This shows that

$$|G| = \sum_{j \in \operatorname{Orb}(i)} |f^{-1}(j)| = \sum_{j \in \operatorname{Orb}(i)} |\operatorname{Stab}(i)| = |\operatorname{Orb}(i)| \cdot |\operatorname{Stab}(i)|$$

□

Corollary 5.16. $|S_n| = n!$

Proof. We prove this by mathematical induction starting from $n = 1$. When $n = 1$, S_n consists of the identity so $|S_1| = 1 = 1!$. In general, assuming that the corollary holds for n, we have prove it for $n+1$. The group S_{n+1} acts transitively on $\{1, \dots, n+1\}$. We want to show that there is a one to one correspondence between $\operatorname{Stab}(n+1)$ and S_n. An element of $f \in \operatorname{Stab}(n+1)$ looks like

$$\begin{pmatrix} 1 & 2 & \dots n & n+1 \\ f(1) & f(2) & \dots f(n) & n+1 \end{pmatrix}$$

Dropping the last column yields a permutation in S_n, and any permutation in S_n extends uniquely to an element of $\operatorname{Stab}(n+1)$ by adding that column. Therefore we have established the correspondence. It follows that $|\operatorname{Stab}(n+1)| = |S_n| = n!$. Therefore

$$|S_{n+1}| = (n+1)|\operatorname{Stab}(n+1)| = (n+1)(n!) = (n+1)!$$

□

5.17 Exercises

1. Given finite sets Y, Z. Prove that $|Y \cup Z| = |Y| + |Z| - |Y \cap Z|$. Recall that the intersection $Y \cap Z = \{x \mid x \in Y \text{ and } x \in Z\}$.

2. If $B \subseteq A$, prove that $|A - B| = |A| - |B|$, where $A - B = \{a \mid a \in A \text{ and } a \notin B\}$. Use this to prove that the set of distinct pairs $\{(x_1, x_2) \in X \times X \mid x_1 \neq x_2\}$ has $|X|^2 - |X|$ elements.

3. We can use the above counting formulas to solve simple exercises in probability theory. Suppose that a 6 sided dice is rolled twice. There are $6 \times 6 = 36$ possible outcomes. Given a subset S of these outcomes, called an *event*, the probability of S occurring is $|S|/36$.

   (a) What is the probability that a five or six is obtained on the first role?

   (b) What is the probability that a five or six is obtained in either (or both) roll(s)?

   (c) What is probability that the same number is rolled twice?

   (d) What is probability that different numbers be obtained for each roll?

   Explain how you got your answers.

4. Let $G \subseteq S_n$ be a subgroup.

   (a) Prove that the stablizer H of an element i is a subgroup of G.

   (b) A subgroup $H \subset G$ is a normal subgroup if $ghg^{-1} \in H$ for all $g \in G$ and $h \in H$. Is the stabilizer a normal subgroup?

5. By the previous results, the order of an element $g \in S_n$ must divide $n!$. We can do much better. Find a better bound using the cycle decomposition.

6. What is the probability that an element of S_5 has order 2?

7. Choose two elements g_1, g_2 from a finite group G. What is the probability that $g_1 g_2 = e$?

8. Determine all the transitive subgroups of S_3.

9. Let $\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \times \ldots \mathbb{Z}_{m_n} = \{(a_1, \ldots, a_n) \mid a_i \in \mathbb{Z}_{m_i}\}$ be the set of vectors.

   (a) Show that this becomes a group using $(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$ with mod m_i arithmetic in each slot.

   (b) Show that the order of this group is $m_1 m_2 \ldots m_n$.

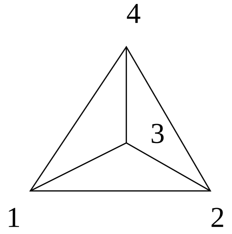
   (c) Let m be the least common multiple of $m_1, \ldots, m_n$. Show that all elements have order dividing m.

10. Prove that Cauchy's theorem holds for the group defined in the previous exercise.

Chapter 6

More counting problems with groups

A polyhedron is a three dimensional version of a polygon. The simplest example is a tetrahedron which is a pyramid with a triangular base.



It is regular if all the triangles, called faces, are equilateral. Let us analyze the rotational symmetries of a regular tetrahedron. Let us call the symmetry group T. We view it as a subgroup of the group S_4 of permutations of the vertices labelled $1, 2, 3, 4$. We can use the orbit-stabilizer theorem to calculate the order of T. Clearly any vertex can be rotated to any other vertex, so the action is transitive. The stabilizer of 4 is the group of rotations keeping it fixed. This consists of the identity I and

$$(123), (132)$$

Therefore $|T| = (4)(3) = 12$. It is easy to list the 9 remaining rotations. There are the rotations keeping 1 fixed:

$$(234), (243)$$

2 fixed:

$$(134), (143)$$

and 3 fixed:

$$(124), (142)$$

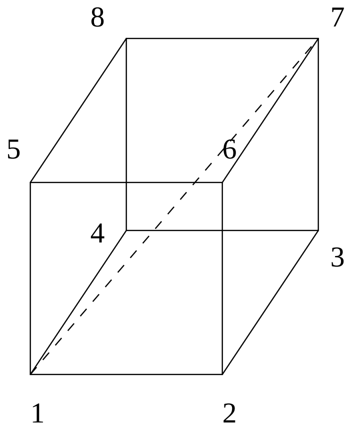


Figure 6.1: cube

We can rotate 180° about the line joining the midpoint of the edges $\overline{13}$ and $\overline{24}$ to get (13)(24). We can do the same thing with other pairs of edges to get

$$(14)(23), (12)(34)$$

In summary

Lemma 6.1. *The symmetry group of a regular tetrahedron is*

$$T = \{I, (123), (132), (134), (143), (124), (142), (13)(24), (14)(23), (12)(34)\}$$

These permutations are exactly the ones than can expressed as an even number of transpositions. This is usually called the alternating group A_4.

Next, we want to analyze the group C of rotational symmetries of the cube

We can view this as a subgroup of S_8. Let us start by writing down all the obvious elements. Of course, we have the identity I. We can rotate the cube 90° about the axis connecting the top and bottom faces to get

$$(1234)(5678)$$

More generally, we have 3 rotations of $90°, 180°, 270°$ fixing each pair of opposite faces. Let us call these type A. There are 2 rotations, other than I, fixing each diagonally opposite pair of vertices such as 1and 7 (the dotted line in picture). Call these type B. For example

$$(254)(368)$$

is type B. We come to the next type, which we call type C. This is the hardest to visualize. To each opposite pair of edges such as $\overline{12}$ and $\overline{78}$, we can connect their midpoints to a get a line L. Now do a 180° rotation about L. Let's count what we have so far:

(I) 1

(A) 3 (rotations) $\times$ 3 (pairs of faces) $= 9$

(B) 2 (rotations) $\times$ 4 (pairs of vertices) $= 8$

(C) 6 (opposite pairs of edges) $= 6$

making 24. To see that this is a complete list, we use the orbit-stabilizer theorem. The action of C is transitive, and Stab(1) consists of I, and two other elements of type B. Therefore $|C| = (8)(3) = 24$. In principle, we have a complete description of C. However, we can do better. There are 4 diagonal lies such as $\overline{17}$. One can see that any non-identity element of C must permute the diagonal lines nontrivially. A bit more formally, we have produced a one to one homomorphism from C to S_4. Since they both have order 24, we can conclude that:

Lemma 6.2. *The symmetry group C of a cube is isomorphic to S_4.*

Let us now turn to counting problems with symmetry.

Question 6.3. *How many dice are there?*

Recall that a die (singular of dice) is gotten by labelling the faces of cube by the numbers 1 through 6. One attempt at a solution goes as follows. Choose some initial labelling, then there as many ways to relabel as there are permutations which is $6! = 720$. This doesn't take into account that there are 24 ways to rotate the cube, and each rotated die should be counted as the same. From this, one may expect that there are $720/24 = 30$ possibilities. This seems more reasonable.

Question 6.4. *How many cubes are there with* 3 *red faces and* 3 *blue?*

Arguing as above, labelling the faces of the cube 1 through 6, there are $\binom{6}{3} = 20$ ways to pick 3 red faces. But this discounts symmetry. On the other hand, dividing by the number of symmetries yields 20/24, which doesn't make sense. Clearly something more sophisticated is required. Let X be a finte set of things such as relabellings of the cube, or colorings of a labelled cube, and suppose that G is a finite set of permutations of X. In fact, we only need to assume that G comes with a homomorphism to S_X. This means that each $g \in G$ determines a permutation of X such that $g_1 g_2(x) = g_1(g_2(x))$ for all $g_i \in G, x \in X$. We say that G acts on X. Given $x \in X$, its orbit $\mathrm{Orb}(x) = \{g(x) \mid g \in G\}$, and let X/G be the set of orbits. Since we really want to x and $g(x)$ to be counted as one thing, we should count the number of orbits. Given $g \in G$, let $Fix(g) = \{x \in X \mid g(x) = x\}$ be the set of fixed points.

Theorem 6.5 (Burnside's Formula)**.** *If G is a finite group acting on a finite set X, then*

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |\,\mathrm{Fix}(g)|$$

Before starting the proof, we define the stablizer $\operatorname{Stab}(x) = \{g \in G \mid g(x) = x\}$. Theorem 5.15 generalizes, with the same proof, to

$$|G| = |\operatorname{Orb}(x)| \cdot |\operatorname{Stab}(x)|$$

Proof of Burnside. Let

$$S = \{(x, g) \in X \times G \mid g(x) = x\}$$

Consider the map $p : S \to G$ given by $p(x, g) = g$. Then $p^{-1}(g) = \operatorname{Fix}(g)$. Therefore proposition 5.4 applied to p yields

$$|S| = \sum_{g \in G} |p^{-1}(g)| = \sum_{g \in G} |\operatorname{Fix}(g)| \tag{6.1}$$

Next consider the map $q : S \to X$ given by $q(x, g) = x$. Then $q^{-1}(x) = \operatorname{Stab}(x)$. Therefore proposition 5.4 applied to q yields

$$|S| = \sum_{x \in X} |q^{-1}(x)| = \sum_{x \in X} |\operatorname{Stab}(x)|$$

Let us write X as disjoint union of orbits $\operatorname{Orb}(x_1) \cup \operatorname{Orb}(x_2) \cup \ldots$, and group terms of the last sum into these orbits

$$|S| = \sum_{x \in \operatorname{Orb}(x_1)} |\operatorname{Stab}(x)| + \sum_{x \in \operatorname{Orb}(x_2)} |\operatorname{Stab}(x)| + \ldots$$

Each orbit $\operatorname{Orb}(x_i)$ has $|G|/|\operatorname{Stab}(x_i)|$ elements by the orbit-stabilizer theorem. Furthermore, for any $x \in \operatorname{Orb}(x_i)$, we have $|\operatorname{Stab}(x)| = |\operatorname{Stab}(x_i)|$. Therefore

$$\sum_{x \in \operatorname{Orb}(x_i)} |\operatorname{Stab}(x)| = \sum_{x \in \operatorname{Orb}(x_i)} |\operatorname{Stab}(x_i)| = \frac{|G|}{|\operatorname{Stab}(x_i)|} |\operatorname{Stab}(x_i)| = |G|$$

Consequently

$$|S| = \sum_{x \in \operatorname{Orb}(x_1)} |G| + \sum_{x \in \operatorname{Orb}(x_2)} |G| + \ldots = |G| \cdot |X/G|$$

Combining this with equation (6.1) yields

$$|G| \cdot |X/G| = \sum_{g \in G} |\operatorname{Fix}(g)|$$

Dividing by $|G|$ yields the desired formula.

□

Let us say that the action of G on X is fixed point free if $\operatorname{Fix}(g) = \emptyset$ unless g is the identity. In this case the naive formula works.

Corollary 6.6. *If the action is fixed point free,*

$$|X/G| = |X|/|G|$$

Coming back to question 6.3. Let X be the set of relabellings of the cube, and $G = C$ the symmetry group of the cube. Then the action is fixed point free, so that $|X/G| = 720/24 = 30$ gives the correct answer.

The solution to question 6.4 using Burnside's formula is rather messy (the answer is 2). So instead, let us consider the simpler question.

