University of Melbourne
School of Mathematics and Statistics

MAST20022
Group Theory and Linear Algebra*

Semester 2, 2018

*The MAST20022 lectures for 2018 will not follow these notes exactly. They are presented as an extra resource to complement the lectures and other suggested material.

These lecture notes were compiled in the School of Mathematics and Statistics of the University of Melbourne for the use of students in the subject MAST20022.

Modified: 24.07.2017 by C. Haesemeyer, July 2018 by L. Reeves

Contents

1 Modular arithmetic and fields

We begin with some number theory, looking at divisibility properties of the *natural numbers*

$$\mathbb{N} = \{1, 2, 3, 4, \dots\}$$

and the *integers*

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

We then introduce a fundamental concept in number theory: the idea of modular arithmetic due to Gauss. This provides examples of the algebraic structures "groups" and "fields" that will be important throughout this course. Modular arithmetic also plays a key role in cryptography, used daily to provide secure transmission of information over the internet.

1.1 Well-ordering

A basic property of the natural numbers is the

Well-ordering property. *Every non-empty subset of $\mathbb{N}$ has a smallest element. In other words, if $A \subset \mathbb{N}$ and $A \neq \emptyset$, then there exists an element $s \in A$ such that $s \leq a$ for all $a \in A$.*

This also applies to subsets of $\mathbb{Z}$ that are bounded below. For example, any non-empty subset of $\{0, 1, 2, 3, 4, 5, \dots\}$ has a smallest element.

Exercise 1. Does well-ordering hold for:

(a) Arbitrary subsets of $\mathbb{Z}$?

(b) Subsets of $\mathbb{R}$ that are bounded below?

(c) Subsets of $\mathbb{Q}$ that are bounded below?

Remark 1.1.1. The well-ordering property of the natural numbers is equivalent to the Principle of Mathematical Induction. Giving a proper proof of this equivalence requires an axiomatic definition of $\mathbb{N}$, which would take us too far afield, but see [2, Theorem C.4].

Principle of mathematical induction. *Assume we are given statements $S(n)$ for all $n \in \mathbb{N}$ such that:*

base case: *the statement $S(1)$ is true, and*

induction step: *for all $n \in \mathbb{N}$, if $S(n)$ is true, then $S(n+1)$ is true.*

Then the statement $S(n)$ is true for all $n \in \mathbb{N}$.

1.2 Divisibility of integers

Adding, subtracting and multiplying integers give rise to integers. But when we divide an integer a by an integer d, we may get a remainder. We can always guarantee that the remainder is smaller than d.

Theorem 1.2.1 (Division with remainder)**.** *If a, d are integers and $d > 0$, there are unique integers q, r such that*

$$a = qd + r \qquad \text{and} \qquad 0 \leq r < d.$$

The number q is called the **quotient** *and r is the* **remainder** *when a is divided by d.*

Proof. Suppose $a > 0$. (The case $a \leq 0$ can be reduced to $a > 0$. Can you see how?)

Consider the set

$$S = \{a - qd \mid q \in \mathbb{Z} \text{ and } a - qd \geq 0\}.$$

Then S is clearly a subset of $\{0, 1, 2, 3, \ldots\}$. It is also non-empty, as $a \in S$: if $q = 0$ then $a - qd = a \geq 0$, so $a = a - qd \in S$. By the well-ordering property, S has a smallest element r. Then $r = a - qd$ for some $q \in \mathbb{Z}$ and we have $r \geq 0$. We claim that $r < d$. If $r \geq d$ then $r - d \geq 0$ and $r - d = a - (q + 1)d$ so $r - d \in S$. This contradicts the fact that r is the smallest element of S. So $0 \leq r < d$ as required. We leave the proof that q and r are unique as Exercise 6. □

Long division of a by d produces the quotient q and remainder r. Alternatively, we can calculate a/d and take $q = \lfloor a/d \rfloor$ to be the largest integer less than or equal to the result; then $r = a - qd$.

Definition 1.2.2. Let a, d be integers. We say that d ***divides*** a if there is an integer q such that $a = qd$. (In other words, the remainder is zero when we divide a by d.) We also say that d is a *factor* (or *divisor*) of a and that a is a *multiple* of d.

Notation 1.2.3. We write $d \mid a$ if d divides a, and $d \nmid a$ if d does not divide a.

Lemma 1.2.4. *Let a, b, c be integers. Then*

(1) If $a \mid b$ and $b \mid c$ then $a \mid c$.

(2) If $a \mid b$ and $a \mid c$ then $a \mid xb + yc$ for all integers x, y.

Proof. (1) If $a \mid b$ then $b = ka$ for some integer $k \in \mathbb{Z}$. If $b \mid c$ then $c = \ell b$ for some $\ell \in \mathbb{Z}$. Then $c = \ell b = \ell(ka) = (\ell k)a$, so $c \mid a$ since $\ell k \in \mathbb{Z}$. The proof of (2) is left as Exercise 4.

□

1.2.1 Exercises

2. Find the quotient and remainder when:

(a) 25 is divided by 3

(b) 68 is divided by 7

(c) -33 is divided by 7.

3. Show that if a, b, c, d are integers such that $a \mid b$ and $c \mid d$ then $ac \mid bd$.

4. Prove that if a, b, c, x, y are integers such that $a \mid b$ and $a \mid c$ then $a \mid xb + yc$. (Lemma 1.2.4 part (2).)

5. Prove that if a, b are positive integers such that $a \mid b$ and $b \mid a$ then $a = b$.

6. (Harder) Show that the remainder r and quotient q in Theorem 1.2.1 are *unique*.

1.3 The Euclidean Algorithm

Definition 1.3.1. The ***greatest common divisor*** of two integers a and b is the non-negative integer d satisfying:

(1) $d \mid a$ and $d \mid b$,

(2) if $c \mid a$ and $c \mid b$ then $c \mid d$.

We write $\gcd(a, b)$ for the greatest common divisor of a and b. If a and b are not both zero, then $\gcd(a, b)$ is the largest natural number that divides both a and b. We also have (by convention) the special case $\gcd(0, 0) = 0$.

Example 1.3.2. $\gcd(4, 6) = 2$, $\gcd(10, -20) = 10$, $\gcd(7, 3) = 1$, $\gcd(0, 5) = 5$.

Definition 1.3.3. Two integers a, b are ***relatively prime*** if $\gcd(a, b) = 1$.

Example 1.3.4. 12 and 35 are relatively prime, but 12 and 34 are not.

For small integers, we can find greatest common divisors easily from the definition: enumerate the sets of divisors of a and of b, and pick the largest number in the intersection. For large integers, there is a much more efficient method, based on the following observations:

Lemma 1.3.5. *Let a, b be integers. Then:*

(1) $\gcd(a, b) = \gcd(b, a) = \gcd(-a, b)$*,*

(2) $\gcd(a, 0) = a$*,*

(3) if $a = bq + r$ where q, r are integers then $\gcd(a, b) = \gcd(b, r)$*.*

Proof. We prove part (3), leaving (1) and (2) as an exercise. If $d \mid a$ and $d \mid b$ then $d \mid r = a - bq$ by Lemma 1.2.4. Similarly, if $d \mid b$ and $d \mid r$ then $d \mid a = bq + r$. It follows that the set of common divisors of a and b is the same as the set of common divisors of b and r. Hence $\gcd(a, b) = \gcd(b, r)$. □

This leads us to the following idea due to Euclid: If we want to compute $\gcd(a, b)$, we may assume that $a \geq b > 0$. Then we can divide a by b, giving $a = bq + r$ with $0 \leq r < b$. Lemma 1.3.5 shows that we can replace a, b by *smaller* integers b, r without changing the gcd. We now repeat the argument, starting with b, r. At each stage, the remainder decreases in size, so eventually we obtain a remainder of 0. The last non-zero remainder is the desired gcd.

Euclidean algorithm.

```
def gcd(a, b):
    """
    INPUT:  integers a and b with a > b > 0
    OUTPUT: gcd(a, b)
    """
    while True:
        q, r = a.quo_rem(b)     # get q, r from a = qb + r
        if r == 0:
            return b
        else:
            a = b
            b = r
```

An instance of this algorithm looks as follows:

$$\begin{array}{rccr}
a = & q_0 \times b & +r_1 & 0 < r_1 < b \\
b = & q_1 \times r_1 & +r_2 & 0 < r_2 < r_1 \\
r_1 = & q_2 \times r_2 & +r_3 & 0 < r_3 < r_2 \\
r_2 = & q_3 \times r_3 & +r_4 & 0 < r_4 < r_3 \\
\vdots & & & \\
r_{n-2} = & q_{n-1} \times r_{n-1} & +r_n & 0 < r_n < r_{n-1} \\
r_{n-1} = & q_n \times r_n & +0 &
\end{array}$$

Then the greatest common divisor is the last non-zero remainder:

$$\gcd(a, b) = r_n.$$

Example 1.3.6. Find $\gcd(4163, 8869)$.

Theorem 1.3.7. *The Euclidean algorithm works as intended. (In other words, it terminates after finitely many steps and returns the gcd.)*

Proof. At each step, the remainder r is strictly smaller than the previous remainder. So each step reduces the remainder by at least one. Since the first remainder is at most $b-1$, after at most b steps the remainder must reach zero[1].

To see that the return value is indeed the gcd, we use Lemma 1.3.5, which shows:

$$\gcd(a,b) = \gcd(b,r_1) = \gcd(r_1,r_2) = \ldots = \gcd(r_n,0) = r_n.$$

□

An important consequence of the Euclidean algorithm is the following:

Theorem 1.3.8. *Let d be the greatest common divisor of two integers a and b. Then there are integers x, y such that*

$$ax + by = d.$$

Proof. Exercise 10. □

Here is how x, y can be calculated systematically by working backwards through the steps of the Euclidean algorithm.

Example 1.3.9. Use the Euclidean algorithm to solve the equation $131x + 71y = 1$.

Theorem 1.3.8 has important consequences.

Theorem 1.3.10. *If $d \mid ab$ and $\gcd(a,d) = 1$ then $d \mid b$.*

Proof. Since $\gcd(a,d) = 1$, we can find $x, y \in \mathbb{Z}$ such that $ax + dy = 1$. Multiplying both sides by b gives $(ab)x + (bd)y = b$. But $d \mid ab$ by assumption and $d \mid bd$, so $d \mid b = (ab)x + (bd)y$ by Lemma 1.2.4(2). □

Definition 1.3.11. A positive integer p is *prime* if it has exactly two positive divisors (which then have to be 1 and p with $p \neq 1$).

Corollary 1.3.12. *If p is prime and $p \mid ab$, then $p \mid a$ or $p \mid b$.*

Proof. If $p \nmid a$ then $\gcd(a,p) = 1$, so $p \mid b$ by Theorem 1.3.10. □

This is the main fact needed to prove that each positive integer has a unique prime factorization:

Fundamental theorem of arithmetic. *Every integer $n \geq 2$ can be written uniquely as a product*

$$n = p_1^{e_1} p_2^{e_2} \ldots p_r^{e_r},$$

where $p_1 < p_2 < \ldots < p_r$ are prime numbers and $e_j \in \mathbb{N}$ for all j.

Proof. See [2, Theorem 1.11]. □

1.3.1 Exercises

7. Using Euclid's Algorithm (by hand) find:

(a) $\gcd(14,35)$ (b) $\gcd(105,165)$ (c) $\gcd(1287,1144)$

(d) $\gcd(1288,1144)$ (e) $\gcd(1287,1145)$

8. Find the greatest common divisor $d = \gcd(a,b)$ for the following pairs of numbers (a,b), and find integers x and y so that $d = xa + yb$.

(a) $(27,33)$, (b) $(27,32)$, (c) $(312,377)$.

[1] This is a very generous estimate on the number of steps necessary in the Euclidean algorithm. Much better estimates exist (search for *Euclidean algorithm* on Wikipedia).

9. (a) Give an example of natural numbers a, b, c such that $a \mid c$ and $b \mid c$ but $ab \nmid c$.

(b) Let $a, b, c \in \mathbb{Z}$ be integers with $\gcd(a, b) = 1$. Prove that if $a \mid c$ and $b \mid c$ then $ab \mid c$.

10. (Harder) Use the well-ordering property of the natural numbers to prove Theorem 1.3.8.
(Hint: Assume that a, b are not both zero, and consider the set

$$S = \{an + bm \mid n, m \in \mathbb{Z} \text{ and } an + bm > 0\}.$$

Show that this contains a smallest positive element d, and that $d = \gcd(a, b)$.)

1.4 Modular arithmetic

Next we introduce the idea of "arithmetic modulo m" where two numbers that differ by a multiple of m are regarded as being the same. A familiar example is "clock arithmetic": two periods of time, as displayed on a 12-hour clock, are indistinguishable if they differ by a multiple of 12.

Definition 1.4.1. Let m be a positive integer. Two integers a and b are ***congruent modulo*** m if m divides $a - b$. We then write

$$a \equiv b \pmod{m}.$$

Equivalently, $a \equiv b \pmod{m}$ exactly when a and b give the same remainder when divided by m.

Example 1.4.2. $3 \equiv 1 \pmod{2}$, $3 \equiv 17 \pmod{7}$, $3 \equiv -15 \pmod{9}$, $4 \equiv 0 \pmod{2}$.
But $6 \not\equiv 1 \pmod{4}$.

Lemma 1.4.3. *Congruence modulo m satisfies the properties:*

(reflexivity) $a \equiv a \pmod{m}$ *for all* a

(symmetry) *if* $a \equiv b \pmod{m}$ *then* $b \equiv a \pmod{m}$

(transitivity) *if* $a \equiv b \pmod{m}$ *and* $b \equiv c \pmod{m}$ *then* $a \equiv c \pmod{m}$

Proof. Exercise 15. □

Definition 1.4.4. A binary relation that satisfies the three properties listed in Lemma 1.4.3 is called an ***equivalence relation***[2].

Arithmetic modulo m is well-behaved:

Theorem 1.4.5. *If* $a \equiv c \pmod{m}$ *and* $b \equiv d \pmod{m}$ *then*

(1) $a + b \equiv c + d \pmod{m}$,

(2) $a - b \equiv c - d \pmod{m}$,

(3) $ab \equiv cd \pmod{m}$,

(4) $a^n \equiv c^n \pmod{m}$ *for every natural number* n.

Proof. We prove (1) and leave the rest as Exercise 16. Since $a \equiv c \pmod{m}$ and $b \equiv d \pmod{m}$, we have $a = c + km$ and $b = d + \ell m$ for some $k, \ell \in \mathbb{Z}$. Then

$$(a + b) - (c + d) = (c + km) + (d + \ell m) - (c + d) = (k + \ell)m$$

and $k + \ell \in \mathbb{Z}$, so $a + b \equiv c + d \pmod{m}$. □

This greatly simplifies calculations modulo m.

Example 1.4.6. To find 29^4 modulo 12, we do not need to compute 29^4. Instead, note that $29 \equiv 5 \pmod{12}$ and so $29^4 \equiv 5^4 \pmod{12}$. We could now calculate 5^4 but we can simplify even more. Note that $5^4 = (5^2)^2$ and $5^2 = 25 \equiv 1 \pmod{12}$. Thus $29^4 \equiv 5^4 \equiv 1^2 \equiv 1 \pmod{12}$.

[2] We know two equivalence relations so far: equality and congruence modulo m.

1.4.1 The integers modulo m

The properties of congruence described in the previous section suggest that we can form a new number system where we regard two numbers that differ by a multiple of m as being the same. The technical way to achieve this identification is to make a set that contains *all* numbers which differ from a fixed number by a multiple of m.

Definition 1.4.7. Let a and m be integers with $m > 0$. The ***congruence class of a modulo m*** is the set of all integers that are congruent to a modulo m. This is denoted $[a]_m$, thus

$$[a]_m = \{x \in \mathbb{Z} \mid x \equiv a \pmod{m}\}.$$

Definition 1.4.8. The set of all congruence classes modulo m is denoted $\mathbb{Z}/m\mathbb{Z}$ and is called ***the integers modulo m***.