Question 6.7. *How many ways can we color a regular tetrahedron with* 2 *red and* 2 *blue faces?*

Let X be the set of such colorings, and let T be the symmetry group. Then

$$\mathrm{Fix}(I) = X$$

has $\binom{4}{2} = 6$ elements. We can see that

$$\mathrm{Fix}(g) = \emptyset$$

for any 3-cycle such as $g = (123)$ because we would need to have 3 faces the same color for any fixed point. For a fixed point of $g = (13)(24)$, the sides adjacent to $\overline{13}$ and $\overline{24}$ would have to be the same color. Therefore

$$|\,\mathrm{Fix}(g)| = 2$$

The same reasoning applies to $g = (14)(23)$ or $(12)(34)$. Thus

$$|X/T| = \frac{1}{12}(6 + 2 + 2 + 2) = 1$$

Of course, this can be figured out directly.

In general, Burnside's formula can be a bit messy to use. In practice, however, there a some tricks to simplify the sum. Given two elements g_1, g_2 of a group, we say that g_1 is *conjugate* to g_2 if $g_1 = hg_2h^{-1}$ for some $h \in G$. Since we can rewrite this as $g_2 = h^{-1}g_1h$, we can see that the relationship is symmetric. Here are a couple of examples

Example 6.8. *Every element g is conjugate to itself because $g = ege^{-1}$.*

Example 6.9. *In the dihedral group D_n, R is conjugate to R^{-1} because $FRF = FRF^{-1} = R^{-1}$.*

An important example is:

Example 6.10. *In S_n, any cycle is conjugate to any other cycle of the same length.*

The relevance for counting problems is as follows.

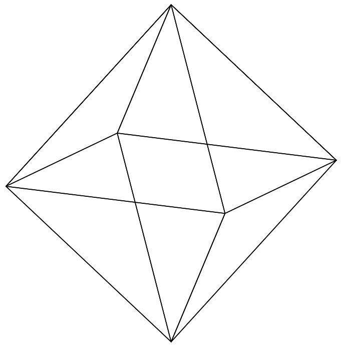
Proposition 6.11. *With the same assumptions as in theorem 6.5, if g_1 is conjugate to g_2, then $|\operatorname{Fix}(g_1)| = |\operatorname{Fix}(g_2)|$*

Proof. Suppose that $g_1 = hg_2h^{-1}$ and $x \in \operatorname{Fix}(g_2)$. Then $g_2 \cdot x = x$. Therefore $g_1(hx) = hg_2h^{-1}hx = hx$. This means that $hx \in \operatorname{Fix}(g_1)$. So we can define a function $f : \operatorname{Fix}(g_2) \to \operatorname{Fix}(g_1)$ by $f(x) = hx$. This has an inverse $f^{-1} : \operatorname{Fix}(g_1) \to \operatorname{Fix}(g_2)$ given by $f(y) = h^{-1}y$. Since f has inverse, it must be one to one and onto. Therefore $|\operatorname{Fix}(g_1)| = |\operatorname{Fix}(g_2)|$. □

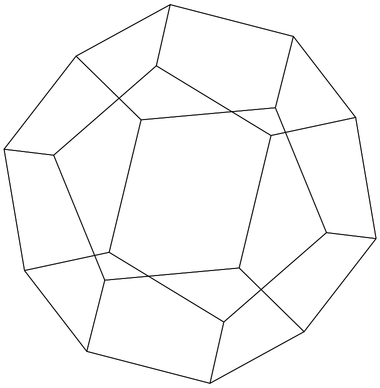
The moral is that we only need to calculate $|\operatorname{Fix}(g)|$ once for every element conjugate to g, and weight the factor in Burnside by the number of such elements.

6.12 Exercises

1. Calculate the order of the (rotational) symmetry group for the octrahedron (which most people would call a diamond)



2. Calculate the order of the (rotational) symmetry group for the dodecahedron



(There are 20 vertices, and 12 pentagonal faces.)

3. Let C^+ be the group of symmetries of the cube including both rotations and reflections such as (12)(34)(56)(78) with labeling in figure 6.1. Calculate the order of C^+. Repeat for the tetrahedron, octahedron and dodecahedron.

4. How many ways are the color the sides of a tetrahedron with 2 red faces, 1 and blue and 1 green. (Even if the answer is obvious to you, use Burnside.)

5. Answer question 6.4 using Burnside.

6. How many ways are there to color the sides of a cube with 2 red faces and 4 blue? (Same instructions as above.)

7. Given a group G, and a subgroup H, show that G acts transitively on G/H by $g(\gamma H) = g\gamma H$. Calculate the stabilizer of γH.

8. Prove Cayley's theorem that every group is isomorphic to a group of permutations. (Hint: Use the action G on itself defined as in the previous problem, and use this to construct a one to one homomorphism $G \to S_G$.)

9. Explain why the statement of example 6.10 holds for $n = 3$, and then do the general case.

Chapter 7

Kernels and quotients

Recall that homomorphism between groups $f : G \to Q$ is a map which preserves the operation and identity (which we denote by $\cdot$ and e). It need not be one to one. The failure to be one to one is easy to measure.

Definition 7.1. *Given a homomorphism between groups $f : G \to Q$, the kernel $\ker f = \{g \in G \mid f(g) = e\}$.*

Lemma 7.2. *A homomorphism is one to one if and only if $\ker f = \{e\}$.*

The proof will be given as an exercise. The kernel is a special kind of subgroup. It's likely that you already encountered this notion in linear algebra in the context of linear transformations. There it also called the kernel or sometimes the null space.

Definition 7.3. *A subgroup $H \subset G$ is called normal if $ghg^{-1} \in H$ for all $g \in G$ and $h \in H$. The operation $h \mapsto ghg^{-1}$ is called conjugation of h by g. So normality of H means that it is closed under conjugation by elements of G.*

Proposition 7.4. *Suppose that $f : G \to Q$ is a homomorphism, then $\ker f$ is a normal subgroup.*

Proof. Let $h_1, h_2 \in H$ and $g \in G$. Then $f(h_1h_2) = f(h_1)f(h_2) = e$, $f(h_1^{-1}) = e$, $f(gh_1g^{-1}) = f(g)f(g)^{-1} = e$. □

Here are some examples.

Example 7.5. *If G is abelian, then any subgroup is normal.*

Example 7.6. *In S_3, $H = \{I, (123), (321)\}$ is a normal subgroup. The subgroup $\{I, (12)\}$ is not normal because (12) is conjugate to (13) and (23).*

We want to prove that every normal subgroup arises as the kernel of a homomorphism. This involves the quotient construction. Given subsets $H_1, H_2 \subset G$ of a group, define their product by

$$H_1H_2 = \{h_1h_2 \mid h_1 \in H_1, h_2 \in H_2\}$$

Lemma 7.7. *If $H \subseteq G$ is normal, then the product of cosets satisfies $(g_1H)(g_2H) = (g_1g_2)H$.*

Proof. By definition, $(g_1H)(g_2H) = \{g_1h_1g_2h_2 \mid h_1, h_2 \in H\}$. Since H is normal, $h_3 = g_2^{-1}h_1g_2 \in H$. Therefore $g_1h_1g_2h_2 = g_1g_2h_3h_2 \in (g_1g_2)H$. This proves $(g_1H)(g_2H) \subseteq (g_1g_2)H$.

For the reverse inclusion $(g_1g_2)H \subseteq (g_1H)(g_2H)$, observe that if $h \in H$, then $g_1g_2h = (g_1e)(g_2h) \in (g_1H)(g_2H)$.

□

Theorem 7.8. *If $H \subseteq G$ is a normal subgroup, then G/H becomes a group with respect to the product defined above. The map $p(g) = gH$ is a homomorphism with kernel H.*

Proof. By the previous lemma, $(gH)(eH) = gH = (eH)(gH)$, $(gH)(g^{-1}H) = H = (g^{-1}H)(gH)$, and $(g_1H)(g_2Hg_3H) = g_1g_2g_3H = (g_1Hg_2H)(g_3H)$. So G/H is a group. Also $p(g_1g_2) = g_1g_2H = (g_1H)(g_2H) = p(g_1)(g_2)$, so p is a homomorphism. Furthermore, $\ker p = \{g \in G \mid gH = H\} = H$. □

When H is normal, we refer to G/H as the *quotient group*. Quotient groups often show up indirectly as follows.

Lemma 7.9. *Let $f : G \to H$ be a homomorphism with kernel $K = \ker f$. Then the image $f(G) = \{f(g) \mid g \in G\}$ is a subgroup isomorphic to G/K. In particular, H is isomorphic to G/K if f is onto.*

The proof will be given as an exercise. The quotient construction can be used to tie up some loose ends from earlier sections. Let n be a positive integer, and let $n\mathbb{Z} = \{nx \mid x \in \mathbb{Z}\}$. This is a subgroup. So we can form the quotient $\mathbb{Z}_n^{new} = \mathbb{Z}/n\mathbb{Z}$. The label "new" is temporary, and is there to distinguish it from $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$. Given an integer x, let $\bar{x} = x + n\mathbb{Z}$. In particular, $x \mapsto \overline{x}$ gives a map from $\mathbb{Z}_n \to \mathbb{Z}_n^{new}$. We leave it as an exercise to show this is a one to one correspondence, and that

$$\overline{x \oplus y} = \overline{x} + \overline{y}$$

where $+$ on the right is addition in the quotient group. Thus, we can conclude that the old and new versions of $\mathbb{Z}_n$ are isomorphic, and we will conflate the two. Recall, in fact, that we never fully completed the proof that the old $\mathbb{Z}_n$ was a group. Now we don't have to!

Normal subgroups can be used to break up complicated groups into simpler pieces. For example, in the exercises, we will see that the dihedral group D_n contains a cyclic subgroup C_n, which is normal and the quotient D_n/C_n is also cyclic. Here we look at the related example of the orthogonal group $O(2)$. This is the full symmetry group of the circle which includes rotations and reflection. The rotations form a subgroup $SO(2)$.

Proposition 7.10. *$SO(2)$ is a normal subgroup of $O(2)$.*

We give two proofs. The first, which uses determinants, gets to the point quickly. However, the second proof is also useful since it leads to the formula (7.1).

First Proof. We start with a standard result.

Theorem 7.11. *For any pair of 2×2 matrices A and B, $\det AB = \det A \det B$.*

Proof. A brute force calculation shows that

$$(a_{11}a_{22} - a_{12}a_{21})(b_{11}b_{22} - b_{12}b_{21})$$

and

$$(a_{11}b_{11} + a_{12}b_{21})(a_{21}b_{12} + a_{22}b_{22}) - (a_{11}b_{12} + a_{12}b_{22})(a_{21}b_{11} + a_{22}b_{22})$$

both can be expanded to

$$a_{11}a_{22}b_{11}b_{22} - a_{11}a_{22}b_{12}b_{21} - a_{12}a_{21}b_{11}b_{22} + a_{12}a_{21}b_{12}b_{21}$$

□

Therefore $\det : O(2) \to \mathbb{R}^*$ is a homomorphism, where $\mathbb{R}^*$ denote the group of nonzero real numbers under multiplication. It follows that $SO(2)$ is the kernel. So it is normal. □

Second Proof. We have to show that $AR(\theta)A^{-1} \in SO(2)$ for any $A \in O(2)$. This is true when $A \in SO(2)$ because $SO(2)$ is a subgroup.

It remains to show that conjugating a rotation by a reflection is a rotation. In fact we will show that for any reflection A

$$AR(\theta)A^{-1} = R(-\theta) \tag{7.1}$$

First let A be the reflection $F = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ about the x-axis. Then an easy calculation shows that $FR(\theta)F^{-1} = FR(\theta)F = R(-\theta)$. Now assume that A is a general reflection. Then

$$A = \begin{bmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{bmatrix} = FR(-\phi)$$

So

$$AR(\theta)A^{-1} = FR(-\phi)R(\theta)R(\phi)F = R(-\theta)$$

as claimed. □

So now we have a normal subgroup $SO(2) \subset O(2)$ which we understand pretty well. What about the quotient $O(2)/SO(2)$. This can identified with the cyclic group $\{\pm 1\} \subset \mathbb{R}^*$ using the determinant.

7.12 Exercises

1. Prove lemma 7.2.

2. Determine the normal subgroups of S_3.

3. Prove lemma 7.9. (Hint: first prove that $f(G)$ is subgroup. Then that $\bar{f}(gH) = f(g)$ is a well defined function which gives an isomorphism $G/K \cong f(G)$.)

4. (a) Given a group G and a normal subgroup H. Let $S \subset G$ be a subset with the property that $S \cap gH$ has exactly one element for every $g \in G$. Show that the restriction of p gives a one to one correspondence $S \to G/H$.

   (b) Show that these conditions hold for $G = \mathbb{R}$, $H = 2\pi\mathbb{Z}$ and $S = [0, 2\pi)$.

5. Prove that $\mathbb{Z}_n$ is isomorphic to the quotient group $\mathbb{Z}/n\mathbb{Z}$ as claimed earlier.

6. Check that $SL_2(\mathbb{R}) = \{A \in GL_2(\mathbb{R}) \mid \det A = 1\}$ is a normal subgroup of $GL_2(\mathbb{R})$.

7. In an earlier exercise in chapter , you showed that the set of upper triangular matrices

$$\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$$

   is a subgroup of $GL_2(\mathbb{R})$. Is it normal?

8. Let $H \subseteq G$ be a normal subgroup $f : G \to K$ be an onto homomorphism, prove that $f(H) = \{f(h) \mid h \in H\}$ is a normal subgroup. What if f is not onto?

9. Given a group G, its *center* $Z(G)$ is the set of elements c which satisfy $cg = gc$ for every $g \in G$.

   (a) Prove that the center is an abelian normal subgroup.

   (b) Does an abelian normal subgroup necessarily lie in the center? (Think about the dihedral group.)

10. Check that the center of S_n, when $n > 2$, is trivial in the sense that it consists of only the identity.

Chapter 8

Rings and modular arithmetic

So far, we have been working with just one operation at a time. But standard number systems, such as $\mathbb{Z}$, have two operations $+$ and $\cdot$ which interact. It is useful to give a name to this sort of thing.

Definition 8.1. *A ring consists of a set R with elements $0, 1 \in R$, and binary operations $+$ and $\cdot$ such that: $(R, +, 0)$ is an Abelian group, $\cdot$ is associative with 1 as the identity, and $\cdot$ distributes over $+$ on the left and right:*

$$x \cdot (y + z) = x \cdot y + x \cdot z$$

$$(y + z) \cdot x = y \cdot x + z \cdot x$$

Definition 8.2. *A ring is commutative if in addition*

$$x \cdot y = y \cdot x$$

Here are some basic examples that everyone should already know.

Example 8.3. *Let $\mathbb{Z}$ (respectively $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$) be the set of integers (respectively rational numbers, real numbers, complex numbers) with the usual operations. These are all commutative rings.*

Example 8.4. *The set $M_{nn}(\mathbb{R})$ of $n \times n$ matrices over $\mathbb{R}$ with the usual matrix operations forms a ring. It is not commutative when $n > 1$.*

We now focus on a new example. Let n be a positive integer, and write $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$, where $\overline{x} = x + n\mathbb{Z}$. We already know that this has an addition given by addition of cosets:

$$\overline{a} + \overline{b} = \overline{a + b}$$

For hereon in, we'll stop writing $\oplus$. We will try to define multiplication the same way by

$$\overline{a}\overline{b} = \overline{ab}$$

However, we have to prove that this definition makes sense. In other words, we have to show that right side depends only on $\overline{a}$ and $\overline{b}$ rather than a and b.

Lemma 8.5. *If $\overline{a} = \overline{a'}$ and $\overline{b} = \overline{b'}$, then $\overline{ab} = \overline{a'b'}$*

Proof. The equality $\overline{x} = \overline{x'}$ holds if and only if $x - x'$ is divisible by n. Therefore $a' = a + nx$ and $b = b' + ny$ for some $x, y \in \mathbb{Z}$. It follows that $a'b' = ab + n(xb' + ya' + nxy)$.

□

Theorem 8.6. *$\mathbb{Z}_n$ is a commutative ring.*

Proof. The laws follow from the fact that $\mathbb{Z}$ is a commutative ring, the definition of the operations in $\mathbb{Z}_n$, and the fact that the map $\mathbb{Z} \to \mathbb{Z}_n$ is onto. For example, here is a proof of the distributive law

$$(\overline{x} + \overline{y})\overline{z} = (\overline{x+y})\overline{z} = \overline{(x+y)z}$$

□

When it's clear we're working in $\mathbb{Z}_n$, we usually just write x instead of $\bar{x}$. To get a feeling for modular multiplication, lets write down the table for $\mathbb{Z}_6$

$\cdot$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

One curious fact is that some nonzero numbers, such as 2, can be multiplied by other nonzero numbers to get 0. We say that such a number is a *zero divisor.*

Lemma 8.7. *An element $\overline{m} \in \mathbb{Z}_n$ is a zero divisor if $m > 1$ and m divides n.*

Proof. We have that $n = mm'$ for some $0 < m' < n$. So that $\overline{m}\overline{m'} = \overline{0}$ □

Also notice that the number 5 has a reciprocal, namely 5.

Definition 8.8. *An element $x \in R$ of a ring is invertible if there exists an element y such that $xy = yx = 1$. Let R^* denote the set of invertible elements. (When R is commutative, invertible elements are also called units.)*

Lemma 8.9. *If R is a ring R^* is a group with respect to multiplication.*

This will be proven in the exercises. The group of invertible elements are easy to determine for the previous examples. For example, $M_{nn}(\mathbb{R})^* = GL_n(\mathbb{R})$.

Given two integers a, b, a common divisor is an integer d such that $d|a$ and $d|b$. The greatest common divisor is exactly that, the common divisor greater than or equal to all others (it exists since the set of common divisors is finite). We denote this by $\gcd(a, b)$.

Lemma 8.10 (Euclid)**.** *If a, b are natural numbers then $\gcd(a, b) = \gcd(b, a \mathit{mod}\, b)$*

Proof. Let $r = a \mathit{mod}\, b$. Then the division algorithm gives $a = qb + r$ for some integer q. SInce $\gcd(b, r)$ divides b and r, it divides $qb + r = a$. Therefore $\gcd(b, r)$ is a common divisor of a and b, so that that $\gcd(b, r) \leq \gcd(a, b)$. On the other hand, $r = a - qb$ implies that $\gcd(a, b)|r$. Therefore $\gcd(a, b)$ is a common divisor of b and r, so $\gcd(a, b) \leq \gcd(b, r)$, which forces them to be equal. □

This lemma leads to a method for computing gcds. For example

$$\gcd(100, 40) = \gcd(40, 20) = \gcd(20, 0) = 20.$$

For our purposes, a *diophantine equation* is an equation with integer coefficients where the solutions are also required to be integers. The simplest examples are the linear ones: given integers a, b, c, find all integers m, n such that $am + bn = c$.

Theorem 8.11. *Given integers a, b, c, $am+bn = c$ has a solution with $m, n \in \mathbb{Z}$ if and only if $\gcd(a, b)|c$.*

Proof. Since $(m', n') = (\pm m, \pm n)$ is a solution of $\pm an' + \pm bm' = c$, we may as well assume that $a, b \geq 0$. We now prove the theorem for natural numbers a, b by induction on the minimum $min(a, b)$.

If $min(a, b) = 0$, then one of them, say $b = 0$. Since $a = \gcd(a, b)$ divides c by assumption, $(c/a, 0)$ gives a solution of $am + bn = c$. Now assume that $a'm + b'n = c'$ has a solution whenever $min(a', b') < min(a, b)$ and the other conditions are fulfilled. Suppose $b \leq a$, and let $r = r(a, b) = a \mathit{mod}\, b$ and $q = q(a, b)$ be given as in theorem 4.5. Then $rm' + bn' = c$ has a solution since $min(r, b) = r < b = min(a, b)$ and $\gcd(b, r) = \gcd(a, b)$ divides c. Let $m = n'$ and $n = m' - qn'$, then

$$am + bn = an' + b(m' - qn') = bm' + rn' = c.$$

□

From the last proof, we can deduce:

Corollary 8.12. *Given $a, b \in \mathbb{Z}$, there exists $m, n \in \mathbb{Z}$ such that $am + bn = \gcd(a, b)$.*

We can now determine the invertible elements

Theorem 8.13. *$m \in \mathbb{Z}_n$ is invertible if and only if* $\gcd(m, n) = 1$ *(we also say that m and n are* relatively prime *or* coprime*).*

Proof. If $gcd(m, n) = 1$, then $mm' + nn' = 1$ or $mm' = -n'n + 1$ for some integers by corollary 8.12. After replacing (m', n') by $(m' + m''n, n' - m'')$ for some suitable m'', we can assume that $0 \leq m' \leq n$. Since have $r(mm', n) = 1$, $mm' = 1$.

The converse follows by reversing these steps. □

Definition 8.14. *A ring is called a division ring if $R^* = R - \{0\}$. A commutative division ring is called a field.*

For example $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are fields. We will see a noncommutative division ring later on. The previous theorem implies the following:

Theorem 8.15. *The ring $\mathbb{Z}_n$ is a field if and only if n is prime.*

Corollary 8.16 (Fermat's little theorem)**.** *When p is a prime and n and integer, then p divides $n^p - n$.*

Proof. If p divides n, then clearly it divides $n^p - n$. Now suppose that p does not divide n, then $\overline{n} \in \mathbb{Z}_p^*$. This is a group of order $p-1$. So by Lagrange's theorem, $\overline{n}$ has order dividing $p - 1$. This implies that $\overline{n}^{p-1} = \overline{1}$, or that $\overline{n}^{p-1} - \overline{1} = \overline{0}$. This implies that p divides $n^{p-1} - 1$ (which is usually taken as the statement of Fermat's little theorem) and therefore $n^p - n$. □

8.17 Exercises

1. Let R be a commutative ring. Prove that $0 \cdot x = 0$. (This might appear to be a completely obvious statement, but it isn't – the only things you know about R are what follows from the axioms.)

2. Let R be a commutative ring. Prove that $(-1) \cdot x = -x$, where $-x$ is the additive inverse of x, that is $(-x) + x = 0$.

3. The Gaussian integers $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$, where $i = \sqrt{-1}$.

   (a) Check that is closed under addition, additive inverses and multiplication, and is therefore a ring.

   (b) Determine the group $\mathbb{Z}[i]^*$ of invertible elements.

4. Check that the Gaussian field $\mathbb{Q}[i] = \{a + bi \mid a, b \in \mathbb{Q}\}$ is a field when equipped with the usual operations.

5. Prove that there are no zero divisors in a field, i.e. if $xy = 0$ then $x = 0$ or $y = 0$.

6. If R_1 and R_2 are commutative rings, define $R = R_1 \times R_2$ with operations $(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$ and $(a_1, a_2) \cdot (b_1, b_2) = (a_1 b_1, a_2 b_2)$. Check that this is a commutative ring with appropriate choice of constants. Show that this has zero divisors.

7. An element x of a commutative ring is nilpotent if $x^N = 0$ for some integer $N \geq 0$. Determine the nilpotent elements of $\mathbb{Z}_n$.

8. Prove that the sum and product of nilpotent elements in a commutative ring are also nilpotent.

9. Sequences of "random" numbers are often generated on a computer by the following method: Choose integers $n \geq 2, a, b, x_0$, and consider the sequence
$$x_{i+1} = (ax_i + b) \, mod \, n.$$
This sequence will eventually repeat itself. The period is the smallest k such that $x_{i+k} = x_i$ for all i large enough. Obviously, short periods are less useful, since the pattern shouldn't be too predictable.

   (a) Prove that the period is at most n.

   (b) Explain why picking a nilpotent in $\mathbb{Z}_n$ would be a really bad choice.