Example 1.4.9. The set $\mathbb{Z}/3\mathbb{Z}$ of integers modulo 3 has 3 elements:

$$\begin{aligned}[0]_3 &= \{\dots, -9, -6, -3, 0, 3, 6, 9, 12, \dots\}\\ [1]_3 &= \{\dots, -8, -5, -2, 1, 4, 7, 10, 13, \dots\}\\ [2]_3 &= \{\dots, -7, -4, -1, 2, 5, 8, 11, 14, \dots\}.\end{aligned}$$

Note that each of the three sets has infinitely many descriptions of the form $[\ell]_3$. For example

$$[0]_3 = [-9]_3 = [12]_3 = [-3174]_3 \text{ and } [2]_3 = [-7]_3 = [8]_3 = [477287]_3.$$

We can use the equality sign here because the *sets* represented by the square brackets are equal. In general, $[a]_m = [b]_m$ if and only if $a \equiv b \pmod{m}$. Thus $\mathbb{Z}/m\mathbb{Z}$ has m elements $[0]_m, [1]_m, \dots, [m-1]_m$.

We want to be able to carry the normal arithmetic operations over to $\mathbb{Z}/m\mathbb{Z}$. There is really only one sensible way to do it.

Definition 1.4.10. For $[a]_m, [b]_m \in \mathbb{Z}/m\mathbb{Z}$, define addition and multiplication by

$$[a]_m + [b]_m = [a+b]_m \qquad \textit{and} \qquad [a]_m \times [b]_m = [ab]_m.$$

There is a potential problem, however. We know that $[0]_3 = [-3174]_3$ and $[2]_3 = [477287]_3$. Our definitions give that

$$[0]_3 + [2]_3 = [2]_3 \qquad \text{and} \qquad [-3174]_3 + [477287]_3 = [474113]_3.$$

Is it true that $[2]_3 = [474113]_3$? The answer is 'YES', since Theorem 1.4.5 can be reinterpreted as:

Lemma 1.4.11. *If $a, b, c, d \in \mathbb{Z}$, $[a]_m = [b]_m$ and $[c]_m = [d]_m$ then*

$$[a+c]_m = [b+d]_m \qquad \textit{and} \qquad [ac]_m = [bd]_m.$$

So we have a 'well-defined' addition and multiplication on $\mathbb{Z}/m\mathbb{Z}$. We also have subtraction defined by $[a]_m - [b]_m = [a-b]_m$.

Example 1.4.12. In $\mathbb{Z}/6\mathbb{Z}$, we have $[3]_6 + [4]_6 = [1]_6$, $[3]_6 \times [5]_6 = [3]_6$, $[3]_6 - [5]_6 = [4]_6$.

Remark 1.4.13. It is often convenient to abbreviate the $[a]_m$ notation for elements of $\mathbb{Z}/m\mathbb{Z}$. So we could write $3 + 4 = 1$, $3 \times 5 = 3$ and $3 - 5 = 4$ in $\mathbb{Z}/6\mathbb{Z}$.

Exercise 11. Write down the addition and multiplication tables for $\mathbb{Z}/5\mathbb{Z}$ and $\mathbb{Z}/6\mathbb{Z}$. You will note that the multiplication tables, in particular, look quite different in the two cases. For example, the multiplication table modulo 6 contains more zeroes.

We should not really expect division in general since we cannot always divide one integer by another to obtain a third integer.

Example 1.4.14. In $\mathbb{Z}/4\mathbb{Z}$, the element $[2]_4$ has no multiplicative inverse $[a]_4$ such that $[2]_4 \times [a]_4 = [1]_4$.

The next result tells us which elements of $\mathbb{Z}/m\mathbb{Z}$ have a multiplicative inverse.

Theorem 1.4.15. *The element $[a]_m$ has a multiplicative inverse in $\mathbb{Z}/m\mathbb{Z}$ if and only if* $\gcd(a, m) = 1$.

(In the simplified notation, this says that a has a multiplicative inverse in $\mathbb{Z}/m\mathbb{Z}$ if and only if $\gcd(a, m) = 1$.)

Proof. ($\Leftarrow$) If $\gcd(a, m) = 1$ then Theorem 1.3.8 gives integers x and y so that $ax + ym = 1$. Then $ax \equiv 1 \pmod{m}$, so $[a]_m[x]_m = [1]_m$ in $\mathbb{Z}/m\mathbb{Z}$.

($\Rightarrow$) If $[a]_m[x]_m = [1]_m$ in $\mathbb{Z}/m\mathbb{Z}$, then $ax \equiv 1 \pmod{m}$ so $ax + bm = 1$ for some integer b. But $\gcd(a, m)$ divides a and m, so also divides $1 = ax + bm$ by Lemma 1.2.4(2). Thus $\gcd(a, m) = 1$. □

Example 1.4.16. Find the multiplicative inverse of $[71]_{131}$ in $\mathbb{Z}/131\mathbb{Z}$.

Corollary 1.4.17. *If p is prime, then every non-zero element of $\mathbb{Z}/p\mathbb{Z}$ has a multiplicative inverse.*

Proof. If $[a]_p \neq [0]_p$ then p does not divide a. Since p is prime, the only positive divisors of p are 1 and p. Hence $\gcd(a, p) = 1$ and $[a]_p$ has a multiplicative inverse. □

1.4.2 Exercises

12. Decide whether the following congruences hold.

(a) $3 \equiv 42 \pmod{13}$, (b) $2 \equiv -20 \pmod{11}$,

(c) $26 \equiv 482 \pmod{14}$, (d) $-2 \equiv 933 \pmod{5}$,

(e) $-2 \equiv 933 \pmod{11}$, (f) $-2 \equiv 933 \pmod{55}$.

13. Simplify the following, writing your answers in the form $a \bmod m$ where $0 \leq a < m$.

(a) $482 \pmod{14}$, (b) $511 \pmod{9}$,

(c) $-374 \pmod{11}$, (d) $933 \pmod{55}$,

(e) $102725 \pmod{10}$, (f) $57102725 \pmod{9}$.

(Hint: $10^n \equiv 1 \pmod{9}$ for all $n \in \mathbb{N}$.)

14. Arithmetic modulo m. Calculate the following, giving answers in the form $a \bmod m$ where $0 \leq a < m$. (Hint: it's easiest to reduce modulo m as soon as possible.)

(a) $24 \times 25 \pmod{21}$ (b) $84 \times 125 \pmod{210}$

(c) $25^2 + 24 \times 3 - 6 \pmod{9}$ (d) $36^3 - 3 \times 19 + 2 \pmod{11}$

(e) $1 \times 2 \times 3 \times 4 \times 5 \times 6 \pmod{7}$ (f) $1 \times 2 \times 3 \times \cdots \times 20 \times 21 \pmod{22}$

15. Prove Lemma 1.4.3.

16. Prove parts (3) and (4) of Theorem 1.4.5.

17. Use congruences modulo 9 to show that the following multiplication is incorrect:

$$326 \times 4471 = 1357546.$$

18. Show that if n is an integer with $n \equiv 7 \pmod{8}$, then the equation

$$n = x^2 + y^2 + z^2$$

has no solutions with x, y, z integers.

19. Let F_n be the n-th Fibonacci number, defined by $F_0 = 0$, $F_1 = 1$ and $F_{k+2} = F_k + F_{k+1}$ for all integers $k \geq 0$.

(1) Use Euclid's algorithm to show that $\gcd(F_n, F_{n+1}) = 1$ for all $n \in \mathbb{N}$.

(2) Find integers x_n, y_n such that $x_n F_n + y_n F_{n+1} = 1$.

20. In the following systems $\mathbb{Z}/m\mathbb{Z}$ write down the set of elements that have multiplicative inverses.

(a) $\mathbb{Z}/7\mathbb{Z}$ (b) $\mathbb{Z}/8\mathbb{Z}$ (c) $\mathbb{Z}/12\mathbb{Z}$

(d) $\mathbb{Z}/13\mathbb{Z}$ (e) $\mathbb{Z}/15\mathbb{Z}$.

21. Using Euclid's algorithm, find the multiplicative inverses of the following (if they exist). Here we use a as an abbreviation for $[a]_m$.

(a) 32 in $\mathbb{Z}/27\mathbb{Z}$, (b) 32 in $\mathbb{Z}/39\mathbb{Z}$, (c) 17 in $\mathbb{Z}/41\mathbb{Z}$,

(d) 18 in $\mathbb{Z}/33\mathbb{Z}$, (e) 200 in $\mathbb{Z}/911\mathbb{Z}$.

22. Find the smallest positive integer giving a remainder of 3 when divided by 7, and a remainder of 8 when divided by 11.

23. (Harder) Prove the **Chinese remainder theorem**:
Let m_1, m_2 be relatively prime integers, and let a_1, a_2 be any integers. Then the simultaneous congruences

$$x \equiv a_1 \pmod{m_1} \qquad \text{and} \qquad x \equiv a_2 \pmod{m_2}$$

have a solution x which is unique modulo $m_1 m_2$.

Generalize to an arbitrary number of congruences.

1.5 Fields

In previous subjects, the vector spaces you have seen were mostly over the real numbers. But vector spaces and matrices have very wide application in mathematics and the physical sciences and the ideas are frequently needed in a context where the scalars are other than the real numbers, or even the complex numbers.

1.5.1 Definition and examples of fields

A field is a mathematical system where the notions of addition, multiplication, subtraction and division, work in the same way that they do for real numbers. We need a precise definition of this concept:

Definition 1.5.1. A ***field*** is a set K together with two binary operations, called addition '+' and multiplication '·' (thus if $a, b \in K$ then $a + b$ and $ab = a \cdot b$ are well-defined elements of K) satisfying the following:

Properties of addition

(1) $a + (b + c) = (a + b) + c$ for all $a, b, c \in K$;

(2) $a + b = b + a$ for all $a, b \in K$;

(3) there is an element $0 \in K$ satisfying $0 + a = a$ for all $a \in K$;

(4) for all $a \in K$, there is an element $-a \in K$ satisfying $a + (-a) = 0$;

Properties of multiplication

(5) $a(bc) = (ab)c$ for all $a, b, c \in K$;

(6) $ab = ba$ for all $a, b \in K$;

(7) there is an element $1 \in K$ with $1 \neq 0$ satisfying $1 \cdot a = a$ for all $a \in K$;

(8) for all $a \in K$ with $a \neq 0$, there is an element $a^{-1} \in K$ satisfying $a(a^{-1}) = 1$;

Connecting addition and multiplication

(9) $a(b+c) = (ab) + (ac)$ for all $a, b, c \in K$.

The following examples (and non-examples) illustrate the idea of a field.

Example 1.5.2.

(a) The real numbers $\mathbb{R}$ form a field.

(b) The complex numbers $\mathbb{C}$ form a field.

(c) The rational numbers $\mathbb{Q}$ form a field.

(d) The integers $\mathbb{Z}$ do **not** form a field. (Does 2 have an inverse in $\mathbb{Z}$?)

(e) All expressions $p(x)/q(x)$ form a field, where $p(x)$ and $q(x)$ are polynomials in x with real coefficients and $q(x)$ is not the zero polynomial. This is called the *field of rational functions* (with coefficients in $\mathbb{R}$).

(f) Let K have two elements $\{0, 1\}$ and the following addition and multiplication tables.

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

This is the simplest case of a very important class of finite fields described in the next result.

Theorem 1.5.3. *Let p be a prime. Then $\mathbb{Z}/p\mathbb{Z}$ is a field (using the operations of addition and multiplication modulo p).*

Proof. This is best done after developing some more theory, at which point it becomes mostly trivial. We have already proved the more interesting bits in Lemma 1.4.11 and Corollary 1.4.17. □

Exercise 24. Show that if m is not a prime, then $\mathbb{Z}/m\mathbb{Z}$ is not a field.

Notation 1.5.4. If p is a prime, we write $\mathbb{F}_p$ for the field $\mathbb{Z}/p\mathbb{Z}$.

1.5.2 Algebraically closed fields

Often we'll be interested in solving polynomial equations in a field. In the following special situation, this is always possible.

Definition 1.5.5. A field K is said to be *algebraically closed* if every non-constant polynomial with coefficients in K has a root in K. In other words, if

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

with $a_0, \ldots, a_{n-1} \in K$, $n \geq 1$, then there exists $\alpha \in K$ such that $p(\alpha) = 0$.

We can give little in the way of examples. The real numbers are not algebraically closed: for example, $x^2 + 1$ is a polynomial with coefficients in $\mathbb{R}$ which has no root in $\mathbb{R}$.

Fundamental theorem of algebra. *The field $\mathbb{C}$ is algebraically closed.*

There are many proofs of this result, most of them requiring nontrivial input from other areas of mathematics (complex or real analysis, topology, differential geometry). For a readable exposition of six such proofs, see the book [1].

Theorem 1.5.6. *Every field lies inside an algebraically closed field.*

For a proof, see [3, Section V.2].

1.5.3 Exercises

25. Show that the set of all real numbers of the form $a + b\sqrt{2}$ with $a, b \in \mathbb{Q}$ forms a field with the usual operations of addition and multiplication of the real numbers. (This is a *subfield* of $\mathbb{R}$.)

26. Show that the set of all real numbers of the form $a + b\sqrt[3]{2}$ with $a, b \in \mathbb{Q}$ does not form a field with the usual operations of addition and multiplication of the real numbers. Is there a way to make a field, similar to the previous example, but which contains $\sqrt[3]{2}$ as well as the rational numbers?

27. Write down the multiplication table for $\mathbb{F}_7$. Find an element a of $\mathbb{F}_7$ so that every non-zero element of $\mathbb{F}_7$ is a power of a.

28. Show that the set of all polynomials with real coefficients does not form a field.

29. (Harder) Let $\mathbb{C}((t))$ denote the set of all *formal Laurent series* of the form

$$c_{-k}t^{-k} + c_{-k+1}t^{-k+1} + \cdots + c_{-1}t^{-1} + c_0 + c_1 t + \cdots + c_s t^s + \ldots$$

with the usual operations of addition and multiplication of Laurent series. Show that $\mathbb{C}((t))$ forms a field. You should ignore the question of whether the series are convergent.

30. Show that the field in Exercise 25 is not algebraically closed.

31. (Harder) Show that, for every prime p, $\mathbb{F}_p$ is not algebraically closed.

2 Linear Algebra

2.1 Revision

2.1.1 Vector spaces and subspaces

We begin with the formal definition of vector spaces.

Definition 2.1.1. Let K be a field. A ***vector space*** over K is a set V with two binary operations, *addition* $V \times V \to V$ (the image of (u, v) will be denoted $u + v$) and *scalar multiplication* $K \times V \to V$ (the image of (a, v) being denoted av). These are required to satisfy the following axioms:

Properties of addition:

(1) $u+(v+w) = (u+v)+w$ for all $u, v, w \in V$

(2) there is an element $0 \in V$ satisfying $0 + v = v + 0 = v$ for all $v \in V$

(3) for each $v \in V$, there is an element $-v \in V$ such that $v + (-v) = (-v) + v = 0$

(4) $u + v = v + u$ for all $u, v \in V$

Properties of scalar multiplication:

(5) $a(u + v) = au + av$ for all $a \in K$, $u, v \in V$

(6) $(a + b)v = av + bv$ for all $a, b \in K$, $v \in V$

(7) $(ab)v = a(bv)$ for all $a, b \in K$, $v \in V$

(8) $1v = v$ for all $v \in V$

Example 2.1.2. (1) Set $K = \mathbb{R}$ and $V = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$ with addition and scalar multiplication defined by:

$$(x, y, z) + (x', y', z') = (x + x', y + y', z + z') \quad \text{and} \quad c(x, y, z) = (cx, cy, cz).$$

This is the standard vector space $\mathbb{R}^3$.

(2) Let K be an arbitrary field and $V = \{(a_1, a_2, \ldots, a_n) \mid a_1, \ldots a_n \in K\}$ with addition and scalar multiplication defined by:

$$\begin{aligned}(a_1, \ldots, a_n) + (b_1, \ldots, b_n) &= (a_1 + b_1, \ldots, , a_n + b_n)\\ c(a_1, \ldots, a_n) &= (ca_1, \ldots, ca_n).\end{aligned}$$

Denote this vector space by K^n. The first example is a special case of this.