Chapter 9

$\mathbb{Z}_p^*$ is cyclic

Given a field K, a polynomial in x is a symbolic expression

$$a_n x^n + a_{n_1} x^{n-1} + \ldots + a_0$$

where $n \in \mathbb{N}$ is arbitrary and the coefficients $a_n, \ldots, a_0 \in K$. Note that polynomials are often viewed as functions but it is important to really treat these as expressions. First of all the algebraic properties become clearer, and secondly when K is finte, there only finitely many functions from $K \to K$ but infinitely many polynomials. We denote the set of these polynomials by $K[x]$. We omit terms if the coefficients are zero, so we can pad out a polynomial with extra zeros whenever convenient e.g. $1 = 0x^2 + 0x + 1$. The highest power of x occurring with a nonzero coefficient is called the degree. We can add polynomials by adding the coefficients

$$f = a_n x^n + a_{n_1} x^{n-1} + \ldots + a_0$$

$$g = b_n x^n + b_{n_1} x^{n-1} + \ldots + b_0$$

$$f + g = (a_n + b_n) x^n + \ldots (a_0 + b_0)$$

Multiplication is defined using the rules one learns in school

$$\begin{aligned} fg &= (a_0 b_0) + (a_1 b_0 + a_0 b_1) x + (a_2 b_0 + a_1 b_1 + a_0 b_2) x^2 + \ldots \\ &= (\sum_{i+j=k} a_i b_j) x^k \end{aligned}$$

Theorem 9.1. *$K[x]$ is a commutative ring with the operations described above.*

Proof. This is fairly routine, so we just list a few steps. Let f and g be as above and

$$h = c_n x^n + c_{n-1} x^{n-1} + \ldots c_0$$

Then

$$f(gh) = (\sum_{i+j+k=\ell} a_i b_j c_k) x^\ell = (fg)h$$

and

$$f(g+h) = \sum_{i+j=k} a_i(b_j + c_j)x^n = fg + fh$$

□

Given a polynomial $f \in K[x]$ and an element $a \in K$, we can substitute x by a in f to get an element that we write as $ev_a(f) = f(a)$. The following is easy to check and true by design.

Lemma 9.2. *Given polynomials f and g, $ev_a(1) = 1$, $ev_a(f+g) = ev_a(f) + ev_a(g)$ and $ev_a(fg) = ev_a(f)ev_a(g)$.*

The lemma says that $ev_a : K[x] \to K$ is a *ring homomorphism.* An element $r \in K$ is called a *zero or root* of a polynomial f if $f(r) = 0$.

Lemma 9.3. *An element $r \in K$ is a root of $f \in K[x]$ if and only if $x - a$ divides f, i.e there exists $g \in K[x]$ with $f = (x-a)g$*

Proof. If $f = (x-r)g$, then clearly $f(r) = (r-r)g(r) = 0$.

Now suppose that $f(r) = 0$. Write $f = a_nx^n + \ldots + a_0$. We want $g = b_{n-1}x^{n-1} + \ldots b_0$ such that $(x-r)g = f$. This equation is equivalent to the system

$$\begin{aligned} b_{n-1} &= a_n \\ b_{n-2} - rb_{n-1} &= a_{n-1} \\ &\ldots \\ b_0 - rb_1 &= a_1 \\ -rb_0 &= a_0 \end{aligned}$$

There are $n+1$ linear equations in n variables $b_{n-1}, \ldots, b_0$, so there is no guarantee that we have a solution. However, if we can eliminate one of the equations by doing "row operations" (adding multiples of one equation to another etc.), then the new system will have n equations, and this would be consistent. Adding r^n times the first equation to the last equation, and then r^{n-1} times the second equation to the last and so on leads to

$$r^nb_{n-1} + r^{n-1}(b_{n-2} - rb_{n-1}) + \ldots - rb_0 = a_nr^n + a_{n-1}r^{n-1} + \ldots a_0$$

or $0 = 0$. So we have eliminated the last equation. The remaining n equations can now be solved to obtain g as above. □

Theorem 9.4. *If $f \in K[x]$ polynomial of degree n, then it has at most n distinct roots. If f has exactly n roots $r_1, \ldots, r_n$, then $f = c(x-r_1)\ldots(x-r_n)$ where c is the leading coefficient of f.*

Proof. Suppose that r_1 is a root of f. Then $f = (x-r_1)g$ by the previous lemma. The roots of f are r_1 together with the roots of g. By induction on the degree, we can assume that g has at most $n-1$ roots. By the same reasoning, if $r_1, \ldots, r_n$ are roots, $f = c(x-r_1)\ldots(x-r_n)$ for some $c \in K$. But clearly c must equal the leading coefficient. □

We now apply these results to the field $K = \mathbb{Z}_p$, where p is a prime. Sometimes this is denoted by $\mathbb{F}_p$ to emphasize that its a field. When the need arises, let us write $\overline{a}$ to indicate we are working $\mathbb{Z}_p$, but we won't bother when the context is clear.

Proposition 9.5. *We can factor $x^p - x = x(x-1)(x-2)\ldots(x-(p-1))$ in $\mathbb{Z}_p[x]$*

Proof. By Fermat's little theorem, $1\ldots, p-1$ are roots. Therefore $x^p - x = x(x-1)(x-2)\ldots(x-p-1)$ in $\mathbb{Z}_p[x]$. □

Corollary 9.6 (Wilson's theorem)**.** $\overline{(p-1)!} = -\overline{1}$

Proof. We have $x^{p-1} - 1 = (x-1)(x-2)\ldots(x-(p-1))$. Now evaluate both sides at 0. □

Corollary 9.7. *The binomial coefficients $\binom{p}{n} = \frac{p!}{n!(p-n)!}$ are divisible by n when $1 < n < p$.*

Proof. Substitue $1+x$ into the above identity to obtain $(1+x)^p - (1+x) = 0$ in $\mathbb{Z}_p$. Now expand using the binomial theorem, which is valid in any field (see exercises), to obtain

$$\sum_{n=1}^{p-1} \overline{\binom{p}{n}} x^n = 0$$

□

The last few results were fairly easy, the next result is not.

Theorem 9.8. *If p is prime, then $\mathbb{Z}_p^*$ is cyclic.*

Proof in a special case. We won't prove this in general, but to get some sense of why this is true, let's prove it when $p = 2q+1$, where q is another prime. This is not typical, but it can certainly happen (e.g. $p = 7, 11, 23, \ldots$). Then $\mathbb{Z}_p^*$ has order $2q$. The possible orders of its elements are $1, 2, q$, or $2q$. There is only element of order 1, namely 1. An element of order 2 is a root of $x^2 - 1$, so it must be $-\overline{1}$. An element of order q satisfies $x^q - 1 = 0$, and be different from 1. Thus there are at most $q-1$ possibilities. So to summarize there are no more $q+1$ elements of orders $1, 2, q$. Therefore there are at least $q-1$ elements of order $2q$, and these are necessarily generators. □

9.9 Exercises

1. Given a field K and a positive integer n, let $\overline{n} = 1 + \ldots + 1$ (n times). K is said to have *positive characteristic* if $\overline{n} = 0$ for some positive n, otherwise K is said to have characteristic 0. In the positive characteristic case, the smallest $n > 0$ with $\overline{n} = 0$ is called the *characteristic*. Prove that the characteristic is a prime number.

2. For any field, prove the binomial theorem

   $$(x+1)^n = \sum_{m=0}^{n} \overline{\binom{n}{m}} x^m$$

   (Recall $\binom{n+1}{m} = \binom{n}{m} + \binom{n+1}{m}$.)

3. Let K be a field and $s \in K$. Let $K[\sqrt{s}]$ be the set of expressions $a + b\sqrt{s}$, with $a, b \in K$. Show that this becomes a commutative ring if we define addition and multiplication as the notation suggests:

   $$(a + bi\sqrt{s}) + (c + d\sqrt{s}) = (a + c) + (b + d)\sqrt{s}$$

   $$(a + b\sqrt{s})(c + d\sqrt{s}) = (ac + bds) + (ad + bc)\sqrt{s}$$

4. Show $K[\sqrt{s}]$ has zero divisors if $x^2 - s = 0$ has a root. If this equation does not have a root, then prove that $K[\sqrt{s}]$ is a field (Hint: $(a+b\sqrt{s})(a-b\sqrt{s}) =?$ and when is it zero?).

5. When p is an odd prime, show that the map $x \mapsto x^2$ from $\mathbb{Z}_p^* \to \mathbb{Z}_p^*$ is not onto. Use this fact to construct a field with p^2 elements and characteristic p.

Chapter 10

Matrices over $\mathbb{Z}_p$

We can now combine everything we've learned to construct a new, and interesting, collection of finite groups. Let p be a prime number. Then $\mathbb{Z}_p$ is a field, which means that we can perform all the usual operations in it, including division. This allows us to do linear algebra over this field pretty much as usual. For instance, we can consider vectors of length n with entries in $\mathbb{Z}_p$. We denote this by $\mathbb{Z}_p^n$. This becomes an abelian group under vector addition:

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$$

where, of course, + on the right denotes addition in $\mathbb{Z}_p$. One can also consider matrices with entries in $\mathbb{Z}_p$. The standard results from basic linear algebra generalize to $\mathbb{Z}_p$ or any field. For example,

Theorem 10.1 (Gaussian elimination)**.** *Let A be an $n \times n$ matrix with entries in a field K. A is invertible if and only if it be can be taken to the identity matrix by a finite sequence of row operations (interchanges, addition of a multiple of one row to another, multiplication of a row by an element of K^*). A is not invertible if and only if it can be taken to a matrix with a zero row.*

Some details will be recalled in the exercises. Let us denote the set of invertible $n \times n$ matrices with entries in K by $GL_n(K)$. This is a group under matrix multiplication. When $K = \mathbb{Z}_p$, this is a finite group. So the first thing we should do is calculate its order. Let us start with the 2×2 case. Let $V = \mathbb{Z}_p^2$ as above, but now represented by 2×1 vectors. $A \in GL_2(\mathbb{Z}_p)$ will act on $v \in \mathbb{Z}_p^2$ by matrix multiplication Av. Set $v = [1, 0]^T \in V$

Lemma 10.2. $\mathrm{Orb}(v) = \mathbb{Z}_p^2 - \{0\}$.

Proof. Given $u \in \mathbb{Z}_p^2 - \{0\}$, we can clearly find a matrix $A \in GL_2(\mathbb{Z}_p)$ with u as its first column. This will satisfy $Av = u$. □

Lemma 10.3. $\mathrm{Stab}(v)$ *is the set of matrices*

$$\left\{ \begin{bmatrix} 1 & x \\ 0 & y \end{bmatrix} \mid y \neq 0 \right\}$$

Proof. The condition $Av = v$ means that the first column is v. □

Theorem 10.4. *The order of $GL_2(\mathbb{Z}_p)$ is $(p^2-1)(p^2-p)$*

Proof. From the last two lemmas and the orbit-stabilizer theorem, the order is $(p^2-1)(p-1)p$ □

Corollary 10.5. *$GL_2(\mathbb{Z}_2)$ is isomorphic to S_3.*

Proof. $GL(\mathbb{Z}_2)$ acts on $\mathbb{Z}_2^2 - \{0\}$ which has 3 elements. Therefore we have a homomorphism $f : G \to S_3$ which is one to one because $\ker f$ consists matrices satisfying $A[1,0]^T = [1,0]^T$ and $A[0,1]^T = [0,1]^T$, and $A = I$ is the only such matrix. Since the order of $GL(\mathbb{Z}_2)$ is 6, f has to be onto as well. □

This isomorphism should be viewed as something of an accident. For $p = 3$, the order is 48 which is not the factorial of anything.

Bolstered by this success, let's try to compute the order of $GL_n(\mathbb{Z}_p)$ which is the group of invertible $n \times n$ matrices.

Theorem 10.6. *The order of $GL_n(\mathbb{Z}_p)$ is $(p^n-1)(p^n-p)\dots(p^n-p^{n-1})$*

Proof. We will apply the same strategy as before. This time $GL_n(\mathbb{Z}_p)$ acts on $\mathbb{Z}_p^n - \{0\}$, and arguing as before, we can see that this is the orbit of $v = [1,0,0,\dots]^T$. So $|GL_n(\mathbb{Z}_p)| = (p^n-1)|\operatorname{Stab}(v)|$ by the orbit-stabilizer theorem. The stabilizer of v is the set of matrices of the form

$$A = \begin{bmatrix} 1 & x_2 & x_3 & \dots \\ 0 & & & \\ 0 & & B & \\ \vdots & & & \end{bmatrix}$$

One can see that for A to be invertible, the $(n-1)\times(n-1)$ block labelled B must be invertible as well, but there are no constraints on the elements labelled x_i. Therefore $A \mapsto (B, x_2, x_3, \dots)$ gives a one to one correspondence between $\operatorname{Stab}(v)$ and $GL_{n-1}(\mathbb{Z}_p) \times \mathbb{Z}_p^{n-1}$. It follows by induction on n that

$$\begin{aligned} |\operatorname{Stab}(v)| &= p^{n-1}|GL_{n-1}(\mathbb{Z}_p)| \\ &= p^{n-1}(p^{n-1}-1)\dots(p^{n-1}-p^{n-2}) \\ &= (p^n-1)\dots(p^n-p^{n-1}) \end{aligned}$$

□

10.7 Exercises

1. A matrix over a field is called elementary if it can be obtained from I by a single row operation. Check that if E is an elementary matrix, it is invertible, and that EA is the matrix obtained from A by a row operation.

2. The one fact from linear algebra, we will just accept is the rank-nullity theorem, which implies that a square matrix is invertible if its kernel contains only 0. If $E_1, \dots, E_N$ are elementary matrices such that $E_N \dots E_1 A = I$, then prove that A is invertible and that $E_N \dots E_1 = A^{-1}$.

3. If $E_1, \dots, E_N$ are elementary matrices such that $E_N \dots E_1 A$ has a row of zeros, prove that A is not invertible. (Hint: show that $\ker A$ contains a nonzero vector.)

4. Determine which of the following matrices over $\mathbb{Z}_2$ is invertible, and find the inverse when it exists.

   (a) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

   (b) $B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

5. Cauchy's theorem would imply that $GL_2(\mathbb{Z}_p)$ would have an element of order p. Show that $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ would work.

6. The determinant $\det : GL_2(\mathbb{Z}_p) \to \mathbb{Z}_p^*$ gives a homomorphism. Show that this is onto, and use this to compute the order of the kernel (which is usually denoted as $SL_2(\mathbb{Z}_p)$).

7. The order of $SL_2(\mathbb{Z}_3)$ is 24, which might lead one to suspect it's isomorphic to S_4. Prove that it isn't by comparing centers (see ex 7 of chap 7).

Chapter 11

The sign of a permutation

Theorem 11.1. *Suppose $n \geq 2$.*

(a) Every permutation in S_n is a product of transpositions.

(b) If the identity $I = \tau_1 \ldots \tau_r$ in S_n is expressed as product of transpositions, r must be even.

Before giving the proof, we need the following lemmas.

Lemma 11.2. *Suppose $a, b, c, d \in \{1, \ldots, n\}$ are mutually distinct elements. We have the following identities among transpositions*

$$(ab) = (ba)$$

$$(ab)(ab) = I \tag{11.1}$$

$$(ac)(ab) = (ab)(bc) \tag{11.2}$$

$$(bc)(ab) = (ac)(cb) \tag{11.3}$$

$$(cd)(ab) = (ab)(cd) \tag{11.4}$$

Proof. The first couple are obvious, the rest will be left as an exercise. □

Lemma 11.3. *Any product of transpositions $\tau_1\tau_2 \ldots \tau_r$, in S_n, is equal to another product of transpositions $\tau_1' \ldots \tau_{r'}'$, such that r and r' have the same parity (in other words, they are either both even or both odd) and n occurs at most once among the τ_i'.*

Proof. Rather than giving a formal proof, we explain the strategy. Use (11.3) and (11.4) to move transpositions containing n next to each other. Then apply (11.1) and (11.2) to eliminate one of the n's. In each of these moves, either r stays the same or drops by 2. Now repeat.

Here are a couple of examples when $n = 4$,

$$(43)(41)(24) = (41)(13)(24) = (41)(24)(13) = (24)(12)(13)$$

$$(34)(12)(34) = (34)(34)(12) = (12)$$

□

Proof of theorem 11.1. We prove both statements by induction on n. The base case $n = 2$ of (a) is clear, the only permutations are (12) and $(12)(12)$. Now suppose that (a) holds for S_n. Let $f \in S_{n+1}$. If $f(n+1) = n+1$, then $f \in \mathrm{Stab}(n+1)$ which can be identified with S_n. So by induction, f is a product of transpositions. Now suppose that $j = f(n+1) \neq n+1$. Then the product $g = (n+1\, j)f$ sends $n+1$ to $n+1$. This implies that g is a product of transpositions $\tau_1\tau_2\ldots$ by the previous case. Therefore $f = (n+1\, j)\tau_1\tau_2\ldots$.

Statement (b) holds when $n = 2$, because $I = (12)^r$ if and only if r is even. Suppose that (b) holds for S_n. Let

$$I = \tau_1\tau_2\ldots\tau_r \tag{11.5}$$

in S_{n+1}. By using these lemma 11.3, we can get a new equation

$$I = \tau'_1\ldots\tau'_{r'} \tag{11.6}$$

where at most one of the τ'_i's contains $n+1$, and r' has the same parity as r. If exactly one of the τ'_i's contains $n+1$, then $\tau'_1\ldots\tau'_{r'}$ will send $n+1$ to a number other than $n+1$. This can't be the identity contradicting (11.6). Therefore none of the τ'_i's contains $n+1$. This means that (11.6) can be viewed as an equation in S_n. So by induction, we can conclude that r' is even. □

Corollary 11.4. *If a permutation σ is expressible as a product of an even (respectively odd) number of transpositions, then* **any** *decomposition of σ as a product of transpositions has an even (respectively odd) number of transpositions.*

Proof. Write

$$\sigma = \tau_1\ldots\tau_r = \tau'_1\ldots\tau'_{r'}$$

where τ_i, τ'_j are transpositions. Therefore

$$I = \tau_r^{-1}\ldots\tau_1^{-1}\tau'_1\ldots\tau'_{r'} = \tau_r\ldots\tau_1\tau'_1\ldots\tau'_{r'}$$

which implies that $r + r'$ is even. This is possible only if r and r' have the same parity. □

Definition 11.5. *A permutation is called even (respectively odd) if it is a product of an even (respectively odd) number of transpositions. Define*

$$\mathrm{sign}(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$

Lemma 11.6. *The map* $\mathrm{sign} : S_n \to \{1, -1\}$ *is a homomorphism.*

Proof. Clearly $\mathrm{sign}(I) = 1$ and $\mathrm{sign}(\sigma\tau) = \mathrm{sign}(\sigma)\,\mathrm{sign}(\tau)$. □

Definition 11.7. *The alternating group $A_n \subset S_n$ is the subgroup of even permutations.*

Observe that A_n is a subgroup, and in fact a normal subgroup, because it equals ker(sign). We can identify S_n/A_n with $\{1, -1\}$. Therefore

Lemma 11.8. $|A_n| = \frac{1}{2}n!$.

Earlier as an exercise, we found that the symmetry group of the dodecahedron had order 60, which is coincidentally the order of A_5. A more precise analysis, which we omit, shows that these groups are in fact isomorphic.

Let us apply these ideas to study functions of several variables. A function $f : X^n \to \mathbb{R}$ is called *symmetric* if

$$f(x_1, \ldots, x_i, \ldots, x_j, \ldots x_n) = f(x_1, \ldots, x_j, \ldots, x_i \ldots x_n)$$

and *antisymmetric* if

$$f(x_1, \ldots, x_i, \ldots, x_j, \ldots x_n) = -f(x_1, \ldots, x_j, \ldots, x_i, \ldots x_n)$$

for all $i \neq j$. For example, when $X = \mathbb{R}$

$$x_1 + x_2 + x_3$$

is symmetric, and

$$(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$$

is antisymmetric. Clearly when f is antisymmetric,

$$f(x_1, \ldots, x_n) = \text{sign}(\sigma) f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

holds for any permutation. A similar equation holds for symmetric functions, with sign(σ) omitted. We define the symmetrization and antisymmetrization operators by

$$\text{Sym}(f) = \frac{1}{n!} \sum_{\sigma \in S_n} f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

$$\text{Asym}(f) = \frac{1}{n!} \sum_{\sigma \in S_n} \text{sign}(\sigma) f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

We'll see in the exercises that these operators produce (anti)symmetric functions.

11.9 Exercises

1. Check the identities in lemma 11.2.

2. Prove that if $\sigma \in S_n$ is odd, then so is σ^{-1}.

3. Prove that a cycle of length r is even if and only if r is odd.

4. Prove that if $G \subseteq S_n$ is a subgroup of odd order, then $G \subseteq A_n$.

5. Prove that $\mathrm{Sym}(f)$ (respectively $\mathrm{Asym}(f)$) is symmetric (respectively antisymmetric), and furthermore that $f = \mathrm{Sym}(f)$ ($f = \mathrm{Asym}(f)$) if and only if f symmetric (antisymmetric).

6. If f is symmetric, prove that $\mathrm{Asym}(f) = 0$.

7. Prove that
$$f(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$
holds for all $\sigma \in A_n$ if and only if f is a sum of a symmetric and antisymmetric function.

Chapter 12

Determinants

The ideas of the previous chapter can be applied to linear algebra. Given an $n \times n$ matrix $A = [a_{ij}]$ over a field K, the *determinant*

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}$$

This is bit like the antisymmetrization considered earlier. There is also symmetric version, without $\operatorname{sign}(\sigma)$, called the permanent. However, as far as I know, it is much less useful. The definition, we gave for the determinant, is not very practical. However, it is theoretically quite useful.

Theorem 12.1. *Given an $n \times n$ matrix A, the following properties hold.*

(a) $\det I = 1$

(b) If B is obtained by multiplying the ith row of A by b then $\det B = b \det A$

(c) Suppose that the ith row of C is the sum of the ith rows of A an B, and all other rows of A, B and C are identical. Then $\det C = \det A + \det B$.

(d) $\det A = \det A^T$.

(e) Let us write $A = [v_1, \dots, v_n]$, where $v_1, v_2, \dots$ are the columns. Then $\det(v_{\tau(1)}, \dots v_{\tau(n)}) = \operatorname{sign}(\tau) \det(v_1, \dots v_n)$

Proof. Item (a) is clear because all the terms $\delta_{1\sigma(1)} \dots \delta_{n\sigma(n)} = 0$ unless $\sigma = I$.

(b)

$$\begin{aligned} \det B &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots (b a_{i\sigma(i)}) \dots a_{n\sigma(n)} \\ &= b \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{i\sigma(i)} \dots a_{n\sigma(n)} \\ &= b \det A \end{aligned}$$

(c) Denote the ith columns of A, B, C by $[\alpha_1, \ldots], [\beta_1, \ldots]$ and $[\alpha_1 + \beta_1, \ldots]$

$$\begin{aligned}\det C &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \ldots (\alpha_{\sigma(i)} + \beta_{\sigma(i)}) \ldots a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \ldots \alpha_{\sigma(i)} \ldots a_{n\sigma(n)} \\ &+ \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \ldots \beta_{\sigma(i)} \ldots a_{n\sigma(n)} \\ &= \det A + \det B\end{aligned}$$

Before proving (d), observe that by the commutative law

$$a_{1\sigma(1)} \ldots a_{n\sigma(n)} = a_{\tau(1)\tau\sigma(1)} \ldots a_{\tau(n)\tau\sigma(n)}$$

for any permutation τ. In particular, setting $\tau = \sigma^{-1}$ gives

$$a_{1\sigma(1)} \ldots a_{n\sigma(n)} = a_{\sigma^{-1}(1)1} \ldots a_{\sigma^{-1}(n)n}$$

Therefore

$$\begin{aligned}\det A^T &= \sum_{\tau \in S_n} \operatorname{sign}(\tau) a_{\tau(1)1} \ldots a_{\tau(n)n} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma^{-1}) a_{\sigma^{-1}(1)1} \ldots a_{\sigma^{-1}(n)n} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma^{-1}(1)1} \ldots a_{\sigma^{-1}(n)n} \\ &= \det A\end{aligned}$$

For (e), using the fact that $\operatorname{sign}(\sigma\tau) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau)$, we obtain $\operatorname{sign}(\sigma\tau)\operatorname{sign}(\tau) = \operatorname{sign}(\sigma)$. Therefore

$$\begin{aligned}\det(v_{\tau(1)}, \ldots, v_{\tau(n)}) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma\tau(1)} \ldots a_{n\sigma\tau(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma\tau)\operatorname{sign}(\tau) a_{1\sigma\tau(1)} \ldots a_{n\sigma\tau(n)} \\ &= \operatorname{sign}(\tau) \sum_{\sigma \in S_n} \operatorname{sign}(\sigma\tau) a_{1\sigma\tau(1)} \ldots a_{n\sigma\tau(n)}\end{aligned}$$

In the last sum, we can set $\eta = \sigma\tau$, and sum over $\eta \in S_n$. This allows us to simplify it to $\operatorname{sign}(\tau)\det(v_1, \ldots v_n)$.