(3) Let $K = \mathbb{R}$ and let $M_{m\times n}(\mathbb{R})$ denote the set of $m \times n$ matrices with entries from $\mathbb{R}$. Then $M_{m\times n}(\mathbb{R})$, furnished with the usual addition and scalar multiplication of matrices, is a vector space. This example also works when we replace $\mathbb{R}$ by a general field.

(4) Let K be a field. Then the set of polynomials with coefficients in K, with the usual addition and scalar multiplication of polynomials, forms a vector space $K[x]$.

(5) As in the previous example, but consider only polynomials of degree at most d, for some fixed natural number d. Call the resulting space $K[x]_{\leq d}$.

(6) The set $\mathbb{R}^{\mathbb{R}} = \mathcal{F}(\mathbb{R}, \mathbb{R})$ of all functions $f\colon \mathbb{R} \to \mathbb{R}$ forms a vector space over the field of real numbers. Addition of two such functions f and g is given by:

$f + g$ is the function defined by $(f + g)\colon x \mapsto f(x) + g(x)$

and scalar multiplication, for $a \in \mathbb{R}$ is given by:

af is the function defined by $(af)\colon x \mapsto af(x)$.

(7) As in the previous example, but allow the set $K^S = \mathcal{F}(S, K)$ of functions $f\colon S \to K$, where S is an arbitrary set and K is a field. This is a vector space over K.

(8) The set of solutions y of the differential equation

$$\frac{d^2y}{dx^2} + 7\frac{dy}{dx} + 23y = 0$$

forms a vector space if we use the addition and scalar multiplication of functions defined above.

(9) Let $K = \mathbb{R}$ and let $V = \mathbb{R}^\infty$ be the set of all sequences $\{a_n\}, a_n \in \mathbb{R}$. Define addition and scalar multiplication by:

$$\{a_n\} + \{b_n\} = \{a_n + b_n\} \qquad \text{and} \qquad x\{a_n\} = \{xa_n\}.$$

Note that this is really a special case of Example 7 since we can regard a sequence as a function $\mathbb{N} \to \mathbb{R}$.

(10) As above but restrict to sequences that satisfy $\lim_{n\to\infty} a_n = 0$.

(11) If we restrict Example 9 to sequences that satisfy $\lim_{n\to\infty} a_n = 1$ then we do **not** obtain a vector space.

Definition 2.1.3. Let V be a vector space over the field K. A ***subspace*** of V is a subset W of V such that W is itself a vector space using the operations of addition and scalar multiplication from V.

If we take a subset of $\mathbb{R}^3$, say $\{(a, b, c) \mid a, b, c \in \mathbb{R}, a + b + c = 0\}$ and start checking whether it is a subspace, we find that many of the checks are essentially trivial. Briefly, we know that the operations behave well because the ambient space, in this case $\mathbb{R}^3$, is a vector space. When we eliminate all the things we don't need to check for this reason, we are left with the following.

Lemma 2.1.4. *Let V be a vector space over K. A subset W of V is a subspace if and only if the following three conditions are satisfied:*

(1) $0 \in W$;

(2) if $u, w \in W$ *then* $u + w \in W$;

(3) if $a \in K$ *and* $w \in W$ *then* $aw \in W$.

Example 2.1.5. (1) The set $W = \{(a, b, c) \mid a, b, c \in \mathbb{R}, a + b + c = 0\}$ is a subspace of $\mathbb{R}^3$.

(2) The set of matrices of trace zero is a subspace of $M_{n\times n}(\mathbb{R})$.

(3) The set of polynomials with zero constant term is a subspace of $K[x]$.

(4) The set of differentiable functions is a subspace of $\mathbb{R}^\mathbb{R} = \mathcal{F}(\mathbb{R}, \mathbb{R})$.

(5) The set of sequences with $\lim_{n\to\infty} a_n = 0$ is a subspace of the space of all sequences.

2.1.2 Spanning, linear dependence, bases

Definition 2.1.6. If S is a subset of a vector space V then a ***linear combination*** of S is an finite sum of the form

$$\sum_{i=1}^{n} a_i s_i \qquad \text{where } a_i \in F, s_i \in S.$$

The set of all linear combinations of elements of S is called the ***span*** of S and is denoted by $\langle S\rangle$. We also say that S is a spanning set for $\langle S\rangle$.

Lemma 2.1.7. *If S is a non-empty subset of V, then $\langle S\rangle$ is a subspace of V.*

Examples:

(1) The set of all linear combinations of the vectors $(1, -2, 3)$ and $(0, 2, 1)$ in $\mathbb{R}^3$ is the set $\{(a, -2a + 2b, 3a + b) : a, b \in \mathbb{R}\}$.

(2) The set of all linear combinations of the matrices

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

in $M_{3\times 3}(\mathbb{R})$ is the set of all matrices of the form

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

where $a, b, c \in \mathbb{R}$.

Definition 2.1.8. We say that a subset S of a vector space V is ***linearly dependent*** if some non-zero linear combination gives the zero vector:

$$\sum_{i=1}^{n} a_i s_i = 0 \quad \text{where } a_i \in F, s_i \in S \text{ and not all } a_i \text{ are zero.}$$

Otherwise, S is said to be ***linearly independent***.

Examples:

(1) The set $\{(1, 2, 3), (2, -1, 0), (-1, 8, 9)\}$ is linearly dependent in $\mathbb{R}^3$.

(2) The set $\{1, x, x^2, 1 + x^3\}$ is linearly independent in $\mathcal{P}(\mathbb{R})$.

(3) The set $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 2 & -29 \\ 0 & 0 \end{bmatrix}\}$ is linearly dependent in $M_{2\times 2}(\mathbb{R})$.

Lemma 2.1.9. *A subset S of a vector space V is linearly dependent if and only if some element s of S is a linear combination of the others.*

In this case removing s from S gives a *smaller* spanning set for the subspace $\langle S \rangle$. Making the spanning set as small as possible leads to the idea of basis.

Definition 2.1.10. A ***basis*** for a vector space V is a linearly independent spanning set.

Examples:

(1) The standard basis for F^n is the set

$$\{e_1 = (1, 0, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)\}.$$

(2) The set $\{(2, 1, 3), (1, 2, 3), (1, 0, 0)\}$ is a basis of $\mathbb{R}^3$.

(3) The set $\{1, x, x^2, 1 + x^3\}$ is a basis of $\mathcal{P}_3(\mathbb{R})$.

(4) The set $\{1, x, x^2, x^3, x^4, \ldots, x^n, \ldots\}$ is a basis of $\mathcal{P}(\mathbb{R})$.

Theorem 2.1.11. *Every vector space has a basis. In fact, every spanning set contains a basis and every linearly independent set can be extended to a basis.*

Theorem 2.1.12. *If $\mathcal{B}_1$ and $\mathcal{B}_2$ are two bases of a vector space then they have the same number of elements. (In general, this means that there exists a bijective function $f \colon \mathcal{B}_1 \to \mathcal{B}_2$.)*

Definition 2.1.13. The *dimension* of a vector space V is the number of elements in a basis. We usually write this as $\dim V$.

By Theorem 2.1.12, we know that this number will not depend on the particular choice of basis.

Examples: For the examples after Definition 2.1.1:

(1) $\mathbb{R}^3$ has dimension 3.

(2) F^n has dimension n.

(3) $M_{m\times n}(\mathbb{R})$ has dimension mn.

(4) $\mathcal{P}_n(F)$ has dimension $n+1$.

(5) Example 8 has dimension 2 (although this needs a bit of work).

(6) All of the other examples have infinite dimension.

Combining subspaces: Let U and W be subspaces of a vector space V. Then the *intersection* $U \cap W = \{v \in V : v \in U \text{ and } v \in W\}$ and the *sum* $U + W = \{u + w : u \in U, w \in W\}$ are both *subspaces* of V. (See exercises 1, 2.) In fact $U + W$ is the smallest subspace containing both U and W.

Lemma 2.1.14. *Let U and W be subspaces of a vector space V and assume that $U + W$ is finite dimensional. Then*

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W.$$

Proof. Let $\{v_1, \dots, v_l\}$ be a basis of $U \cap W$. Then $\{v_1, \dots, v_l\}$ is a linearly independent set in U and so can be extended to a basis $\{v_1, \dots, v_l, u_1, \dots, u_m\}$ of U. Similarly $\{v_1, \dots, v_l\}$ can be extended to a basis $\{v_1, \dots, v_l, w_1, \dots, w_n\}$ of W. We claim that $\{v_1, \dots, v_l, u_1, \dots, u_m, w_1, \dots, w_n\}$ is a basis of $U + W$.

Since every element of U is a linear combination of $\{v_1, \dots, v_l, u_1, \dots, u_m\}$ and every element of W is a linear combination of

$$\{v_1, \dots, v_l, w_1, \dots, w_n\},$$

it is clear that the sum of an element of U and an element of W is a linear combination of

$$\{v_1, \dots, v_l, u_1, \dots, u_m\} \cup \{v_1, \dots, v_l, w_1, \dots, w_n\} \\ = \{v_1, \dots, v_l, u_1, \dots, u_m, w_1, \dots, w_n\}. \quad (2.1)$$

So $\{v_1, \dots, v_l, u_1, \dots, u_m, w_1, \dots, w_n\}$ spans $U + W$.

Suppose that we have

$$\sum_i a_i v_i + \sum_j b_j u_j + \sum_k c_k w_k = 0 \text{ with } a_i, b_j, c_k \in F.$$

Then $\sum_k c_k w_k$ is a linear combination of elements of U and so lies in $U \cap W$. Thus $\sum_k c_k w_k$ can be written as a linear combination of the basis $\{v_1, \dots, v_l\}$ of $U \cap W$. Thus we have

$$\sum_k c_k w_k = \sum_i d_i v_i \text{ for some } d_i \in F.$$

But

$$\{v_1, \dots, v_l, w_1, \dots, w_n\}$$

is a basis of W and so linearly independent. Thus each c_k and each d_i is zero. Now we have $\sum_i a_i v_i + \sum_j b_j u_j = 0$. But $\{v_1, \dots, v_l, u_1, \dots, u_m\}$ is a basis of U and so linearly independent. Thus each a_i and b_j is zero. Hence

$$\{v_1, \dots, v_l, u_1, \dots, u_m, w_1, \dots, w_n\}$$

is linearly independent and so is a basis for $U + W$.

We now have $\dim(U \cap W) = l$, $\dim U = l + m$, $\dim W = l + n$ and $\dim(U + W) = l + m + n$. The result follows immediately.

□

2.1.3 Linear transformations

Informally, a linear transformation is a function between vector spaces over the same field which preserves the operations of addition and scalar multiplication.

Definition 2.1.15. Let V and W be vector spaces over the same field F. A function $f : V \to W$ is a ***linear transformation*** if

(1) $f(u+v) = f(u) + f(v)$ for all $u, v \in V$;

(2) $f(av) = af(v)$ for all $a \in F, v \in V$.

Examples:

(1) Rotation about the origin through a fixed angle θ is a linear transformation on $\mathbb{R}^2$.

(2) Rotation about any line through the origin and through a fixed angle θ is a linear transformation on $\mathbb{R}^3$.

(3) Differentiation is a linear transformation on $\mathcal{P}(\mathbb{R})$.

(4) Let $\mathcal{C}$ denote the subspace of $\mathbb{R}^{\mathbb{R}}$ consisting of continuous functions. Let the function $I : \mathcal{C} \to \mathcal{C}$ be given by defining $I(f)$ to be the function whose value at t is

$$I(f)[t] = \int_0^t f(x)\,dx.$$

Then I is a linear transformation.

(5) The functions $f, g : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^2$ and $g(x) = x + 2$ are *not* linear transformations.

Definition 2.1.16. Let $f : V \to W$ be a linear transformation. The ***nullspace*** (or ***kernel***) of f is $\{v \in V : f(v) = 0\}$. The ***range*** (or ***image***) of f is $\{f(v) : v \in V\}$.

Examples:

(1) Rotation in $\mathbb{R}^2$ has nullspace 0 and range the whole of $\mathbb{R}^2$.

(2) Differentiation on $\mathcal{P}(\mathbb{R})$ has nullspace $\langle 1 \rangle$ and range $\mathcal{P}(\mathbb{R})$.

The following should not be too surprising, nor too hard to prove.

Lemma 2.1.17. *Let $f : V \to W$ be a linear transformation. The nullspace of f is a subspace of V and the range of f is a subspace of W.*

Definition 2.1.18. Let $f : V \to W$ be a linear transformation. The dimension of the nullspace of f is called the ***nullity*** of f and the dimension of the range of f is called the ***rank*** of f.

Lemma 2.1.19. *Let $f : V \to W$ be a linear transformation and assume that V is finite dimensional. The nullity of f plus the rank of f is equal to the dimension of V.*

Sketch of proof. Denote the nullspace of f by N. Since it is a subspace of V it will have a basis $\mathcal{B} = \{v_1, \ldots, v_m\}$. So m is the nullity of f. Since $\mathcal{B}$ is a basis of N, it is linearly independent in N. Since N is a subspace of V, $\mathcal{B}$ is also linearly independent in V. So we can extend $\mathcal{B}$ to a basis of $\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ of V. So the dimension of V is n.

We claim that $\{f(v_{m+1}), \ldots, f(v_n)\}$ is a basis of the range of V. We must show that $\{f(v_{m+1}), \ldots, f(v_n)\}$ is linearly independent and that every element of the range of V can be expressed as a linear combination of $\{f(v_{m+1}), \ldots, f(v_n)\}$. We leave the details as Exercise 47.

We will have shown that f has nullity m and rank $n - m$ where n is the dimension of V. The theorem now follows. □

2.1.4 Matrix representations

Any $n \times m$ matrix $A \in M_{n\times m}(F)$ gives a linear transformation $f_A : F^m \to F^n$ defined by matrix multiplication: $f_A(x) = Ax$ for $x \in F^m$ where we think of vectors in F^m, F^n as *column vectors.* Note that the ith column of A is $f_A(e_i)$ where e_i is the ith standard basis vector for F^m.

Conversely, any linear transformation $f\colon V \to W$ between finite dimensional vector spaces V and W over a field F can be represented by a matrix: Let $\mathcal{B}_V = \{v_1, v_2, \ldots, v_m\}$ be an ordered basis for V and $\mathcal{B}_W = \{w_1, w_2, \ldots, w_n\}$ be an ordered basis for W. Then $f(v_i) \in W$ for each $i = 1, \ldots, m$ and we can write $f(v_i)$ uniquely as a linear combination of the basis vectors in $\mathcal{B}_W$. We form an $n \times m$ matrix A with these coefficients as the ith *column*.

Definition 2.1.20. This matrix A is called the ***matrix of*** f with respect to the bases $\mathcal{B}_V$ and $\mathcal{B}_W$.

Explicitly, if

$$\begin{array}{rcccccccc} f(v_1) & = & a_{11}w_1 & + & a_{21}w_2 & + & \ldots & + & a_{n1}w_n \\ f(v_2) & = & a_{12}w_1 & + & a_{22}w_2 & + & \ldots & + & a_{n2}w_n \\ \vdots & = & \vdots & & \vdots & & & & \vdots \\ f(v_m) & = & a_{1m}w_1 & + & a_{2m}w_2 & + & \ldots & + & a_{nm}w_n \end{array}$$

with each $a_{ij} \in F$. Then $A = (a_{ij})$.

It is often the case that $V = W$ and $\mathcal{B}_V = \mathcal{B}_W$. Then we say that f has matrix A with respect to $\mathcal{B}_V$.