□

Corollary 12.2. *A matrix A with a row of zeros has* $\det A = 0$

Proof. By (b), $\det A = 0 \cdot \det A$. □

In the exercises, we will use the above theorem to show that A behaves in the expected way under elementary row operations. This can be summarized as

Lemma 12.3. *If A, E are both $n \times n$ with E elementary, then $\det(E) \neq 0$ and $\det(EA) = \det(E)\det(A)$.*

Much of the importance of determines stems from the following facts.

Theorem 12.4. *Given an $n \times n$ matrix A, the following statements are equivalent*

(a) *A is invertible.*

(b) $\det A \neq 0$.

(c) *$Av = 0$ implies $v = 0$.*

Proof. By theorem 10.1, a square matrix A is either a product of elementary matrices when its invertible, or a product of elementary matrices and a matrix B with a zero row otherwise. In the first case $\det A \neq 0$ by lemma 12.3. In the second case, $\det A$ is proportional to $\det B = 0$. This proves the equivalence of (a) and (b).

If A is invertible and $Av = 0$, then $v = A^{-1}Av = 0$. Suppose A is not invertible. Then $\det A^T = \det A = 0$ by what was just proved. Therefore $A^T = FB$ where F is a product of elementary matrices, and B hasa row of zeros. For simplicity, suppose that the first row is zero. Set $v = (F^T)^{-1}[1, 0, \dots, 0]^T$. This vector is nonzero and $Av = B^T F^T v = 0$. This proves the equivalence of (a) and (c). □

Theorem 12.5. *The determinant gives a homomorphism* $\det : GL_n(K) \to K^*$.

Proof. If A and B are invertible $n \times n$ matrices, write A as a product of elementary matrices, then $\det AB = \det A \det B$ follows from lemma 12.3. □

Let A be an $n \times n$ matrix. An element $\lambda \in K$ is an *eigenvalue* of A if there exists a nonzero vector $v \in K^n$, called an *eigenvector*, such that

$$Av = \lambda v$$

or equivalently

$$(\lambda I - A)v = 0$$

Theorem 12.6. *The expression $p(x) = \det(xI - A)$ is a polynomial of degree n, called the characteristic polynomial of A. λ is an eigenvalue if and only if it is a root of $p(x)$.*

Proof. Clearly

$$p(x) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)(x\delta_{1\sigma(1)} - a_{1\sigma(1)}) \dots (x\delta_{n\sigma(n)} - a_{n\sigma(n)})$$

is a polynomial of degree n. From the definition of $p(x)$, $p(\lambda) = 0$ if and only if $\lambda I - A$ is not invertible. By theorem 12.4, this is equivalent to $\lambda I - A$ having a nonzero kernel, or in other words for λ to be an eigenvalue. □

Corollary 12.7. *A has at most n distinct eigenvalues.*

12.8 Exercises

1. Let E be obtained from I by interchanging two rows. Check that $\det E = -1$ and $\det(EA) = -\det A$.

2. Let E be obtained from I by multiplying a row by $k \in K^*$. Check that $\det E = k$ and $\det(EA) = k \det A$.

3. Let E be obtained from I by adding a multiple of one row to another. Check that $\det E = 1$ and $\det(EA) = \det A$.

4. Suppose that a square matrix A can be subdivided into blocks

$$\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

as indicated with B and D square. Prove that $\det A = \det B \det D$.

5. The permanent of an $n \times n$ matrix A is

$$\operatorname{Perm} A = \sum_{\sigma \in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)}$$

Prove that $\operatorname{Perm} A = \operatorname{Perm} A^T$.

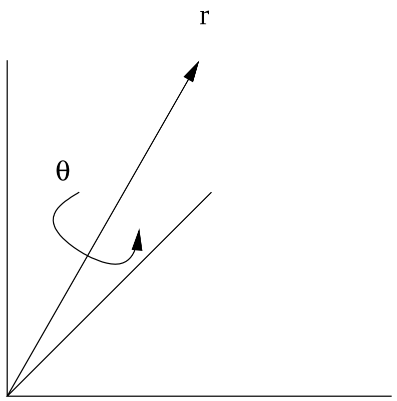
6. Calculate the cardinality of $\{A \in GL_n(\mathbb{Z}_p) \mid \det A = 0\}$.

7. Suppose that A is an $n \times n$ matrix with n eigenvalues $\lambda_1 \dots, \lambda_n$ allowing repetitions, i.e. the characteristic polynomial factors as $(x - \lambda_1) \dots (x - \lambda_n)$. Show that $\det A = (-1)^n \lambda_1 \dots \lambda_n$ and $\operatorname{trace} A = \lambda_1 + \dots + \lambda_n$, where $\operatorname{trace} A$ is defined as $a_{11} + a_{22} + \dots a_{nn}$.

Chapter 13

The 3 dimensional rotation group

A rotation in space is a transformation $R : \mathbb{R}^3 \to \mathbb{R}^3$ determined by a unit vector $r \in \mathbb{R}^3$ and an angle $\theta \in R$ as indicated in the picture below.



A bit more precisely, the transformation $R = R(\theta, r)$ has the line through r as the axis, and the plane perpendicular to the line is rotated by the angle θ in the direction given by the right hand rule (the direction that the fingers of right hand point if the thumb points in the direction of r). R is a linear transformation, so it is represented by a matrix that we denote by the same symbol. It is invertible with inverse $R(-\theta, r)$. Therefore the set of rotations is a subset of $GL_3(\mathbb{R})$. We will show that it is a subgroup, and in particular that the product of two rotations is again a rotation. This is fairly obvious if the rotations share the same axis, but far from obvious in general. The trick is characterize the matrices that arise from rotations. Recall that a 3×3 matrix A is orthogonal if its columns are *orthonormal*, i.e. they unit vectors such that the dot product of any two is zero. This is equivalent to $A^T A = I$.

Lemma 13.1. *If A is orthogonal,* $\det A = \pm 1$.

Proof. From $A^T A = I$ we obtain $\det(A)^2 = \det(A^T)\det(A) = 1$. $\square$

We already saw in the exercises to chapter 3 that the set of orthogonal matrices $O(3)$ forms a subgroup of $GL_3(\mathbb{R})$. Let $SO(3) = \{A \in O(3) \mid \det A = 1\}$.

Lemma 13.2. *$SO(3)$ is a subgroup of $O(3)$.*

Proof. If $A, B \in SO(3)$, then $\det(AB) = \det(A)\det(B) = 1$ and $\det(A^{-1}) = 1^{-1} = 1$. Also $\det(I) = 1$. $\square$

Proposition 13.3. *Every rotation matrix lies in $SO(3)$.*

Proof. Given a unit vector $v_3 = r$ as above, fix $R = R(\theta, r)$. By Gram-Schmid we can find two more vectors, so v_1, v_2, v_3 is orthonormal. Therefore $A = [v_1 v_2 v_3]$ is an orthogonal matrix. After possibly switching v_1, v_2, we can assume that v_1, v_2, v_3 is right handed or equivalently that $\det A = 1$. Then

$$\begin{aligned} R(v_1) &= \cos\theta v_1 + \sin\theta v_2 \\ R(v_2) &= -\sin\theta v_1 + \cos\theta v_2 \\ R(v_3) &= v_3 \end{aligned}$$

and therefore

$$RA = AM$$

where

$$M = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Since $M, A \in SO(3)$, it follows that $R = AMA^{-1} \in SO(3)$. $\square$

In principle, the method of proof can be used to calculate $R(\theta, [a, b, c]^T)$ explicitly. In fact, I did find an expression with the help of a computer algebra package:

$$\begin{bmatrix} a^2 + \cos(\theta) - a^2\cos(\theta) & -c\sin(\theta) + ab - ab\cos(\theta) & ac - ac\cos(\theta) + b\sin(\theta) \\ ab - ab\cos(\theta) + c\sin(\theta) & b^2 + \cos(\theta) - b^2\cos(\theta) & -a\sin(\theta) + bc - bc\cos(\theta) \\ -b\sin(\theta) + ac - ac\cos(\theta) & bc - bc\cos(\theta) + a\sin(\theta) & -b^2 + b^2\cos(\theta) - a^2 + a^2\cos(\theta) + 1 \end{bmatrix}$$

However, the formula is pretty horrendous and essentially useless. We will see a better way to do calculations shortly (which is in fact what I used to calculate the previous matrix).

We want to prove that every matrix in $SO(3)$ is a rotation. We start by studying their eigenvalues. In general, a real matrix need not have any real eigenvalues. However, this will not be a problem in our case.

Lemma 13.4. *A 3×3 real matrix has a real eigenvalue.*

Proof. The characteristic polynomial $p(\lambda) = \lambda^3 + a_2\lambda^2 + \ldots$ has real coefficients. Since λ^3 grows faster than the other terms, $p(\lambda) > 0$ when $\lambda \gg 0$, and $p(\lambda) < 0$ when $\lambda \ll 0$. Therefore the graph of $y = p(x)$ must cross the x-axis somewhere, and this would give a real root of p. (This intuitive argument is justified by the intermediate value theorem from analysis.)

□

Lemma 13.5. *If $A \in O(3)$, 1 or -1 is an eigenvalue.*

Proof. By the previous lemma, there exists a nonzero vector $v = [x, y, z]^T \in \mathbb{R}^3$ and real number λ such that $Av = \lambda v$. Since a multiple of v will satisfy the same conditions, we can assume that the square of the length $v^T v = x^2 + y^2 + z^2 = 1$. It follows that

$$\lambda^2 = (\lambda v)^T(\lambda v) = (Av)^T(Av) = v^T A^T A v = v^T v = 1$$

□

Theorem 13.6. *A matrix in $SO(3)$ is a rotation.*

Proof. Let $R \in SO(3)$. By the previous lemma, ± 1 is an eigenvalue.

We divide the proof into two cases. First suppose that 1 is eigenvalue. Let v_3 be an eigenvector with eigenvalue 1. We can assume that v_3 is a unit vector. We can complete this to an orthonormal set v_1, v_2, v_3. The vectors v_1 and v_2 form a basis for the plane $v_3^\perp$ perpendicular to v_3. The matrix $A = [v_1, v_2, v_3]$ is orthogonal, and we can assume that it is in $SO(3)$ by switching v_1 and v_2 if necessary. It follows that

$$RA = [Rv_1, Rv_2, Rv_3] = [Rv_1, Rv_2, v_3]$$

remains orthogonal. Therefore Rv_1, Rv_2 lie in $v_3^\perp$. Thus we can write

$$\begin{aligned} R(v_1) &= av_1 + bv_2 \\ R(v_2) &= cv_1 + dv_2 \\ R(v_3) &= v_3 \end{aligned}$$

The matrix

$$A^{-1}RA = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

lies in $SO(3)$. It follows that the block $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ lies in $SO(2)$, which means that it is a plane rotation matrix $R(\theta)$. It follows that $R = R(\theta, v_3)$.

Now suppose that -1 is an eigenvalue and let v_3 be an eigenvector. Defining A as above, we can see that

$$A^{-1}RA = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

This time the upper 2×2 is block lies $O(2)$ with determinant -1. This implies that it is a reflection. This means that there is a nonzero vector v in the plane $v_3^\perp$ such $Rv = v$. Therefore R also $+1$ as an eigenvalue, and we have already shown that R is a rotation. □

From the proof, we extract the following useful fact.

Corollary 13.7. *Every matrix in $SO(3)$ has $+1$ as an eigenvalue. If the matrix is not the identity then the corresponding eigenvector is the axis of rotation.*

We excluded the identity above, because everything would be an axis of rotation for it. Let us summarize everything we've proved in one statement.

Theorem 13.8. *The set of rotations in $\mathbb{R}^3$ can be identified with $SO(3)$, and this forms a group.*

13.9 Exercises

1. Check that unlike $SO(2)$, $SO(3)$ is not abelian. (This could get messy, so choose the matrices with care.)

2. Given two rotations $R_i = R(\theta_i, v_i)$, show that the axis of $R_2R_1R_2^{-1}$ is R_2v_1. Conclude that a normal subgroup of $SO(3)$, different from $\{I\}$, is infinite.

3. Check that
$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
has $1, e^{\pm i\theta}$ as complex eigenvalues. With the help of the previous exercise show that this holds for any rotation $R(\theta, v)$.

4. Show the map $f : O(2) \to SO(3)$ defined by
$$f(A) = \begin{bmatrix} A & 0 \\ 0 & \det(A) \end{bmatrix}$$
is a one to one homomorphism. Therefore we can view $O(2)$ as a subgroup of $SO(3)$. Show that this subgroup is the subgroup $\{g \in SO(3) \mid gr = \pm r\}$, where $r = [0, 0, 1]^T$.

5. Two subgroups $H_i \subseteq G$ of a group are *conjugate* if for some $g \in G$, $H_2 = gH_1g^{-1} := \{ghg^{-1} \mid h \in H_1\}$. Prove that $H_1 \cong H_2$ if they are conjugate. Is the converse true?

6. Prove that for any nonzero vector $v \in \mathbb{R}^3$, the subgroup $\{g \in SO(3) \mid gv = \pm v\}$ (respectively $\{g \in SO(3) \mid gv = v\}$) is conjugate, and therefore isomorphic, to $O(2)$ (respectively $SO(2)$). (Hint: use the previous exercises.)

Chapter 14

Finite subgroups of the rotation group

At this point, it should it should come as no surprise that finite subgroups of the $O(2)$ are groups of a symmetries of a regular polygon. We prove a slightly more precise statement.

Theorem 14.1. *A finite subgroup of $SO(2)$ is cyclic, and a finite subgroup of $O(2)$ not contained in $SO(2)$ is dihedral.*

Recall that the dihedral group D_n is defined by generators and relations $R^n = I, F^2 = I$ and $FRF = R^{-1}$. We include the "degenerate" cases $D_1 \cong \mathbb{Z}_2$, and $D_2 \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ (see exercises).

Proof. First suppose that $G \subset SO(2)$ is finite. The elements of G are of course rotations through some angle $\theta \in [0, 2\pi)$. Let $R \in G - \{I\}$ be the rotation with the smallest possible θ. Let $S \in G - \{I\}$ be another element with angle ϕ. Since $\phi \geq \theta$, we can write $\phi = n\theta + \psi$, where $n \geq 0$ is an integer and $\psi \geq 0$. By choosing n as large as large as possible, we can assume that $\psi < \theta$. Since ψ is the angle of SR^{-n}, we must have $\psi = 0$. This proves that $S = R^n$. So G is generated by R, and therefore cyclic.

Now suppose that $G \subset O(2)$ is finite but not contained in $SO(2)$. Then there exists $F \in G$ with $\det F = -1$. This is necessarily a reflection so that $F^2 = I$. $G \cap SO(2)$ is cyclic with generator R by the previous paragraph. Let us suppose that R has order n. We have that $\det FR = -1$, so it is also a reflection. This means that $FRFR = I$ or $FRF = R^{-1}$. Together with the relations $F^2 = I$ and R^n, we see that $G \cong D_n$. □

Let us now turn to finite subgroups of $SO(3)$. Since $O(2) \subset SO(3)$, we have the above examples. We also have symmetry groups of a regular tetrahedron, cube or dodecahedron. Remarkably, the converse is also true. We will be content to prove a weaker statement.

Theorem 14.2. *Let $G \subset SO(3)$ be a finite subgroup. Then either G is cyclic, dihedral or else it has order $12, 24$ or 60.*

The proof will be broken down into a series of lemmas. Let us suppose that $G \subset SO(3)$ is a nontrivial finite subgroup. Then G acts on the sphere S of radius one centered at the origin. We define a point of S to be a *pole* of G if it is fixed by at least one $g \in G$ with $g \neq I$. Let P be the set of poles. For $g \neq I$, there are exactly two poles $\pm p$, where the axis of g meets S. It follows that P is a finite set with even cardinality. We will see in an exercise that G acts on P. So, we can partition P into a finite number, say n, of orbits. Choose one point p_i, in each orbit.

Lemma 14.3.

$$2\left(1-\frac{1}{|G|}\right) = \sum_{i=1}^{n}\left(1-\frac{1}{|\operatorname{Stab}(p_i)|}\right) \tag{14.1}$$

Proof. By Burnside's formula

$$n = \frac{1}{|G|}\sum_{g\in G}|\operatorname{Fix}(g)|$$

As noted above $|\operatorname{Fix}(g)| = 2$, when $g \neq I$. Therefore, with the help of the orbit-stabilizer theorem

$$\begin{aligned} n &= \frac{1}{|G|}\left(2(|G|-1)+|P|\right)\\ &= \frac{1}{|G|}\left(2(|G|-1)+\sum_{1}^{n}|\operatorname{Orb}(p_i)|\right)\\ &= \frac{1}{|G|}\left(2(|G|-1)+\sum_{1}^{n}\frac{|G|}{|\operatorname{Stab}(p_i)|}\right)\end{aligned}$$

This can be rearranged to get

$$2\left(1-\frac{1}{|G|}\right) = \sum_{1}^{n}\left(1-\frac{1}{|\operatorname{Stab}(p_i)|}\right)$$

□

Lemma 14.4. *With above notation, if $G \neq \{I\}$ then either $n = 2$ or 3 in* (14.1).

Proof. Since $|G| \geq 2$ and $|\operatorname{Stab}(p_i)| \geq 2$, we must have

$$1 \leq 2\left(1-\frac{1}{|G|}\right) < 2$$

and

$$\frac{n}{2} \leq \sum\left(1-\frac{1}{|\operatorname{Stab}(p_i)|}\right) < n$$

The only way for (14.1) to hold is for $n = 2, 3$. □

Lemma 14.5. *If $n = 2$, G is cyclic.*

Proof. Since $\mathrm{Stab}(p_i) \subseteq G$, we have

$$\left(1 - \frac{1}{|\mathrm{Stab}(p_i)|}\right) \le \left(1 - \frac{1}{|G|}\right) \tag{14.2}$$

But (14.1) implies

$$2\left(1 - \frac{1}{|G|}\right) = \left(1 - \frac{1}{|\mathrm{Stab}(p_1)|}\right) + \left(1 - \frac{1}{|\mathrm{Stab}(p_2|}\right)$$

and this forces equality in (14.2) for both $i = 1, 2$. This implies that $G = \mathrm{Stab}(p_1) = \mathrm{Stab}(p_2)$. This means that $g \in G$ is a rotation with axis the line L connecting p_1 to 0 (or p_2 to 0, which would have to be the same). It follows that g would have to be a rotation in the plane perpendicular to L. So that G can be viewed as subgroup of $SO(2)$. Therefore it is cyclic by theorem 14.1. □

We now turn to the case $n = 3$. Let us set $n_i = |\mathrm{Stab}(p_i)|$ and arrange them in order $2 \le n_1 \le n_2 \le n_3$. (14.1) becomes

$$2\left(1 - \frac{1}{|G|}\right) = \left(1 - \frac{1}{n_1}\right) + \left(1 - \frac{1}{n_2}\right) + \left(1 - \frac{1}{n_3}\right)$$

or

$$1 + \frac{2}{|G|} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3}$$

The left side is greater than one, so we have a natural constraint.

Lemma 14.6. *The only integer solutions to the inequalities*

$$2 \le n_1 \le n_2 \le n_3$$

$$\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} > 1$$

are as listed together with the corresponding orders of G.

(a) $(2, 2, n_3)$ *and* $|G| = 2n_3$.

(b) $(2, 3, 3)$ *and* $|G| = 12$.

(c) $(2, 3, 4)$ *and* $|G| = 24$.

(d) $(2, 3, 5)$ *and* $|G| = 60$.

To complete the proof of theorem 14.2, we need the following

Lemma 14.7. *A subgroup $G \subset SO(3)$ corresponding to the triple $(2, 2, n)$ is isomorphic to D_n.*

Proof. We will deal with $n = 2$ in the exercises, so let us assume that $n > 2$. Let $H = \text{Stab}(p_3)$. This has order n. Note that $\text{Stab}(-p_3) = H$. We must have an element $F \in G$ which takes p_3 to $-p_3$, because otherwise we would have two orbits with stabilizers of order $n > 2$ contradicting our assumptions. Let $K \subseteq G$ be the subgroup generated by F and the elements of H. Let $k = |K|$. Then $k = qn$ with $q > 1$ because $H \subsetneqq K$. This forces $k = 2n$. Therefore $G = K$. This implies $G \subset \{g \in SO(3) \mid gp_3 = \pm p_3\} \cong O(2)$ (by a previous exercise). Theorem 14.1 implies that G is dihedral.

□

14.8 Exercises

1. Let D_2 be generated by F, R with relations $F^2 = R^2 = I$ and $FRF = R^{-1}$. Prove that this is abelian, and that the map $f : \mathbb{Z}_2 \times \mathbb{Z}_2 \to D_2$ given by $f(1, 0) = F$ and $f(0, 1) = R$ gives an isomorphism. This group is usually called the Klein four group.

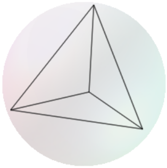
2. Suppose that (G, e) is a group of order 4 such that every element satisfies $g^2 = e$. Prove that $D_2 \cong G$.

3. Let $G \subset SO(3)$ be a finite group and P be the set of poles. Show that if $p \in P$, and $g \in G$, then $gp \in P$.

4. Prove lemma 14.6.

5. Let $G \subset SO(3)$ be a subgroup corresponding to the triple $(2, 2, 2)$ in the sense of lemma 14.6. Prove that $G \cong D_2$.

6. Consider a regular tetrahedron inscribed in the unit sphere S.



Show that the set of poles P of the symmetry group T of the tetrahedron consists of the vertices, midpoints of edges and midpoints of faces extended to S. Show that the T action on P has three orbits, where one of them has a stabilizer of order 2 and the remaining two have stabilizers of order 3.

7. Determine the poles of the symmetry group of the cube, and determine the orbits and stabilizers as in the previous exercise.

Chapter 15

Quaternions

The two dimensional rotation group can be naturally identified with the multiplicative group of complex numbers with $|z| = 1$. This idea can be extended to handle rotations in $\mathbb{R}^3$, and this will be explained in the next chapter. We start by describing the ring of quaternions, which was discovered by Hamilton in order to generalize complex numbers. The ring of quaternions is given by

$$\mathbb{H} = \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\}$$

where i, j, k are symbols. Alternatively, we can think of $a + bi + cj + dk$ as a more suggestive way of writing the vector $(a, b, c, d) \in \mathbb{R}^4$. We define

$$0 = 0 + 0i + 0j + 0k$$

$$1 = 1 + 0i + 0j + 0k$$

$$(a+bi+cj+dk)+(a'+b'i+c'j+d'k) = (a+a')+(b+b')i+(c+c')j+(d+d')k$$

$$\begin{aligned}(a + bi + cj + dk) \cdot (a' + b'i + c'j + d'k) = {} & (aa' - bb' - cc' - dd') + (ab' + ba' + cd' - dc')i \\ & + (ac' - bd' + ca' + db')j + (ad' + bc' - cb' + da')k\end{aligned}$$

To put it another way, multiplication is determined by the rules:

$$1 \text{ is the identity}$$

$$i^2 = j^2 = k^2 = -1$$

$$ij = -ji = k$$

$$jk = -kj = i$$

$$ki = -ik = j.$$

Theorem 15.1. *With the above rules, $\mathbb{H}$ becomes a noncommutative ring.*

Proof. All the laws, except the associative law for multiplication, are not difficult to verify. In principle, associativity can be checked by a long and messy calculation. Instead, we will embed $\mathbb{H}$ into the ring $M_{22}(\mathbb{C})$ with the help of the Pauli spin matrices[1] used in physics:

$$\sigma_i = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, \sigma_j = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \sigma_k = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

The i within the matrices is the complex number $\sqrt{-1}$ of course. We define a map $f : \mathbb{H} \to M_{22}(\mathbb{C})$ by

$$\begin{aligned} f(a + bi + cj + dk) &= aI + b\sigma_i + c\sigma_j + d\sigma_k \\ &= \begin{bmatrix} a + di & bi - c \\ bi + c & a - di \end{bmatrix} \end{aligned} \tag{15.1}$$

which is clearly a homomorphism of additive groups. If

$$f(a + bi + cj + dk) = 0$$

then clearly $a = b = c = d = 0$ by (15.1). So f is one to one. A calculation shows that

$$\begin{aligned} \sigma_i^2 = \sigma_j^2 = \sigma_k^2 = -I \\ \sigma_i\sigma_j = -\sigma_j\sigma_i = \sigma_k \end{aligned} \tag{15.2}$$

etc. So that f takes a product of quaternions uv to the product of matrices $f(u)f(v)$. Associativity of products is now automatic. More explicitly, $u(vw) = (uv)w$, because $f(u(vw)) = f(u)f(v)f(w) = f((uv)w)$. □

We could have simply defined the set of quaternions to be the set of matrices of the form (15.1). But this would hide the fact that quaternions should be viewed as a generalization of complex numbers. Many familiar constructions from complex numbers carry over to $\mathbb{H}$. We define the conjugate, norm and real and imaginary parts of a quaternion by

$$\overline{a + bi + cj + dk} = a - bi - cj - dk$$

$$|a + bi + cj + dk| = \sqrt{a^2 + b^2 + c^2 + d^2}$$

$$Re(a + bi + cj + dk) = a$$

$$Im(a + bi + cj + dk) = bi + cj + dk$$

Let us say that a quaternion is imaginary if its real part is zero

Theorem 15.2. *Let $q \in \mathbb{H}$ then*

(a) $\overline{\overline{q}} = q$.