Examples:

(1) The rotation about the origin through an angle of θ in $\mathbb{R}^2$ is a linear transformation taking $(1,0)$ to $(\cos\theta, \sin\theta) = \cos\theta(1,0) + \sin\theta(0,1)$ and $(0,1)$ to $(-\sin\theta, \cos\theta) = -\sin\theta(1,0) + \cos\theta(0,1)$. So its matrix with respect to the basis $\{(1,0),(0,1)\}$ is

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

(2) Differentiation gives a linear transformation $D : \mathcal{P}_3(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$. The matrix with respect to the bases $\{1, x, x^2, x^3\}$ and $\{1, x, x^2\}$ is

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Interpretation of the matrix representation:

Given an (ordered) basis $\mathcal{B}_V = \{v_1, v_2, \ldots, v_m\}$ for a vector space V, each vector $v \in V$ can be written *uniquely* as a linear combination

$$v = a_1v_1 + \ldots + a_mv_m, \quad \alpha_i \in F.$$

This allows us to introduce ***coordinates*** on V: the column vector

$$[v]_{\mathcal{B}_V} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \in F^m$$

is called the ***coordinate vector*** of v with respect to the basis $\mathcal{B}_V$.

Then the effect a linear transformation $f : V \to W$ on coordinate vectors is just multiplication by the matrix A representing f:

$$[f(v)]_{\mathcal{B}_W} = A[v]_{\mathcal{B}_V}.$$

In summary, we have

$$\begin{tikzcd} v \in V \arrow[r, "\text{apply } f"] \arrow[d, "\text{take coords}"'] & f(v) \in W \arrow[d, "\text{take coords}"] \\ {[v]_{\mathcal{B}_V} \in F^m} \arrow[r, "\text{mult by } A"] & {[f(v)]_{\mathcal{B}_W} \in F^n.} \end{tikzcd}$$

2.1.5 Change of basis

Any linear transformation will have different matrices for different bases of the underlying vector spaces. It is very useful to be able to choose a basis so that the matrix is as simple as possible. To do this, we need to be able to see the effect on the matrix of changing the basis.

Let V, W be finite dimensional vector spaces over a field F and let $f : V \to W$ be a linear transformation. Let $\mathcal{B}_V = \{v_1, v_2, \ldots, v_m\}$ be a basis for V and $\mathcal{B}_W = \{w_1, w_2, \ldots, w_n\}$ be a basis for W. Suppose that $\mathcal{B}'_V = \{v'_1, v'_2, \ldots, v'_m\}$ is a new basis for V and $\mathcal{B}'_W = \{w'_1, w'_2, \ldots, w'_n\}$ is a new basis for W. Then we can convert $\mathcal{B}_V$-coordinates to $\mathcal{B}'_V$-coordinates using the matrix P with ith column $[v_i]_{\mathcal{B}'_V}$. Similarly we can convert $\mathcal{B}_W$-coordinates to $\mathcal{B}'_W$-coordinates using the matrix Q with ith column $[w_i]_{\mathcal{B}'_W}$.

Explicitly, $P = (p_{ij})$ and $Q = (q_{ij})$ where

$$v_i = \sum_{j=1}^{m} p_{ji} v'_j \text{ and } w_i = \sum_{j=1}^{m} q_{ji} w'_j.$$

Theorem 2.1.21. *The matrices P and Q are invertible and the matrix of f with respect to the bases $\mathcal{B}'_V$ and $\mathcal{B}'_W$ is*

$$QAP^{-1},$$

where A is the matrix of f with respect to the bases $\mathcal{B}_V$ and $\mathcal{B}_W$.

Thus we have the following diagram:

$$\begin{tikzcd}
{[v]_{\mathcal{B}_V}} \arrow[r, "A"] \arrow[d, "P"'] & {[f(v)]_{\mathcal{B}_W}} \arrow[d, "Q"] \\
{[v]_{\mathcal{B}'_V}} \arrow[r, "QAP^{-1}"] & {[f(v)]_{\mathcal{B}'_W}.}
\end{tikzcd}$$

In the most important case where $V = W$ and $\mathcal{B}_V = \mathcal{B}_W$, we also have $P = Q$ and so, if A is the matrix of f with respect to the old basis then PAP^{-1} is the matrix of f with respect to the new basis.

Example: Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation defined by
$f(x, y) = (3x - y, -x + 3y)$. Using the standard basis $\mathcal{B} = \{(1, 0), (0, 1)\}$ we find the matrix of f is

$$A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}.$$

Now let's calculate the matrix with respect to the basis $\mathcal{B}' = \{(1, 1), (-1, 1)\}$. We have

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

and

$$f(-1, 1) = (-4, 4) = 0(1, 1) + 4(-1, 1).$$

Thus the matrix for f with respect to basis $\mathcal{B}'$ is the diagonal matrix

$$A' = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}.$$

This makes it easy to understand the effect of the transformation f: it just stretches by a factor 2 in the $(1, 1)$ direction and by a factor 4 in the $(-1, 1)$ direction.

Alternatively we can use the change of basis formula in the previous theorem. The transition matrix from $\mathcal{B}'$ to the standard basis $\mathcal{B}$ is $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ so the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$ is the *inverse* of this:

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Then

$$A' = PAP^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix},$$

as before.

Definition 2.1.22. Two $n \times n$ matrices A and B are said to be ***similar*** if $B = PAP^{-1}$ for some invertible matrix P.

Thus similar matrices represent the same linear transformation with respect to different bases.

2.1.6 Exercises

32. If U and W are subspaces of a vector space V, show that $U + W = \{u + w : u \in U, w \in W\}$ is also a subspace.

33. Show that, if U_1 and U_2 are subspaces of a vector space V then $U_1 \cap U_2$ is also a subspace.

34. If U_1 and U_2 are subspaces of a vector space V and $U_1 \cup U_2 = V$, show that either $U_1 = V$ or $U_2 = V$.

35. Decide whether the following sets of vectors are (i) linearly dependent and (ii) a basis, in $\mathbb{F}_7^4$.

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

(b) $\{(1, 2, 3, 1), (4, 6, 2, 0), (0, 1, 5, 1)\}$;

(c) $\{(1, 2, 3, 1), (4, 6, 2, 0), (0, 1, 5, 2), (0, 1, 0, 0), (0, 1, 0, 1)\}$.

36. Decide whether the following sets of matrices are linearly independent in the space $M_{2\times 2}(\mathbb{R})$:

(a) $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$;

(b) $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$;

(c) $\left\{ \begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 4 & -6 \\ 3 & 8 \end{bmatrix} \right\}$.

37. Show that any subset of a linearly independent set is also linearly independent.

38. Let F be a field and let $E_{ij} \in M_{m\times n}(F)$ be the matrix with 1 in the i, j position and 0 elsewhere. Show that $\{E_{i,j} : 1 \le i \le m, 1 \le j \le n\}$ is a basis of $M_{m\times n}(F)$.

39. Show that the space $\mathcal{P}(F)$ does not have finite dimension.

40. What is the dimension of the space $M_{3\times 3}(\mathbb{F}_5)$?

41. Let B be the matrix $\begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}$. Show that the function $g : M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ given by $A \mapsto AB$ for $A \in M_{2\times 2}(\mathbb{R})$ is a linear transformation.

42. Find the matrix of the linear transformation of Question 41 with respect to the basis

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

found in Question 38.

43. Find the matrix, with respect to the standard basis of $\mathbb{R}^2$, of the reflection in the x-axis. Now let $\mathcal{B}$ be the basis $\{(a, b), (c, d)\}, ad - bc \neq 0$ of $\mathbb{R}^2$. Write down a change of basis matrix for the change from the standard basis to $\mathcal{B}$ and so calculate the matrix of the reflection with respect to this new basis.

44. Calculate the nullity and rank of the linear transformation f on $\mathbb{R}^3$ given by (here e_1, e_2, e_3 is the standard basis)

$$f(e_1) = e_1 - e_2; f(e_2) = e_2 - e_3; f(e_3) = e_1 - e_3.$$

45. Calculate the nullity and rank of the linear transformation f on $\mathbb{Z}_7^3$ given by

$$\begin{aligned} f\left(([1]_7, [0]_7, [0]_7)\right) &= ([1]_7, [2]_7, [3]_7); \\ f\big(([0]_7, [1]_7, [0]_7]_)\big) &= ([3]_7, [4]_7, [5]_7); \\ f\left(([0]_7, [0]_7, [1]_7)\right) &= ([5]_7, [1]_7, [4]_7). \end{aligned}$$

46. Let $f \colon V \to V$ be a linear transformation on a finite dimensional vector space V. Show that the nullity of f is zero if and only if f is surjective.

47. Complete the proof of Lemma 2.1.19.

2.2 Normal forms

In this section, we shall consider a linear transformation $f : V \to V$ on a vector space V and study the problem of finding a basis $\mathcal{B}$ of V so that the matrix of f with respect to $\mathcal{B}$ is as simple as possible. Often we can choose a *diagonal* matrix representing f, but in general the best we can find is the *Jordan normal form* (see Section 2.2.4).

2.2.1 Eigenvalues and eigenspaces, invariant subspaces

Definition 2.2.1. Suppose that $f(v) = av$ for some non-zero $v \in V$ and some $a \in F$. Then a is called an ***eigenvalue*** of f and v is said to be an eigenvector corresponding to a. The set of all solutions to the equation $f(v) = av$ which correspond to a fixed eigenvalue a is a subspace of V, called the ***eigenspace*** corresponding to a.

Similarly we define eigenvalues, eigenvectors and eigenspaces for any square matrix A.

Examples:

(1) A rotation fixing the origin on $\mathbb{R}^3$ has an eigenvalue of 1 and a corresponding eigenspace of dimension 1, the axis of the rotation.

(2) A reflection fixing the origin on $\mathbb{R}^2$ has two eigenvalues, 1 and -1. The eigenspace corresponding to 1 is the line of reflection. The eigenspace corresponding to -1 is the perpendicular to the line of reflection.

(3) A reflection fixing the origin on $\mathbb{R}^3$ has three eigenvalues, 1 (twice) and -1. The eigenspace corresponding to 1 is the plane of reflection. The eigenspace corresponding to -1 is the line perpendicular to the plane of reflection.

(4) The eigenvalues of the (linear transformation corresponding to the) matrix

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

satisfy $\det(A - \lambda I) = 0$. So the eigenvalues are $2, -1, 0$. The corresponding eigenspaces are generated by the eigenvectors

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ 8 \\ 2 \end{bmatrix}$$

respectively. Each eigenspace has dimension 1.

(5) Let $\mathcal{D}$ denote the subspace of functions in $\mathbb{R}^{\mathbb{R}}$ which are differentiable infinitely often. Then differentiation gives a linear transformation
$\delta : \mathcal{D} \to \mathcal{D}$ and

$$\delta(e^{ax}) = \frac{d}{dx} e^{ax} = ae^{ax}.$$

So *every* real number a is an eigenvalue of δ with corresponding eigenvector e^{ax}.

Observe that, if a is an eigenvalue and V_a is the corresponding eigenspace, then $v \in V_a$ implies $f(v) \in V_a$ (because $f(v) = av$). Subspaces with this property are of special interest.

Definition 2.2.2. Let f be a linear transformation on a vector space V and let W be a subspace of V. We say that W is ***f-invariant*** if $f(w) \in W$ for every $w \in W$.

If W is f-invariant then f defines a linear transformation $f_W : W \to W$ called the ***restriction*** of f to W (forget all elements of V which are not in W). Invariant subspaces are useful when we are trying to pick a 'good' basis with which to represent a linear transformation.

Lemma 2.2.3. *Let $f : V \to V$ be a linear transformation and let W be an f-invariant subspace, with $\dim V = n$ and $\dim W = m$. Let $\mathcal{B}_1 = \{w_1, \ldots, w_m\}$ be a basis for W, and extend it to a basis*

$\mathcal{B} = \{w_1, \ldots, w_m, w_{m+1}, \ldots, w_n\}$ for V. Then the matrix of f with respect to $\mathcal{B}$ is of the "block" form

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

where A, B, D are matrices and A is the $m \times m$ matrix of f_W.

Proof. The first m columns represent the images of the elements of $\mathcal{B}_1$. These images lie in W, as W is f-invariant, and so can be expressed in terms of the elements of $\mathcal{B}_1$ only. It follows that the first m columns have non-zero entries only in the first m rows. □

Examples:

(1) Let f be a rotation on $\mathbb{R}^3$. Then the plane perpendicular to the axis of rotation is an invariant subspace of f. If we use two (orthonormal) vectors from this plane together with a unit vector along the axis of rotation, the matrix for the rotation becomes

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

(2) Suppose that a linear transformation on $\mathbb{R}^3$ has matrix

$$\begin{bmatrix} 3 & 2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

with respect to the basis $\{e_1, e_2, e_3\}$. Then the subspace $W = \langle e_1, e_2 \rangle$ is f-invariant and the matrix of f_W with respect to the basis $\{e_1, e_2\}$ of W is

$$\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}.$$

In order to make more progress with this, we need the idea of a complement to a subspace.

Definition 2.2.4. Let V be a vector space and let W be a subspace of V. Then a subspace U of V is a ***complement*** to W if $U \cap W = \{0\}$ and $U + W = V$. We then write $U \oplus W = V$, and say that V is a ***direct sum*** of U and W.

Examples:

(1) In $\mathbb{R}^3$, a complement to a plane through the origin is any line through the origin which does not lie in the plane.

(2) In $\mathbb{R}^4$, the subspaces $\langle (1,0,0,0), (0,1,0,0) \rangle$ and $\langle (0,0,1,0), (0,0,0,1) \rangle$ are complementary.

(3) In $\mathcal{P}(\mathbb{R})$, the subspaces $\langle 2, 1+x, 1+x+x^3 \rangle$ and $\langle x^2 + 3x^4, x^4, x^5, x^6, \ldots, x^n, \ldots \rangle$ are complementary.

Lemma 2.2.5. *Let V be a finite dimensional vector space and let U, W be subspaces of V. The following are equivalent.*

(1) U is a complement of W;

(2) There is a basis $\mathcal{B}$ of V of the form $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ where $\mathcal{B}_1$ is a basis of U, $\mathcal{B}_2$ is a basis of W and $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$.

(3) $U \cap W = \{0\}$ and $\dim U = \dim V - \dim W$;

(4) $V = U + W$ and $\dim U = \dim V - \dim W$.

Proof. We shall make frequent use of Lemma 2.1.14.

(1) implies (2). Let $\mathcal{B}_1$ be a basis of U and $\mathcal{B}_2$ be a basis of W. Then it is easy to check that $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$ and $\mathcal{B}_1 \cup \mathcal{B}_2$ spans $U + W = V$. But $\dim V = \dim(U+W) = \dim U + \dim W$ and so $\mathcal{B}_1 \cup \mathcal{B}_2$ is a spanning set of V which has the same number of elements as a basis. It must therefore be a basis.

(2) implies (3). From (2) we can quickly deduce that $\dim V = \dim U + \dim W$ and that $V = U + W$. Thus $\dim(U \cap W) = \dim U + \dim W - \dim(U+W) = 0$; hence $U \cap W = \{0\}$.

(3) implies (4). Assuming (3) we have

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W) = \dim U + \dim W = \dim V.$$

Thus $U + W = V$.

(4) implies (1). Assuming (4) we have that $V = U + W$. Also, $\dim(U \cap W) = \dim U + \dim W - \dim V = 0$ and so $U \cap W = \{0\}$. □

We can now obtain an even better version of Lemma 2.2.3.

Lemma 2.2.6. *Let V be a vector space and let f be a linear transformation on V. Let U, W be complementary subspaces of V (i.e. $U \oplus W = V$). Suppose that both U and W are f-invariant. Choose an ordered basis $\mathcal{B}$ of V of the form $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$ where $\mathcal{B}_1$ is a basis of U and $\mathcal{B}_2$ is a basis of W. Then the matrix of f with respect to $\mathcal{B}$ is of the "block diagonal" form:*

$$\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$$

where A is the matrix of f_U and D is the matrix of f_W.

Proof. See Exercise 54. □

Thus we can now see a way to simplify the matrix for a linear transformation f. We must find complementary f-invariant subspaces. The next section will give us a way to do that.

2.2.2 Minimal polynomials

Let V be a vector space of finite dimension n over a field F. Let $f : V \to V$ be a linear transformation, and let A be a matrix representing f with respect to some basis. Given any polynomial

$$p(X) = a_0 + a_1 X + \ldots + a_k X^k, a_i \in F$$

we can apply it to the matrix A:

$$p(A) = a_0 I + a_1 A + a_2 A^2 + \ldots + a_k A^k,$$

where I is the $n \times n$ identity matrix. Similarly we define

$$p(f) = a_0 1_V + a_1 f + a_2 f^2 + \ldots + a_k f^k,$$

where $1_V : V \to V$ is the identity transformation and $f^k = f \circ \ldots \circ f : V \to V$ is f composed with itself k times.

Now A lies in the vector space $M_{n\times n}(F)$ of all $n \times n$ matrices over F. The powers of A, which represent the powers of f, also lie in $M_{n\times n}(F)$. So $\{I_n, A, A^2, \ldots, A^k, \ldots\}$ is an apparently infinite subset of the finite dimensional vector space $M_{n\times n}(F)$. Thus these powers must be linearly dependent and so there is some expression $\sum_i a_i A^i = 0$ with $a_i \in F$, not all zero. There is therefore a similar expression $\sum_i a_i f^i = 0$. If we let $q(X)$ be the polynomial $\sum_i a_i X^i = 0$ we can write this as $q(f) = 0$. Observe that we can always divide through $q(X)$ by the coefficient of the term of highest degree to ensure that, in the result, this coefficient is 1. We call such a polynomial ***monic***. That is, a ***monic polynomial*** is one in which the term of highest degree has coefficient 1.

Definition 2.2.7. Let f be a linear transformation on a vector space V of finite dimension. The ***minimal polynomial*** of f is the monic polynomial $m(X)$ of lowest degree such that $m(f) = 0$.

We define the minimal polynomial of a matrix A in a similar way.

Examples:

(1) Let f denote a reflection, in a line through the origin, in $\mathbb{R}^2$. Then the minimal polynomial of f is $X^2 - 1$. It is clear that f^2 is the identity on $\mathbb{R}^2$. Also, if the minimal polynomial were to have smaller degree, then it would be of degree 1 and f would be a scalar multiple of the identity, which is false.

(2) Let f be a linear transformation with matrix

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then the minimal polynomial of f is $(X-2)(X-3)X$.

(3) Let f be a linear transformation with matrix

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then the minimal polynomial of f is $(X-2)(X-1)$.

Lemma 2.2.8. *If $m(X)$ is the minimal polynomial of f and if $q(X)$ is a polynomial with coefficients in F such that $q(f) = 0$ then $m(x)$ divides $q(X)$.*

Proof. Use 'division with remainder' of polynomials to write

$$q(X) = s(X)m(X) + r(X)$$

where $r(X)$ is a polynomial which is either 0 or is of degree less than the degree of $m(X)$. But then $q(f) = s(f)m(f) + r(f)$; that is, $0 = s(f)0 + r(f)$. Thus $r(f) = 0$. But this will contradict the definition of $m(X)$ unless $r(X) = 0$. So $r(X) = 0$ and so $q(X) = s(X)m(X)$; that is, $m(X)$ divides $q(X)$. □

This helps a little in finding the minimal polynomial. If we can find some polynomial $q(X)$ with $q(f) = 0$ then we can look amongst its factors for the minimal polynomial. In fact we shall see later that the characteristic polynomial is a candidate for $q(X)$. The next lemma tells us that only factors big enough to have all of the eigenvalues as roots need be checked.

Lemma 2.2.9. *The roots of the minimal polynomial of f are precisely the eigenvalues of f.*

Proof. Let $m(X)$ be the minimal polynomial of f.

Let a be an eigenvalue of f with corresponding eigenvector $v \neq 0$. Then $f(v) = av$. It is easily checked that $f^k(v) = a^k v$ and so $m(f)(v) = m(a)v$. But $m(f) = 0$ and so $m(a)v = 0$. Since $v \neq 0$ we must have $m(a) = 0$; that is, a is a root of $m(X)$.

Suppose, conversely, that a is a root of $m(X)$. Write $m(X) = (X-a)^k p(X)$ where $X - a$ does not divide $p(X)$. Since $m(X)$ is the *minimal* polynomial, $p(f)$ is not the zero linear transformation on V. So there exists $u \in V$ such that $v = p(f)(u) \neq 0$. Observe that

$$0 = m(f)(u) = (f-a)^k\left(p(f)(u)\right) = (f-a)^k(v).$$

Choose l least so that $(f-a)^l(v) = 0$ and set $w = (f-a)^{l-1}(v)$. then $w \neq 0$ and $(f-a)(w) = 0$. Hence $f(w) = aw$; that is, a is an eigenvalue of f. □

Although the lemma tells us that the minimal polynomial has the eigenvalues as roots, it does not tell us about the multiplicity of these roots. The roots of the characteristic polynomial are also the eigenvalues, with possibly different multiplicity. We will come to the exact relationship between the two polynomials later.

The following result is an important way of constructing more invariant subspaces.

Lemma 2.2.10. *Let $f : V \to V$ be a linear transformation on a vector space V over F. Let $p(X)$ be any polynomial with coefficients in F. Then the null-space of $p(f)$ is an f-invariant subspace of V.*

Proof. See Exercise 55. □

The following lemma, although rather technical, is the key step in decomposing a linear transformation into simple pieces.

Lemma 2.2.11. *Suppose that the minimal polynomial $m(X)$ of f can be factored as a product $m(X) = p(X)q(X)$ where $p(X)$ and $q(X)$ are polynomials, with coefficients from F, which have no common factor (except constants). Then V is a direct sum of f-invariant subspaces*

$$V = W_p \oplus W_q,$$

where W_p and W_q are the nullspaces of $p(f)$ and $q(f)$ respectively. Further, the restrictions f_{W_p} and f_{W_q} have minimal polynomials $p(x)$ and $q(x)$ respectively.

Proof. We will use the following **fact** that may not be familiar to everybody: If $p(X)$ and $q(X)$ are two polynomials which have no common factor (except constants), then there are polynomials $k(X)$ and $l(X)$ such that

$$k(X)p(X) + l(X)q(X) = 1.$$

(This is the polynomial version of Theorem 1.3.8, and it can be proved in the same way: use the Euclidean algorithm for finding the greatest common factor.)

Lemma 2.2.10 tells us that W_p and W_q are f-invariant. We must show that $V = W_p \oplus W_q$.

(1) To show $W_p \cap W_q = \{0\}$: Suppose that $v \in W_p \cap W_q$. Because $v \in W_p$, we have that $p(f)(v) = 0$; because $v \in W_q$, we have that $q(f)(v) = 0$. But, we also have

$$\mathrm{id}_V = k(f)p(f) + l(f)q(f).$$

Thus

$$\begin{aligned} v &= (k(f)p(f) + l(f)q(f))\,(v) \\ &= k(f)\,(p(f)(v)) + l(f)\,(q(f)(v)) \\ &= k(f)\,(0) + l(f)\,(0) = 0 + 0 = 0 \end{aligned}$$

and so $W_p \cap W_q = \{0\}$.

(2) To show $W_p + W_q = V$: Suppose now that $v \in V$. Then

$$\mathrm{id}_V = p(f)k(f) + q(f)l(f)$$

so

$$v = p(f)\big(k(f)(v)\big) + q(f)\big(l(f)(v)\big) = p(f)(u) + q(f)(w)$$

where we have put $u = k(f)(v)$ and $w = l(f)(v)$. But $q(f)\big(p(f)(u)\big) = m(f)(u) = 0$ and so $p(f)(u) \in W_q$. Similarly, $q(f)(w) \in W_p$. Thus

$$v = p(f)(u) + q(f)(w) \in W_q + W_p$$

and so $V = W_q + W_p$. Thus $W_q \oplus W_p = V$.

(3) Restricting f to W_p we see that $p(f_{W_p}) = p(f)_{W_p} = 0$. We must show that no polynomial $p_1(X)$ of degree smaller than the degree of $p(X)$ satisfies $p_1(f_{W_p}) = 0$. As in the last paragraph we have that, for every $v \in V$, $p(f)\,(q(f)(v)) = m(f)(v) = 0$ and so $q(f)(v) \in W_p$. If $p_1(f_{W_p}) = 0$ for some polynomial $p_1(X)$ of degree smaller than that of $p(X)$, then we would have that $p_1(f)\,(q(f)(v)) = 0$ for all $v \in V$. That is, setting $m_1(X) = p_1(X)q(X)$, $m_1(f) = 0$. But since $p_1(X)$ has degree smaller than that of $p(X)$, $m_1(X)$ would have degree smaller than that of $m(X)$ which is not possible. So $p(X)$ is the minimal polynomial of f_{W_p} and, similarly, $q(X)$ is the minimal polynomial of f_{W_q}. □

Combining the previous lemmas now shows us how to choose a basis so that the matrix has 'block diagonal' form.

Theorem 2.2.12. *Let f be a linear transformation on a finite dimensional vector space. Suppose that the minimal polynomial of f takes the form $m(X) = q_1(X) \dots q_k(X)$ where $q_i(X)$ has no common factor with $q_j(X)$ if $i \neq j$. Let W_i be the nullspace of $q_i(f)$. Suppose that $\mathcal{B}_i$ is an ordered basis for W_i. Then*

$$\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_k)$$

is an ordered basis for V and the matrix of f with respect to this basis is of the form

$$\begin{bmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_k \end{bmatrix}$$

where A_i is the matrix of f_{W_i} with respect to $\mathcal{B}_i$.

Proof. We combine the result of Lemma 2.2.11 with the result of Lemma 2.2.6, using induction on k. □

The final step in our argument is to work out the simplest possibilities for the diagonal blocks A_i that are guaranteed by this last theorem.

2.2.3 Triangular form and the Cayley-Hamilton Theorem

We have seen in Theorem 2.2.12 that any linear transformation on a finite dimensional vector space can be represented by a matrix in 'block diagonal' form. We are therefore left with the problem of understanding these diagonal blocks or, equivalently, linear transformations which have a minimal polynomial which is a power of an *irreducible* polynomial (i.e. one with no lower degree factors).

We shall now **assume**, for this subsection and the next, that **the field F of scalars is algebraically closed**; for example, F could be the field $\mathbb{C}$ of complex numbers.

Then an irreducible polynomial is of the form $X - a$ for some $a \in F$ and we want to consider minimal polynomials of the type $(X - a)^m$.

Definition 2.2.13. A matrix $A = (a_{ij})$ is ***(upper) triangular*** if $a_{ij} = 0$ for all $i > j$.

Lemma 2.2.14. *Suppose that f has a minimal polynomial of the form $(X - a)^m$. Then there is a basis of V with respect to which the matrix of f is triangular. Further, $m \leq \dim V$.*

Proof. We define a set of subspaces of V as follows.

$$V_i = \{v \in V : (f - a)^i(v) = 0\}, i = 1, 2, \dots, m.$$

Note that

$$\{0\} \subseteq V_1 \subseteq V_2 \subseteq \dots \subseteq V_m = V.$$

Note also that if $v \in V_i$ then $f(v) \in V_i$ and $(f - a)(v) \in V_{i-1}$.

Choose a basis of V by starting with a basis $\mathcal{B}_1$ of V_1. Extend this to a basis $\mathcal{B}_2$ of V_2, then to a basis $\mathcal{B}_3$ of V_3 and so on until we have a basis $\mathcal{B}_m$ of $V_m = V$.

Now we look at the matrix for f using the basis $\mathcal{B}_m$. If $v \in \mathcal{B}_i$ but $v \notin \mathcal{B}_{i-1}$, then

$$f(v) = av + \sum_j^n a_j v_j \quad \text{for some } a_j \in F \quad \text{and } v_j \in \mathcal{B}_{i-1}$$

Thus the matrix A of f with respect to this basis is upper triangular with a's on the main diagonal. Note that $A - aI_n$ is upper triangular with zeroes on the diagonal. We leave as Exercise 52 the fact that $(A - aI_n)^n = 0$. Then, $(f - a)^n$ is zero and so $(X - a)^m$ divides $(X - a)^n$ by Lemma 2.2.8. That is, $m \leq n$. □

Theorem 2.2.15 (Triangular Form Theorem)**.** *Let V be a finite dimensional vector space over an algebraically closed field F and let f be a linear transformation on V. Then there is a basis of V so that the matrix of f with respect to this basis is triangular.*

Proof. This just combines Theorem 2.2.12 with Lemma 2.2.14. □

Definition 2.2.16. Let f be a linear transformation on a finite dimensional vector space V. The ***characteristic polynomial*** $c(x)$ of f is the polynomial $\det(x1_V - f)$, i.e. $\det(xI - A)$ where A is any matrix representing f.

Note that if two matrices A, B represent the same linear transformation then A, B will similar, i.e. $B = PAP^{-1}$ with P invertible. Then
$(xI - B) = P(xI - A)P^{-1}$, so $\det(xI - B) = \det(xI - A)$. Hence the resulting characteristic polynomial will not depend on the choice of matrix representing the linear transformation. Note also that this characteristic polynomial is *monic* and that $\det(xI - A) = \pm \det(A - xI)$.

Theorem 2.2.17 (Cayley-Hamilton Theorem)**.** *Let f be a linear transformation on a finite dimensional vector space. Then f satisfies its characteristic polynomial. That is, if $c(x)$ is the characteristic polynomial then $c(f)$ is the zero linear transformation.*

Proof. Choose a basis for V as described in Theorem 2.2.12. Thus the matrix A of f with respect to this basis will have the form

$$A = \begin{bmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_k \end{bmatrix}$$

where each A_i is a triangular matrix with a_i (say) on the diagonal. Suppose that A_i is $n_i \times n_i$. Then an easy calculation tells us that the characteristic polynomial of A is

$$c(X) = \prod_{i=1}^{k} (X - a_i)^{n_i}.$$

But the minimal polynomial of f is

$$m(X) = \prod_{i=1}^{k} (X - a_i)^{m_i}$$

and, by Lemma 2.2.14, $m_i \leq n_i$. Thus $m(f)$ divides $c(f)$ and so $c(f) = 0$ since $m(f) = 0$. □

Examples:

(1) Let f denote a reflection in $\mathbb{R}^3$. Then the characteristic polynomial of f is $(X - 1)^2(X + 1)$ (for example, choose a basis consisting of two vectors in the 'mirror' and one perpendicular to the mirror). The minimal polynomial is $X^2 - 1$.

(2) Let f be a linear transformation with matrix

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then the characteristic polynomial is $(X - 2)^2(X - 1)$ whereas the minimal polynomial is $(X - 2)(X - 1)$.

Applications of the Cayley-Hamilton theorem include calculation of inverses and matrix powers A^k (see the exercises).

Remark: We have been assuming in this subsection that the field F is algebraically closed and we have used this for our proof of the Cayley-Hamilton Theorem. But the Cayley-Hamilton Theorem is true over any field. Think of a matrix A, rather than a linear transformation, with coefficients over an arbitrary field K. Then its characteristic polynomial has coefficients in K. There will be an algebraically closed field F containing K and we can regard A as a matrix over F. The Cayley-Hamilton theorem over F will then apply to show that A satisfies its characteristic polynomial.

2.2.4 Jordan normal form

Triangular form is a practical way to represent a linear transformation but is not the best possible way. For this we need the *Jordan normal form* (JNF) — also known as the *Jordan canonical form* (JCF). We shall describe this but we shall not attempt to prove the result or to show how to calculate Jordan normal form in general.

Definition 2.2.18. The $n \times n$ ***Jordan block*** matrix $J(a, n)$ (for $a \in F$) is the $n \times n$ matrix

$$A = \begin{bmatrix} a & 1 & 0 & \dots & 0 \\ 0 & a & 1 & \dots & 0 \\ \vdots & \dots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a & 1 \\ 0 & \dots & 0 & 0 & a \end{bmatrix}.$$

A straightforward calculation shows that the characteristic polynomial and the minimal polynomial of $J(a, n)$ are both $(X - a)^n$.

Examples:

The matrices

$$J(2,3) = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \quad J(0,4) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad J(4,1) = [4]$$

are all Jordan blocks.

Recall that we are assuming in this section that the field of scalars F is **algebraically closed** (e.g. $F = \mathbb{C}$.)