[1] Actually, we are using i times the Pauli matrices, which is more convenient for our purposes.

(b) $\overline{q_1 + q_2} = \overline{q_2} + \overline{q_1}$.

(c) $\overline{q_1 q_2} = \overline{q_2}\,\overline{q_1}$.

(d) $q\overline{q} = |q|^2$.

(e) $|q_1 q_2| = |q_1||q_2|$

(f) *If q is imaginary $q^2 = -|q|^2$.*

The first two statements are easy. The remainder are left as exercises.

Corollary 15.3. $\mathbb{H}$ *forms a division ring. If $q \neq 0$, its inverse*

$$q^{-1} = \frac{1}{|q|^2}\overline{q}$$

In particular, $\mathbb{H}^* = \mathbb{H} - \{0\}$ is a group under multiplication.

Lagrange proved that every positive integer is a sum of four squares of integers. For example,

$$10 = 3^2 + 1^2 + 0^2 + 0^2$$
$$20 = 4^2 + 2^2 + 1^2 + 0^2$$
$$30 = 5^2 + 2^2 + 1^2 + 0^2$$

Although we won't prove the theorem, we will explain one step because it gives a nice application of quaternions.

Proposition 15.4. *If x and y are both expressible as a sum of four squares of integers, then the same is true of xy.*

Proof. By assumption, we can find two quaternions q_1 and q_2 with integer coefficients such that $x = |q_1|^2$ and $y = |q_2|^2$. The product $q_1 q_2$ is also a quaternion with integer coefficients. By theorem 15.2, $xy = |q_1 q_2|^2$. □

15.5 Exercises

1. Check (15.2) and use this to show that in addition

$$\sigma_j \sigma_k = -\sigma_k \sigma_j = \sigma_i$$

$$\sigma_k \sigma_i = -\sigma_i \sigma_k = \sigma_j$$

hold.

2. Prove part (c) and (d) of theorem 15.2.

3. Prove part (e) and (f).

4. Check that the set $Q = \{1, -1, i, -i, j, -j, k, -k\}$ is a subgroup of $\mathbb{H}^*$ which is not abelian and not isomorphic to D_4. So it is a group of order 8 that we have not seen before, called the quaternion group.

5. Show that $\{\pm 1\} \subset Q$ is a normal subgroup, and that the quotient $Q/\{\pm 1\}$ is isomorphic to D_2.

6. Let
$$\tilde{T} = \{\pm 1, \pm i, \pm j, \pm k, \frac{1}{2}(\pm 1 \pm i \pm j \pm k)\}$$
where the signs on the terms in the last sum can be chosen independently of each other. Check that $\tilde{T} \subset \mathbb{H}^*$ is a subgroup of order 24. This is called the binary tetrahedral group.

7. You have probably encountered the scalar (or dot or inner) product $\langle , \rangle$ and vector (or cross) product $\times$ on $\mathbb{R}^3$ before. Identifying vectors $(a, b, c) \in \mathbb{R}^3$ with imaginary quaternions $ai + bj + ck$, the scalar product is an $\mathbb{R}$-valued operation given by
$$\langle bi + cj + dk, ei + fj + gk\rangle = be + cf + dg$$
The vector product is an $\mathbb{R}^3$-valued distributive operation satisfying
$$v \times w = -w \times v$$
$$i \times j = k,\ j \times k = i,\ k \times i = j.$$
If $v, w \in \mathbb{R}^3$, show that these are related quaternionic product by
$$v \cdot w = -\langle v, w\rangle + v \times w$$

Chapter 16

The Spin group

We return to the study of rotations. We saw earlier a rotation can be represented by a 3×3 matrix in $SO(3)$. However, as we saw this description is cumbersome for doing calculations. We will give an alternative based on the ring of quaternions which makes this easy. Define the spin group

$$\mathrm{Spin} = \{q \in \mathbb{H} \mid |q| = 1\}$$

Using theorem 15.2, we can see that this is a subgroup of $\mathbb{H}^*$, so it really is a group. The word "spin" comes from physics (as in electron spin); at least I think it does. Usually this group is called Spin(3), but we won't consider any of the other groups in this series.

Lemma 16.1. *If $q \in \mathrm{Spin}$ and $v \in \mathbb{H}$ is imaginary, then $qv\overline{q}$ is imaginary.*

Proof. $Re(v) = 0$ implies that $\overline{v} = -v$. Therefore

$$\overline{qv\overline{q}} = q\overline{v}\overline{q} = -qv\overline{q}$$

This implies $qv\overline{q}$ is imaginary. □

We will identify $\mathbb{R}^3$ with imaginary quaternions by sending $[x, y, z]$ to $xi + yj + zk$. The previous lemma allows us to define a transformation $\mathrm{Rot}(q) : \mathbb{R}^3 \to \mathbb{R}^3$ by $\mathrm{Rot}(q) = qv\overline{q}$ for $q \in \mathrm{Spin}$. This is a linear transformation, therefore it can be represented by a 3×3 matrix.

Lemma 16.2. $\mathrm{Rot} : \mathrm{Spin} \to GL_3(\mathbb{R})$ *is a homomorphism.*

Proof. We have that $\mathrm{Rot}(q_1q_2) = \mathrm{Rot}(q_1)\,\mathrm{Rot}(q_2)$ because $\mathrm{Rot}(q_1q_2)(v) = q_1q_2v\overline{q}_2\overline{q}_1 = \mathrm{Rot}(q_1)\,\mathrm{Rot}(q_2)(v)$. And the lemma follows.

□

Lemma 16.3. $\mathrm{Rot}(q)$ *is an orthogonal matrix.*

Proof. We use the standard characterization of orthogonal matrices that these are exactly the square matrices for which $|Av| = |v|$ for all vectors v. If $v \in \mathbb{R}^3$, $|\,\mathrm{Rot}(q)(v)|^2 = |qv\overline{q}|^2 = |q|^2|v|^2|\overline{q}|^2 = |v|^2$. □

Lemma 16.4. $\mathrm{Rot}(q) \in SO(3)$.

Proof. There are a number of ways to see this. Geometrically, an orthogonal matrix lies in $SO(3)$ if it takes a right handed orthonormal basis to another right handed basis. In terms of the vector cross products, right handed means that the cross product of the first vector with the second vector is the third. In the exercise 7 of the last chapter, we saw that the imaginary part of the product of two imaginary quaternions is the vector cross product of the corresponding vectors. The right handed basis i, j, k gets transformed to $\mathrm{Rot}(q)i, \mathrm{Rot}(q)j, \mathrm{Rot}(q)k$. Since $qi\overline{q}qj\overline{q} = qij\overline{q} = qk\overline{q}$, we have $\mathrm{Rot}(q)i \times \mathrm{Rot}(q)j = \mathrm{Rot}(q)k$. So this is again right handed. □

Lemma 16.5. *If r is an imaginary quaternion with $|r| = 1$, and $a, b \in \mathbb{R}$ satisfy $a^2 + b^2 = 1$, then* $\mathrm{Rot}(a + br)$ *is a rotation about r.*

Proof. Let $q = a + br$. It clearly satisfies $|q| = 1$. The lemma follows from

$$\mathrm{Rot}(q)(r) = (a + br)r(a - br) = (ar - b)(a - br) = r$$

□

It remains to determine the angle.

Theorem 16.6. *For any unit vector r viewed as an imaginary quaternion,*

$$\mathrm{Rot}(\cos(\theta) + \sin(\theta)r)$$

is $R(2\theta, r)$.

Proof. Pick a right handed system orthonormal vectors v_1, v_2, v_3 with $v_3 = r$. Then by exercise 7 of the last chapter, $v_1v_2 = v_3$, $v_2v_3 = v_1$, and $v_3v_1 = v_2$. Let $q = \cos(\theta) + \sin(\theta)r$. We have already seen that $\mathrm{Rot}(q)v_3 = v_3$. We also find

$$\begin{aligned} Rot(q)v_1 &= (\cos\theta + \sin\theta v_3)v_1(\cos\theta - \sin\theta v_3) \\ &= (\cos^2\theta - \sin^2\theta)v_1 + (2\sin\theta\cos\theta)v_2 \\ &= \cos(2\theta)v_1 + \sin(2\theta)v_2 \end{aligned}$$

and

$$\begin{aligned} Rot(q)v_2 &= (\cos\theta + \sin\theta v_3)v_2(\cos\theta - \sin\theta v_3) \\ &= -\sin(2\theta)v_1 + \cos(2\theta)v_2 \end{aligned}$$

which means that $\mathrm{Rot}(q)$ behaves like $R(2\theta, r)$. □

Corollary 16.7. *The homomorphism* $\mathrm{Rot} : \mathrm{Spin} \to SO(3)$ *is onto, and $SO(3)$ is isomorphic to* $\mathrm{Spin}/\{\pm 1\}$.

Proof. Any rotation is given by $R(2\theta, r)$ for some θ and r, so Rot is onto. The kernel of Rot consists of $\{1, -1\}$. Therefore $SO(3) \cong \mathrm{Spin}/\{\pm 1\}$. □

So in other words, a rotation can be represented by an element of Spin uniquely up to a plus or minus sign. This representation of rotations by quaternions is very economical, and, unlike $R(\theta, r)$, multiplication is straigthforward.

16.8 Exercises

1. Suppose we rotate $\mathbb{R}^3$ counterclockwise once around the z axis by $90°$, and then around the x axis by $90°$. This can expressed as a single rotation. Determine it.

2. Given a matrix $A \in M_{nn}(\mathbb{C})$. Define the adjoint $A^* = \bar{A}^T$. In other words the ijth entry of A^* is $\overline{a}_{ji}$. (This should not be confused with the matrix built out of cofactors which also often called the adjoint.) A matrix A called unitary if $A^*A = I$ and special unitary if in addition $\det A = 1$. Prove that the subset $U(n)$ (or $SU(n)$) of (special) unitary matrices in $GL_n(\mathbb{C})$ forms a subgroup.

3. Let $a + bi + cj + dk \in$ Spin, and let $A \in M_{22}(\mathbb{C})$ be given by (15.1). Prove that $A \in SU(2)$. Prove that this gives an isomorphism Spin $\cong SU(2)$.

4. Consider the quaternion group $Q = \{\pm 1, \pm i, \pm j, \pm k\}$ studied in a previous exercise. Show this lies in Spin and that its image in $SO(3)$ *is* the subgroup

$$\left\{ \begin{bmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm 1 \end{bmatrix} \mid \text{there are 0 or 2 } -1\text{'s} \right\}$$

   Find the poles (see chapter 14) and calculate the orders of their stabilizers.

5. Let

$$V = \{\frac{1}{\sqrt{3}}[1,1,1]^T, \frac{1}{\sqrt{3}}[-1,-1,1]^T, \frac{1}{\sqrt{3}}[-1,1,-1]^T, \frac{1}{\sqrt{3}}[1,-1,-1]^T\}$$

   and let

$$\tilde{T} = \{\pm 1, \pm i, \pm j, \pm k, \frac{1}{2}(\pm 1 \pm i \pm j \pm k)\}$$

   be the subgroup of Spin defined in an exercise in the previous chapter. Show that the image T of $\tilde{T}$ in $SO(3)$ has order 12, and that it consists of the union of the set of matrices in exercise 5 and

$$\{R(\theta, r) \mid \theta \in \{\frac{\pi}{6}, \frac{\pi}{3}\}, r \in V\}$$

6. Continuing the last exercise. Show that the T acts as the rotational symmetry group of the regular tetrahedron with vertices in V.

Bibliography

[1] M. Artin, Algebra

[2] M. Armstrong, Groups and symmetry

[3] H. Weyl, Symmetry