Theorem 2.2.19 (Jordan Normal Form)**.** *Let f be a linear transformation on a finite dimensional vector space V. Then there is a basis of V with respect to which the matrix of f takes the form*

$$\begin{bmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_k \end{bmatrix}$$

where each A_i is a Jordan block matrix. This expression for the matrix of f is unique up to re-ordering the diagonal blocks $A_1, \dots, A_k$.

Equivalently, every matrix is similar to a matrix in Jordan normal form, and this Jordan form is unique up to re-ordering the Jordan blocks.

Question: How can we determine the Jordan normal form?

There is a fairly easy observation which gives us information about the JNF just by knowing the characteristic and minimal polynomials.

Lemma 2.2.20. *Suppose that a is an eigenvalue of f and let $(X - a)^m$ be the highest power of $X - a$ dividing the minimal polynomial and let $(X - a)^n$ be the highest power of $X - a$ dividing the characteristic polynomial. Then m is the size of the largest Jordan block $J(a, l)$ that occurs in the JNF of f and n is the sum of the sizes of the Jordan blocks $J(a, l)$ that occur in the JNF of f.*

Proof. The claim about the characteristic polynomial is clear. A little computation is needed to check the claim about the minimal polynomial. □

Theorem 2.2.21. *A linear transformation f of a finite dimensional vector space V can be represented by a diagonal matrix (with respect to a suitable basis) if and only if the minimal polynomial of f has no repeated roots. In particular, if* $\dim V = n$ *and f has n distinct eigenvalues then f can be represented by a diagonal matrix.*

Proof. The first thing to appreciate is that the uniqueness of the JNF guarantees that if f can be represented by a diagonal matrix then that matrix is the JNF of f. A JNF is diagonal if and only if each Jordan block which occurs has size 1. By the previous lemma, this happens if and only if each eigenvalue occurs as a root of the minimal polynomial with multiplicity 1. □

Examples:

(1) The matrices

$$\begin{bmatrix}3&1&0\\0&3&0\\0&0&2\end{bmatrix}\quad\begin{bmatrix}3&1&0\\0&3&1\\0&0&3\end{bmatrix}\quad\begin{bmatrix}3&1&0&0\\0&3&1&0\\0&0&3&0\\0&0&0&2\end{bmatrix}$$

are all in JNF.

(2) The matrix

$$A=\begin{bmatrix}2&1&0&0\\0&3&0&0\\0&0&2&0\\0&0&-2&1\end{bmatrix}$$

has characteristic polynomial $(X-1)(X-2)^2(X-3)$. Thus its minimal polynomial is either $(X-1)(X-2)^2(X-3)$ or $(X-1)(X-2)(X-3)$. Direct computation shows that $(A-I)(A-2I)(A-3I)=0$. Hence its minimal polynomial is $(X-1)(X-2)(X-3)$ and so its JNF is

$$\begin{bmatrix}1&0&0&0\\0&2&0&0\\0&0&2&0\\0&0&0&3\end{bmatrix}.$$

For matrices of small size, we can gain much information from the minimal and characteristic polynomials. For example, for a 2×2 matrix the characteristic polynomial is $(X-a)(X-b)$ for some $a,b\in F$. The minimal polynomial then divides $(X-a)(X-b)$ and has the same roots. If $a\neq b$ then the minimal polynomial is also $(X-a)(X-b)$ and the JNF is $\begin{bmatrix}a&0\\0&b\end{bmatrix}$. If $a=b$ then the minimal polynomial is either $(X-a)^2$ in which case the JNF is $\begin{bmatrix}a&1\\0&a\end{bmatrix}$ or $(X-a)$ in which case the JNF is $\begin{bmatrix}a&0\\0&a\end{bmatrix}$.

We can do a similar analysis for 3×3 matrices. In the following, a,b,c are assumed to be different elements of F and each row represents a different type of possibility for the characteristic polynomial, the minimal polynomial and the JNF.

Characteristic	Minimal	JNF
$(X-a)(X-b)(X-c)$	$(X-a)(X-b)(X-c)$	$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$
$(X-a)^2(X-b)$	$(X-a)^2(X-b)$	$\begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$
$(X-a)^2(X-b)$	$(X-a)(X-b)$	$\begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$
$(X-a)^3$	$(X-a)^3$	$\begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}$
$(X-a)^3$	$(X-a)^2$	$\begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix}$
$(X-a)^3$	$(X-a)$	$\begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix}$

Knowing the *dimensions of eigenspaces* of a matrix A also helps in determining the Jordan normal form. It's easy to check that:

(1) For each Jordan block in the JNF there is exactly *one* linearly independent eigenvector.

(2) the dimension of the eigenspace for an eigenvalue λ is the number of Jordan blocks with λ on the diagonal.

Example: Find the Jordan normal form for

$$A = \begin{bmatrix} 2 & 2 & -1 \\ -1 & -1 & 1 \\ -1 & -2 & 2 \end{bmatrix}.$$

The characteristic polynomial is $c(x) = (x-1)^3$ so there is only one eigenvalue, $\lambda = 1$. Using row reduction, we find the corresponding eigenspace Nullspace$(A - I)$ has dimension 2. Thus the Jordan normal form J has 2 blocks, hence

$$J = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2.2.5 Exercises

48. Find the minimal polynomials of the matrices:

$$\begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}.$$

49. Show that the matrices

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

have the same minimal polynomial. Do they have the same characteristic polynomial?

50. Show that the matrix

$$A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$$

has minimal polynomial $X^2 - 2X - 8$. Use this to determine the inverse of A.

51. Show that a linear transformation f is invertible if and only if its minimal polynomial has non-zero constant term. Assuming f is invertible, how can the inverse be calculated if the minimal polynomial is known?

52. Suppose that A is an $n \times n$ upper triangular matrix with zeroes on the diagonal. Prove that $A^n = 0$.

53. Let f be a linear transformation on a vector space V with minimal polynomial $X^2 - 1$ and suppose that $2 \neq 0$ in the field of scalars. (Thus, for example, $\mathbb{F}_2$ is not allowed as the field of scalars.) Show directly that the subspaces $\{v \in V : f(v) = v\}$ and $\{v \in V : f(v) = -v\}$ are complementary subspaces of V. Find a diagonal matrix representing f.

54. Prove Lemma 2.2.6.

55. Prove Lemma 2.2.10.

56. Show that the linear transformation $\mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_n(\mathbb{R})$ given by differentiation cannot be represented by a diagonal matrix.

57. If f is a linear transformation on a finite dimensional vector space V satisfying $f^2 = f$, explain how to find a diagonal matrix representing f.

58. Suppose that linear transformations f and g on a vector space V commute; that is, that $fg = gf$. Show that an eigenspace of f will be g-invariant. If the field F of scalars is algebraically closed and V is finite dimensional, deduce that f and g have a common eigenvector.

59. Find the Jordan normal form of the following matrices:

$$\begin{bmatrix} -1 & 1 \\ -1 & -3 \end{bmatrix}, \begin{bmatrix} -1 & 3 & 0 \\ 0 & 2 & 0 \\ 2 & 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}.$$

60. For each of the following pairs of minimal and characteristic polynomials, find all possibilities for the Jordan normal form:

Minimal polynomial	Characteristic polynomial
$X^2(X+1)^2$	$X^2(X+1)^4$
$(X-3)^2$	$(X-3)^5$
X^3	X^7
$(X-1)^2(X+1)^2$	$(X-1)^4(X+1)^4$.

61. Which of the following pairs of matrices (over $\mathbb{C}$) are similar?

(a) $\begin{bmatrix} -1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix};$

(b) $\begin{bmatrix} -1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 5 \\ 0 & -1 \end{bmatrix};$

(c) $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$

62. Given a 4×4 matrix A over $\mathbb{C}$ and given the minimal and characteristic polynomials of A, describe the possibilities for the JNF of A. (There will be one case where there are two possibilities.)

63. Show that any JNF matrix J is a sum $J = D + N$ where D is diagonal and N is nilpotent; that is $N^k = 0$ for some k. Deduce that any linear transformation f of a finite dimensional complex vector space can be written in the form $f = d + n$ where d is diagonalisable and n is nilpotent.

64. In the language of the previous question, show that $JN = NJ$ and $JD = DJ$. Deduce that $fd = df$ and $fn = nf$.

65. (Harder) Show that the Jordan normal form of a complex matrix A is completely determined by the dimensions of the nullspaces of $(A - \lambda I)^i$, $i = 1, 2, 3, \dots$ for all the eigenvalues λ of A.

2.3 Inner product spaces

Inner products are generalizations of the *dot product* in $\mathbb{R}^3$ or $\mathbb{R}^n$. They give a way to introduce *geometry* in a vector space. You have already seen some of the properties of inner products on real vector spaces. Inner products on complex vector spaces have many of the same properties but there are sufficient differences that we need to cover the topic from scratch.

In this chapter, the **field F will be either the real numbers or the complex numbers**. Each has an absolute value function $F \to \mathbb{R}_+$ and each has a 'conjugate function' $F \to F$. For the complex numbers, this is the usual complex conjugation and for the real numbers it is simply the identity function.

2.3.1 Complex inner products

Definition 2.3.1. An ***inner product*** in a vector space V over F is a function $V \times V \to F$ satisfying the following (we write the value of the function as (v, w) where $v, w \in V$ and $(v, w) \in F$):

(1) $(v, w) = \overline{(w, v)}$ for all $v, w \in V$;

(2) $(au + bv, w) = a(u, w) + b(v, w)$ for all $u, v, w \in V$ and $a, b \in F$;

(3) $(v, v) \geq 0$ for all $v \in V$ and $(v, v) = 0$ if and only if $v = 0$.

(Here and later, when we say $(v, v) \geq 0$ we will mean '(v, v) is real and non-negative'.)

A real inner product space is often called a *Euclidean space* and a complex inner product space is often called a *unitary* space.

Note: (1) Taking $v = w$ in condition (1) gives $(v, v) = \overline{(v, v)}$; hence (v, v) is always *real.*

(2) Conditions (1) and (2) imply that $(w, au + bv) = \overline{a}(w, u) + \overline{b}(w, v)$.

Examples:

(1) Set $V = \mathbb{R}^n$ and define

$$((a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n)) = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

This gives an inner product, the standard 'dot product'.

(2) Set $V = \mathbb{C}^n$ and define

$$((a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n)) = a_1\overline{b_1} + a_2\overline{b_2} + \cdots + a_n\overline{b_n}.$$

This again gives an inner product, the 'complex dot product'.

(3) Let V be any n-dimensional real space and let $\{v_1, \dots, v_n\}$ be a basis of V. Define an inner product on V by

$$(a_1v_1 + \cdots + a_nv_n, b_1v_1 + \cdots + b_nv_n) = a_1b_1 + \cdots + a_nb_n.$$

This gives an inner product; it is not hard to see that Example 1 is a special case of this.

(4) Let V be any n-dimensional complex space and let $\{v_1, \dots, v_n\}$ be a basis of V. Define an inner product on V by

$$(a_1v_1 + \cdots + a_nv_n, b_1v_1 + \cdots + b_nv_n) = a_1\overline{b_1} + a_2\overline{b_2} + \cdots + a_n\overline{b_n}.$$

This gives an inner product; again, Example 2 is a special case of this.

(5) Let $V = M_{n\times n}(F)$. Define an inner product by

$$(A, B) = \operatorname{trace}(A\overline{B}^T)$$

where trace(C), for a square matrix C, is the sum of the diagonal entries.

(6) Recall that $\mathcal{P}(F)$ denotes the space of all polynomials with coefficients from $F = \mathbb{R}$ or $\mathbb{C}$. We can define an inner product on $\mathcal{P}(F)$ by

$$(p(x), q(x)) = \int_0^1 p(x)\overline{q(x)}dx.$$

(7) Let $V = C([a, b], F)$ be the vector space of all continuous functions $f : [a, b] \to F$ where $[a, b]$ is the closed interval $\{t : a \leq t \leq b\}$. Then we can define an inner product on V by

$$(f, g) = \int_a^b f(t)\overline{g(t)}\,dt.$$

Definition 2.3.2. If V is an inner product space:

(1) The ***length*** of a vector $v \in V$ is $\|v\| = \sqrt{(v, v)}$.

(2) Two elements $v, w \in V$ are ***orthogonal*** if $(v, w) = 0$.

(3) A subset S of V is said to be ***orthonormal*** if $v, w \in S$ implies that

$$(v, w) = \begin{cases} 0 & \text{if } v \neq w \\ 1 & \text{if } v = w \end{cases}$$

Theorem 2.3.3. *Let W be a finite dimensional inner product space. Then*

(1) any orthonormal set in W is linearly independent;

(2) any orthonormal set in W can be extended to an orthonormal basis.

Proof. We leave the proof of the first part as an exercise. The second part is the Gram-Schmidt orthogonalisation process which most of you will have seen for the space $\mathbb{R}^n$.

We sketch the argument. Let $\mathcal{O} = \{v_1, \ldots, v_m\}$ be an orthonormal set. By the first part it is linearly independent and so can be extended to a basis $\mathcal{O} \cup \mathcal{B}$. Let $w \in \mathcal{B}$ and set $w' = w - \sum_{i=1}^m (w, v_i)v_i$. It is easily checked that $(w', v_i) = 0$ for $i = 1, \ldots, m$. Thus $\mathcal{O} \cup \{w'/\|w'\|\}$ will be an orthonormal set strictly containing $\mathcal{O}$. Now repeat the process until you have a basis. □

Examples:

(1) We can extend the orthonormal set $\{\frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{3}}(1, -1, 1)\}$ of $\mathbb{R}^3$ by adding the vector $w = (0, 0, 1)$ to form a basis and then forming

$$\begin{aligned} w' &= (0, 0, 1) - 0.\frac{1}{2}(1, 1, 0) - 1.\frac{1}{3}(1, -1, 1) \\ &= \frac{1}{3}(-1, 1, 2). \end{aligned}$$

So $\{\frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{3}}(1, -1, 1), w'/\|w'\| = \frac{1}{\sqrt{6}}(-1, 1, 2)\}$ will be an orthonormal basis of $\mathbb{R}^3$.

(2) The set $\{1\}$ is an orthonormal set in $\mathcal{P}_1(\mathbb{R})$. If we extend it to a basis and apply the Gram-Schmidt orthogonalisation process, we obtain an orthonormal basis $\{1, \sqrt{3}(2x - 1)\}$.

(3) The set

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

is an orthonormal basis of $M_{2\times 2}(\mathbb{R})$.

Theorem 2.3.4 (Bessel's inequality)**.** *Let $S = \{v_1, \dots, v_n\}$ be an orthonormal subset of an inner product space V. Let $v \in V$ and set $a_i = (v, v_i)$ for $i = 1, \dots, n$. Then*

$$\sum_{i=1}^{n} |a_i|^2 \leq \|v\|^2.$$

The vector $v - \sum_{i=1}^{n} (v, v_i) v_i$ is orthogonal to each v_j. In particular, if S is a basis of V, then

$$v = \sum_{i=1}^{n} a_i v_i$$

and

$$\sum_{i=1}^{n} |a_i|^2 = \|v\|^2.$$

Proof. We have

$$\begin{aligned} 0 &\leq \|v - \sum_{i=1}^{n} a_i v_i\|^2 = (v - \sum_{i=1}^{n} a_i v_i, v - \sum_{i=1}^{n} a_i v_i) \\ &= (v, v) - \sum_{i=1}^{n} a_i (v_i, v) - \sum_{i=1}^{n} \overline{a_i} (v, v_i) + \sum_{i,j=1}^{n} a_i \overline{a_j} (v_i, v_j) \\ &= (v, v) - \sum_{i=1}^{n} |a_i|^2 - \sum_{i=1}^{n} |a_i|^2 + \sum_{i=1}^{n} |a_i|^2 \\ &= (v, v) - \sum_{i=1}^{n} |a_i|^2. \end{aligned}$$

which completes the proof of the first part.

For the second part, note that

$$(v - \sum_{i=1}^{n} a_i v_i, v_j) = (v, v_j) - \sum_{i=1}^{n} a_i (v_i, v_j) = (v, v_j) - a_j = 0.$$

Finally, if S is a basis of V then $v - \sum_{i=1}^{n} a_i v_i$ will be orthogonal to every element of a basis and so must be zero. Further, equality holds in the first inequality. □

Example: Let $\mathcal{F}$ be the subspace of $\mathbb{R}^{\mathbb{R}}$ consisting of functions which are periodic of period 2π and which are 'sufficiently good to have a convergent Fourier series'. We could require the functions to be twice differentiable, for example, but weaker conditions will also do. Define an inner product on $\mathcal{F}$ by

$$(f, g) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) g(t) dt.$$

The functions $\cos nt$ and $\sin nt$ will certainly lie in $\mathcal{F}$ and

$$\{\frac{1}{\sqrt{2}}, \cos t, \sin t, \cos 2t, \sin 2t, \dots\}$$

is an infinite orthonormal set in $\mathcal{F}$.

It is **not** true that this orthonormal set is a basis of $\mathcal{F}$ but the theory of Fourier series tells us that every function f in $\mathcal{F}$ can be expressed in the form

$$\begin{aligned} f = (f, \frac{1}{\sqrt{2}}) \frac{1}{\sqrt{2}} &+ (f, \sin t) \sin t + (f, \cos t) \cos t + \dots + \\ &+ (f, \sin nt) \sin nt + (f, \cos nt) \cos nt + \dots. \end{aligned} \tag{2.2}$$

This expression is an infinite series and so does not fit into our theory of vector spaces. But the expression here is very similar to that in the last sentence of the previous theorem and suggests that there is some generalisation of the current theory into which we could fit the theory of Fourier series.

Lemma 2.3.5 (Schwarz's inequality)**.** *If v, w are elements of an inner product space V then*

$$|(v,w)| \leq \|v\|.\|w\|.$$

Proof. If $w = 0$ then both sides are zero. If not, then $\{w/\|w\|\}$ is an orthonormal set and the result follows directly from Bessel's inequality. □

If V is a real inner product space, this allows us to define the ***angle*** θ between two non-zero vectors v, w in V by

$$\cos\theta = \frac{(v,w)}{\|v\| \, .\|w\|}, \text{ and } 0 \leq \theta \leq \pi.$$

Examples:

(1) If we take $V = \mathbb{R}^n$ and the dot product, this becomes

$$|\sum_{i=1}^{n} a_i b_i| \leq (\sum_{i=1}^{n} a_i^2)^{\frac{1}{2}} (\sum_{i=1}^{n} b_i^2)^{\frac{1}{2}}$$

for any real numbers a_i, b_i.

(2) If we take $V = \mathbb{C}^n$ and the complex dot product, we have

$$|\sum_{i=1}^{n} a_i \overline{b_i}| \leq (\sum_{i=1}^{n} |a_i|^2)^{\frac{1}{2}} (\sum_{i=1}^{n} |b_i|^2)^{\frac{1}{2}}$$

for any complex numbers a_i, b_i. Note that, by replacing b_i by $\overline{b_i}$, we can obtain a complex inequality identical to the previous example.

(3) If we take the inner product space of Example 7 above, then we have

$$\left|\int_a^b f(t)\overline{g(t)}\,dt\right| \leq \left(\int_a^b f(t)^2\,dt\right)^{\frac{1}{2}} \left(\int_a^b g(t)^2\,dt\right)^{\frac{1}{2}}.$$

Schwarz's inequality also allows us to define ***distance*** on an inner product space. If $u, v \in V$ then we can define the distance between u and v to be $\delta(u,v) = \|u - v\|$. Clearly, $\delta(u,u) \geq 0$ and $\delta(u,v) = \delta(v,u)$. The other major requirement of a 'distance function' is that $\delta(u,v) \leq \delta(u,w) + \delta(w,v)$, the 'triangle inequality'.

To see the latter, we will show that $\|x+y\|^2 \leq (\|x\| + \|y\|)^2$ for any $x, y \in V$. Then if we take square roots and replace x by $u - w$ and y by $w - v$ we will have the triangle inequality.

So

$$\begin{aligned}
\|x+y\|^2 &= (x+y, x+y) = (x,x) + (x,y) + (y,x) + (y,y) \\
&= \|x\|^2 + (x,y) + \overline{(x,y)} + \|y\|^2 \\
&= \|x\|^2 + 2\mathrm{Re}\,((x,y)) + \|y\|^2 \\
&\leq \|x\|^2 + 2|(x,y)| + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2 .
\end{aligned}$$

2.3.2 Orthogonal complements

Definition 2.3.6. Let V be an inner product space and let W be a subspace of V. The ***orthogonal complement*** $W^{\perp}$ of W in V is

$$W^{\perp} = \{v \in V : (v,w) = 0 \text{ for all } w \in W\}.$$

Lemma 2.3.7. *Let V be a finite dimensional inner product space and let W be a subspace of V. Then $W^\perp$ is a subspace of V, $W^\perp \cap W = \{0\}$ and $V = W + W^\perp$. Thus, V is a direct sum $V = W \oplus W^\perp$ and $\dim W^\perp = \dim V - \dim W$.*

Proof. It is an easy check that $W^\perp$ is a subspace of V. If $w \in W^\perp$ then $(w, W) = 0$. If also, $w \in W$, then $(w, w) = 0$ and so $w = 0$; that is, $W^\perp \cap W = \{0\}$.

Choose an orthonormal basis $\mathcal{B}_1$ for W and extend this to an orthonormal basis $\mathcal{B}_1 \cup \mathcal{B}_2$ for V. Because $\mathcal{B}_1 \cup \mathcal{B}_2$ is orthonormal, each vector in $\mathcal{B}_2$ is orthogonal to each vector in $\mathcal{B}_1$ and so $\mathcal{B}_2 \subseteq W^\perp$. Every element v of V can be written as a sum $v = v_1 + v_2$ with $v_i \in \langle \mathcal{B}_i \rangle$. But then $v_1 \in W$ and $v_2 \in W^\perp$. Thus $V = W + W^\perp$.

The final sentence follows from Lemma 2.2.5. □

Examples:

(1) The orthogonal complement to a plane through the origin in $\mathbb{R}^3$ is the normal through the origin.

(2) The orthogonal complement to a line through the origin in $\mathbb{R}^3$ is the plane through the origin to which it is normal.

(3) The orthogonal complement to the set of diagonal matrices in $M_{n\times n}(\mathbb{R})$ is the set of matrices with zero entries on the diagonal.

(4) If A is an $m \times n$ matrix with real coefficients then the *row space* of A is the orthogonal complement of the *nullspace* of A. (See Exercise 71.)

2.3.3 Adjoints, self-adjoint, Hermitian, normal

Definition 2.3.8. Let $f\colon V \to V$ be a linear transformation on an inner product space V. The ***adjoint*** f^* of f is a linear transformation $f^*\colon V \to V$ satisfying

$$(f(v), w) = (v, f^*(w)) \quad \text{for all } v, w \in V. \tag{$*$}$$

Lemma 2.3.9. *If V is finite dimensional then the adjoint f^* exists and is unique.*

Idea of proof. To see that $(*)$ really does define a function f^* we need to show that, given w there is a unique $w_1 \in V$ such that $(f(v), w) = (v, w_1)$ for all $v \in V$; we can then set $f^*(w) = w_1$. If V is finite-dimensional, this is always possible; the proof is left as Exercise 74. Once this is done, we know that f^* is a well-defined function. It is then not hard to see that f^* is a linear transformation; the proof is left as Exercise 75. □

Note: An adjoint f^* does not always exist if V is infinite dimensional (see Exercise 72).

The next lemma is often useful for working with adjoints.

Lemma 2.3.10. *If $f, g : V \to V$ are linear transformations on an inner product space V satisfying*

$$(f(v), w) = (g(v), w) \textit{ for all } v, w \in V$$

then $f = g$.

Proof. We have

$$(f(v) - g(v), w) = 0 \text{ for all } v, w \in V.$$

Taking $w = f(v) - g(v)$ gives

$$(f(v) - g(v), f(v) - g(v)) = 0 \text{ for all } v \in V.$$

Hence $f(v) - g(v) = 0$ for all $v \in V$ by the positivity of inner products. So $f(v) = g(v)$ for all $v \in V$ and $f = g$. □

Some properties of adjoints: If $f, g : V \to V$ are linear transformations on a finite dimensional inner product space V and $\alpha \in \mathbb{C}$ then

(1) $(f+g)^* = f^* + g^*$,

(2) $(\alpha f)^* = \overline{\alpha} f^*$,

(3) $(fg)^* = g^* f^*$ (where fg denotes the composition of $f \circ g$),

(4) $(f^*)^* = f$

These follow easily from the definition of adjoint and the previous lemma (see Exercises 68 and 76).

If we have a matrix for f, then what is the matrix of f^*? To get a nice answer we need to choose the matrix with respect to an orthonormal basis.

Lemma 2.3.11. *Let V be an inner product space with an orthonormal basis $\mathcal{B} = \{v_1, \ldots, v_n\}$. Suppose that a linear transformation f has a matrix A with respect to $\mathcal{B}$. Then the matrix A^* of f^* with respect to $\mathcal{B}$ is given by*

$$(A^*)_{ij} = \overline{A_{ji}}\,;$$

that is, A^* is the 'complex conjugate transpose' of A.

Proof. Suppose that $A = (a_{ij})$ and $A^* = (b_{ij})$. Then for any i, j,

$$(f(v_i), v_j) = (\sum_k a_{ki} v_k, v_j) = \sum_k a_{ki} (v_k, v_j) = a_{ji}$$

and

$$(v_i, f^*(v_j)) = (v_i, \sum_k b_{kj} v_k) = \sum_k \overline{b_{kj}} (v_i, v_k) = \overline{b_{ij}}$$

and so $b_{ij} = \overline{a_{ji}}$ as required. □

Definition 2.3.12. If A is an $n \times n$ matrix over F then A^* is the matrix defined by $(A^*)_{ij} = \overline{A_{ji}}$.

We also call A^* the adjoint of A.

Next we look at some important kinds of linear transformations

Definition 2.3.13. Let f be a linear transformation on an inner product space V.

(1) We say that f is ***self-adjoint*** if $f = f^*$. If V is real, this is often called ***symmetric***. If V is complex, it is called ***Hermitian***.

(2) We say that f is an ***isometry*** if $f^* f = 1_V$ where 1_V is the identity transformation on V. If V is real, this is often called ***orthogonal***. If V is complex, it is called ***unitary***.

(3) We say that f is ***normal*** if $f f^* = f^* f$.

Similar terminology applies to matrices.

Examples:

(1) If a linear transformation is represented by a symmetric matrix with respect to an orthonormal basis, then it is self-adjoint. For example $\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$ is self-adjoint.

(2) A rotation of $\mathbb{R}^n$ is orthogonal.

(3) The matrix $\begin{bmatrix} 1 & 2-i \\ 2+i & 3 \end{bmatrix}$ is self-adjoint (Hermitian).

(4) Skew-symmetric real matrices (that is, matrices A that satisfy $A = -A^*$) are normal but not self-adjoint.

(5) The matrix $\begin{bmatrix} 1 & 1 \\ i & 3+2i \end{bmatrix}$ is normal but not self-adjoint or skew-symmetric or orthogonal.

Isometries are so called because they are *distance preserving* transformations.

Lemma 2.3.14. *Let f be a linear transformation on an inner product space V. The following are equivalent.*

(1) $f^*f = 1_V$*;*

(2) $(f(u), f(v)) = (u, v)$ *for all* $u, v \in V$*;*

(3) $\|f(v)\| = \|v\|$ *for all* $v \in V$.

Note: If V is finite dimensional, (1) means that $f^* = f^{-1}$.

Proof. If (1) is true and if $u, v \in V$ then

$$(f(u), f(v)) = (u, f^*(f(v))) = (u, 1_V(v)) = (u, v)$$

and so (2) is true. Set $u = v$ and take square roots to obtain (3) from (2). Assume (3) and consider the linear transformation $g = f^*f - 1_V$. Then

$$g^* = (f^*f - 1_V)^* = (f^*f)^* - 1_V = f^*f^{**} - 1_V = f^*f - 1_V = g$$

and so g is self-adjoint. For all $v \in V$,

$$(g(v), v) = (f^*(f(v)) - v, v) = (f(v), f(v)) - (v, v) = \|f(v)\|^2 - \|v\|^2 = 0.$$

We leave it as Exercise 77 to show that this means that g is the zero linear transformation and so $f^*f = 1_V$. □

Lemma 2.3.15. *Let W be an f-invariant subspace of V. Then $W^\perp$ is f^*-invariant.*

Proof. Suppose that $u \in W^\perp$. If $w \in W$ it follows that $f(w) \in W$ because W is f-invariant and so

$$(w, f^*(u)) = (f(w), u) = 0 \text{ as } u \in W^\perp.$$

Hence $f^*(u) \in W^\perp$ and so $W^\perp$ is f^*-invariant. □

Lemma 2.3.16. *Let f be a linear transformation over a finite dimensional real vector space V. Then V has an f-invariant subspace of dimension at most 2.*

Proof. Consider the minimal polynomial $m(X)$ of f. Since the field of scalars is the real numbers, any irreducible factor $p(X)$ of $m(X)$ has degree at most 2. (If you have not seen this fact before it may require some thought; remember that the polynomial factors completely over the complex numbers and that complex roots of a real polynomial occur in complex conjugate pairs.) Set $m(X) = p(X)q(X)$ and set $W = \{q(f)(v) : v \in V\}$. Then W is a non-zero subspace (otherwise we would have $q(f) = 0$ on V) and $p(f)(w) = 0$ for any $w \in W$.

Choose any $w \in W$ with $w \neq 0$. If $p(X) = X - a$ for some $a \in \mathbb{R}$ then $f(w) = aw$ and so the subspace $\langle w \rangle$ is 1-dimensional and invariant. If $p(X) = X^2 + aX + b$ for some $a, b \in \mathbb{R}$ then the subspace $\langle w, f(w) \rangle$ is 2-dimensional and invariant. □

Theorem 2.3.17. *Let f be an orthogonal linear transformation over a finite dimensional real vector space V. Then there is an orthonormal basis of V of the form*

$$\{u_1, v_1, u_2, v_2, \dots u_k, v_k, w_1, \dots w_l\}$$

so that, for some $\theta_1, \dots, \theta_k$,

$$f(u_i) = (\cos\theta_i)u_i + (\sin\theta_i)v_i \quad \textit{and} \quad f(v_i) = -\sin(\theta_i)u_i + (\cos\theta_i)v_i$$

and $f(w_i) = \pm w_i$.

Proof. We shall prove this by induction on the dimension of V. By Lemma 2.3.16, V has an f-invariant subspace W which is 1 or 2 dimensional. By Lemma 2.3.15, $W^{\perp}$ is f^*-invariant. Since f is orthogonal, $f^* = f^{-1}$ and so $W^{\perp}$ is f^{-1}-invariant; that is, $f^{-1}(W^{\perp}) = W^{\perp}$ and so $f(W^{\perp}) = W^{\perp}$. That is, $W^{\perp}$ is f-invariant.

We leave as Exercise 79 that the two linear transformations obtained by restriction, f_W and $f_{W^{\perp}}$, are still orthogonal. Since W has dimension either 1 or 2 then it has an orthonormal basis of the kind described (we leave this as Exercise 80). Since the dimension of $W^{\perp}$ is less than that of V, we can apply the inductive hypothesis to deduce that $W^{\perp}$ has a basis of the kind described. Combining these two bases (possibly with some re-ordering) we have a basis for V as required in the statement of the theorem. □

Examples:

(1) In dimension 2, the possibilities for orthogonal matrices *up to similarity* are

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

for some θ. The first represents a reflection, the last represents a rotation through θ.

(2) In dimension 3, there must be a real eigenvalue because the characteristic polynomial has degree 3 and real polynomials of odd degree must have a real root. We can summarise the possibilities as follows:

three eigenvalues equal to 1: the identity;

two eigenvalues equal to 1 and one equal to -1: a reflection;

one eigenvalue equal to 1, the other two either both -1 or complex: a rotation

one eigenvalue equal to -1, the other two either both -1 or complex: an 'improper rotation', i.e. the product of a rotation and a reflection.

2.3.4 Exercises

66. Find the length of

(a) $(2+i, 3-2i, -1)$ in the inner product space of Example 2.

(b) $x^2 - 3x + 1$ in the inner product space of Example 6.

(c) $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ in the inner product space of Example 5.

67. An exercise (from an anonymous textbook) claims that, for all elements u, v of an inner product space, $\|u+v\| + \|u-v\| = 2\|u\| + 2\|v\|$. Prove that this is false. Can you guess what was intended?

68. If f and g are linear transformations on an inner product space V, show that

$$(f+g)^* = f^* + g^* \qquad \text{and} \qquad (fg)^* = g^* f^*.$$

69. If A is a transition matrix between orthonormal bases, show that A is isometric.

70. Suppose that f is a linear transformation on a finite dimensional inner product space V.

(a) If f is self-adjoint, show that the eigenvalues of f are real;

(b) if f is isometric, show that the eigenvalues of f have absolute value 1.

71. Suppose that f is a linear transformation on a finite dimensional inner product space V. Show that the range of f^* is the orthogonal complement of the nullspace of f. Deduce that the rank of f is equal to the rank of f^*. Deduce that the row-rank of a square matrix is equal to its column rank.

72. (Harder) Show that the function δ of differentiation on the inner product space of Example 6 (after Definition 2.3.1) has no adjoint. (Hint: Try to find what $\delta^*(1)$ should be.)

73. Show that a triangular matrix which is self-adjoint or unitary is diagonal.

74. Let V be a finite dimensional inner product space and f a linear transformation on V. Show that, given a vector $w \in V$, there exists a unique vector $w_1 \in V$ such that $(f(v), w) = (v, w_1)$ for all $v \in V$. (Hint: First show that it will be enough to consider only those v which lie in some fixed orthonormal basis of V.)

75. Deduce that Definition 2.3.8 does define a linear transformation.

76. Let f be a linear transformation on a finite dimensional inner product space V. Show, without using matrices, that $(f^*)^* = f$.

77. Let g be a self-adjoint linear transformation on a finite dimensional inner product space V. Suppose that $(g(v), v) = 0$ for all $v \in V$.

(a) Show that $(g(u), w) + (g(w), u) = 0$ for all $u, w \in V$ (replace v by $u + w$).

(b) Deduce that g is the zero linear transformation if the space is a real space. (This is the time to use the fact that g is self-adjoint).

(c) Assume now that the space is complex; deduce that $(g(u), w)$ is imaginary for all $u, w \in V$.

(d) Deduce that $(g(iu), w)$ is imaginary for all $u, w \in V$ and so that $(g(u), w) = 0$ for all $u, w \in V$.

(e) Deduce that g is zero in the complex case also.

78. Let f be a linear transformation on a finite dimensional inner product space V. Suppose that W is an f-invariant and f^*-invariant subspace of V. Show that $(f_W)^* = (f^*)_W$.

79. Let f be an isometry on a finite dimensional inner product space V. Suppose that W is an f-invariant subspace of V. Show that f_W is also an isometry.

80. Let V be a two dimensional real inner product space and let f be an isometry of V. Show that f can be represented by a matrix of the form:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \epsilon\sin\theta & \epsilon\cos\theta \end{bmatrix}$$

where $\epsilon = \pm 1$.

2.4 The Spectral Theorem and applications

The main result of this section shows that for the important class of *normal* linear transformations we can find a basis made up of orthonormal eigenvectors.

2.4.1 The Theorem

Theorem 2.4.1 (Spectral Theorem; first version)**.** *Let f be a normal linear transformation on a finite dimensional complex inner product space V. Then there is an orthonormal basis for V such that the matrix of f with respect to this basis is diagonal.*

We can also state this in terms of matrices: Let A be a normal matrix. Then there exists a unitary matrix U such that $U^{-1}AU = U^*AU = D$ is diagonal.

Before we prove the spectral theorem we need the following preliminary result.

Lemma 2.4.2. *Let f be a normal linear transformation on a finite dimensional complex inner product space V. Then there is a non-zero element of V which is an eigenvector for both f and f^*. The two corresponding eigenvalues are complex conjugates.*

Proof. Firstly, let a be an eigenvalue of f and let V_a be the subspace of eigenvectors corresponding to a. It is easily checked that V_a is f-invariant. We claim that it is also f^*-invariant. If $v \in V_a$ then $f(f^*(v)) = f^*(f(v))$, as f is normal, and so $f(f^*(v)) = f^*(av) = af^*(v)$ as $v \in V_a$. Thus $f^*(v)$ is also an a-eigenvector of f and so $f^*(v) \in V_a$. Hence V_a is f^*-invariant.

Thus we can consider $f^*_{V_a}$. This will have an eigenvector in V_a. Call this eigenvector w and suppose that b is the corresponding eigenvalue. Thus $f^*(w) = bw$. Also, as $w \in V_a$, $f(w) = aw$. Note that

$$a(w, w) = (aw, w) = (f(w), w) = (w, f^*(w)) = (w, bw) = \bar{b}(w, w)$$

and so $a = \bar{b}$. □

More generally, let $f, g : V \to V$ be two linear transformations on a complex vector space V. If f, g *commute* (i.e. $fg = gf$), then they have a common eigenvector (see exercise 11, section 1.3.5).

Proof of spectral theorem. We shall prove the theorem by induction on $\dim V$. If $\dim V = 1$, the result is immediate. So we shall suppose that $\dim V > 1$ and that the theorem is true for all spaces of dimension less than V.

By Lemma 2.4.2 we can choose an element of V which is a non-zero eigenvector for both f and f^*. Let $W = \langle v \rangle^\perp$. Since $\langle v \rangle$ is both f and f^* invariant, W will be both f^* and $f^{**} = f$ invariant, by Lemma 2.3.15.

Thus we can apply the inductive hypothesis to W and f_W. (Check that f_W is a normal linear transformation on W, using Exercise 82). We obtain an orthonormal basis of W with respect to which the matrix of f_W is diagonal. Adding the vector $v/\|v\|$ gives an orthonormal basis of V consisting of eigenvectors for f. □

The first version of the spectral theorem is a straightforward statement which is, hopefully, easy to understand. We are now going to give another version which is less easy to understand. But there is a reason for this obfuscation. The second version is one which generalises much more easily to infinite dimensional spaces of the right kind (Hilbert spaces) where matrices are not available (or, at least, much less useful). In Hilbert space, the statement will be similar but the summations will be replaced by integrals.

Theorem 2.4.3 (Spectral Theorem; second version)**.** *Let f be a normal linear transformation on a finite dimensional complex inner product space space V. There exist self-adjoint (Hermitian) linear transformations $e_1, \dots, e_k$ and scalars $a_1, \dots a_k$ such that:*

(1) $a_i \neq a_j$ *if* $i \neq j$*;*

(2) $e_i^2 = e_i$ *and no* e_i *is zero;*

(3) $\sum_{i=1}^k e_i = 1_V$*;*

(4) $\sum_{i=1}^k a_i e_i = f$*.*

(In fact, e_i is the orthogonal projection onto the a_i-eigenspace.)

Proof. Find the orthonormal basis given by the first version of the theorem and let A be the corresponding diagonal matrix. Let $a_1, \dots a_k$ be the distinct eigenvalues of f and so the distinct diagonal entries of A. Let E_i be the diagonal matrix with an entry of 1 wherever a_i occurs in A and an entry of 0 elsewhere. Then $\sum_{i=1}^k E_i$ is the identity matrix and $\sum_{i=1}^k a_i E_i = A$. Also $E_i^2 = E_i$. Let e_i be the linear transformation which has the matrix E_i with respect to the chosen orthonormal basis. The required properties of e_i are now easy to check. □

2.4.2 Polar form

There are strong similarities between complex numbers and complex matrices (or linear transformations over a complex vector space). Just as we have absolute values, real and imaginary parts and polar decomposition for complex numbers, we have similar ideas for normal matrices. We shall try to make most of our ideas eventually independent of the matrix representation of the linear transformation so they can be more easily generalised.

Firstly, some observations.

Lemma 2.4.4. *Let f be a linear transformation on a complex inner product space V.*

(1) If f is unitary then the eigenvalues of f are of absolute value 1.

(2) If f is self-adjoint then the eigenvalues of f are real.

Proof. (1) Suppose that f is unitary and that $f(v) = av$ for some $a \in \mathbb{C}$ and $0 \neq v \in V$. Then

$$a\overline{a}(v, v) = (av, av) = (f(v), f(v)) = (f^*(f(v)), v) = (v, v)$$

and so $a\overline{a} = 1$.

(2) Suppose that f is self-adjoint and that $f(v) = av$ for some $a \in \mathbb{C}$ and $0 \neq v \in V$. Then

$$a(v, v) = (av, v) = (f(v), v) = (v, f^*(v)) = (v, f(v)) = (v, av) = \overline{a}(v, v)$$

and so $a = \overline{a}$; that is, a is real. □

It follows that a diagonal matrix for a self-adjoint linear transformation f has real entries. It is reasonable to define a self-adjoint matrix to be *non-negative* if it has non-negative (real) eigenvalues and *positive* if it has positive eigenvalues. There are some more convenient versions of the definition, however.

Lemma 2.4.5. *Let f be a linear transformation on a finite dimensional complex inner product space V. The following are equivalent:*

(1) f is self-adjoint and all eigenvalues of f are non-negative

(2) $f = g^2$ for some self-adjoint g;

(3) $f = hh^*$ for some h;

(4) f is self-adjoint and $(f(v), v) \geq 0$ for all $v \in V$.

Proof. Suppose that (1) is true and let A be the diagonal matrix for f guaranteed by the Spectral Theorem. Suppose that $A = \mathrm{diag}(a_1, \ldots, a_n)$. Then $a_i \geq 0$ so there are $b_i \geq 0$ with $b_i^2 = a_i$. Let $B = \mathrm{diag}(b_1, \ldots, b_n)$. Then $B^2 = A$. Let g be the linear transformation corresponding to B.

If (2) is true then (3) is trivial; take $h = g$.

Suppose that (3) is true. Then $f^* = (hh^*)^* = h^{**}h^* = hh^* = f$ and so f is self-adjoint. Also,

$$(f(v), v) = (hh^*(v), v) = (h^*(v), h^*(v)) \geq 0$$

and so (4) is true.

Finally, suppose that (4) is true. If a is an eigenvalue of f with associated eigenvector v then $(f(v), v) = (av, v) = a(v, v) \geq 0$ and so $a \geq 0$ as $(v, v) > 0$.

Thus we have proved that (1) implies (2) implies (3) implies (4) implies (1) and so the equivalence of all four.

□

It is not difficult to show that, just as any complex number z can be expressed in the form $z = |z| \exp^{i \arg z}$ with $|z|$ real and non-negative and $\exp^{i \arg z}$ of absolute value 1, we can write any normal linear transformation f as a product of a non-negative self-adjoint linear transformation with a unitary linear transformation. We sketch the argument. Find a diagonal matrix A for f; say $A = \mathrm{diag}(z_1, \ldots, z_n)$. Write $z_i = p_i u_i$ with p_i real and positive and u_i of absolute value 1. Set $P = \mathrm{diag}(p_1, \ldots, p_n)$ and $U = \mathrm{diag}(u_1, \ldots, u_n)$. Then $A = PU$. Let p and u be the linear transformations corresponding to the matrices P and U. Then $f = pu$ is the required decomposition.

In fact we do not need to assume that f is normal.

Theorem 2.4.6. *Let f be a linear transformation on a finite dimensional complex inner product space. Then there exists a non-negative linear transformation p and a unitary linear transformation u such that $f = pu$.*

Proof. We shall give the proof only in the case that f is invertible.

By Lemma 2.4.5, ff^* is non-negative (in fact positive if f is invertible) and so, by Lemma 2.4.5, we can find a non-negative p such that $p^2 = ff^*$. Since f is assumed invertible, so also is p. Then p^{-1} will also be self-adjoint. Consider $p^{-1}f$. We have

$$p^{-1}f\left(p^{-1}f\right)^* = p^{-1}(ff^*)(p^{-1})^* = p^{-1}(p^2)(p^{-1}) = 1_V.$$

Thus $u = p^{-1}f$ is unitary and so $f = pu$ as required. □

2.4.3 Commuting normal matrices

The spectral theorem gives us a tool to deal with matrices which commute and also with the problem of deciding which matrices commute with a given matrix.

Theorem 2.4.7. *Let f and g be normal linear transformations on a finite dimensional complex inner product space V. Suppose that $fg = gf$. Then there is an orthonormal basis for V such that the matrices of both f and g with respect to this basis are diagonal.*

Proof. Let $V_1, \dots, V_m$ be the eigenspaces of f in V and suppose that the corresponding eigenvectors are $a_1, \dots, a_m$. We know from the Spectral Theorem (or by direct checking) that V_i and V_j are orthogonal if $i \neq j$.

We claim that each V_i is g-invariant. For, if $v \in V_i$ and $w = g(v)$,then

$$f(w) = f(g(v)) = g(f(v)) = g(a_i v) = a_i g(v) = a_i w$$

and so $g(v) = w \in V_i$. Thus we can consider g_{V_i}. It is not difficult to check that g_{V_i} is still normal and so we can apply the Spectral Theorem to g_{V_i}. We then find an orthonormal basis $\mathcal{B}_i$ of V_i which consists of eigenvectors of g_{V_i} and so of g. It clearly consists of eigenvectors of f since every element of V_i is an eigenvector of f.

The set $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_m$ is an orthonormal basis of V and consists of vectors which are eigenvectors for both f and g. It follows that the matrices of f and g with respect to this basis are diagonal. □

Theorem 2.4.8. *Let f and g be normal linear transformations on a finite dimensional complex inner product space V. Then f and g commute if and only if there is a normal linear transformation h and polynomials $p(X)$ and $q(X)$ such that $f = p(h)$ and $g = q(h)$.*

Proof. It is clear that two linear transformations of the form $p(h)$ and $q(h)$ must commute so we turn to the converse.

Suppose that f and g are normal linear transformations satisfying $fg = gf$. By the previous theorem, we can find an orthonormal basis of V so that the matrices A of f and B of g, with respect to this basis, are diagonal. Suppose that $A = \operatorname{diag}(a_1, \dots, a_n)$ and $B = \operatorname{diag}(b_1, \dots, b_n)$.

Set $C = \operatorname{diag}(1, \dots, n)$. There will be polynomials p and q so that $p(i) = a_i$ and $q(i) = b_i$. (This requires the theory of 'interpolation' of polynomials. If you haven't seen it before, try to do it from first principles for small values of n.) Thus $p(C) = A$ and $q(C) = B$.

Let h be the linear transformation corresponding to C. Then $p(h) = f$ and $q(h) = g$, as required.

□

2.4.4 Exercises

81. Show that if $A = UDU^*$ where D is a diagonal matrix and U is unitary, then A is a normal matrix. (The spectral theorem implies that the converse is true).

82. Show that a linear transformation $f : V \to V$ on a complex inner product space V is normal if and only if $(f(u), f(v)) = (f^*(u), f^*(v))$ for all $u, v \in V$.

83. Show that every normal matrix A has a square root; that is, a matrix B so that $B^2 = A$.

84. Must every complex square matrix have a square root?

85. Two linear transformations f and g on a finite dimensional complex inner product space are *unitarily equivalent* if there is a unitary linear transformation u such that $g = u^{-1}fu$. Two matrices are *unitarily equivalent* if their linear transformations, with respect to some fixed orthonormal basis, are *unitarily equivalent.*

Decide whether the following matrices are unitarily equivalent.

(a)

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

(b)

$$\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

(c)

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix}.$$

86. Are f and f^* always unitarily equivalent?

87. If f is a normal linear transformation on a finite dimensional complex inner product space, and if $f^2 = f^3$, show that $f = f^2$. Show also that f is self-adjoint.

88. If f is a normal linear transformation on a finite dimensional complex inner product space show that $f^* = p(f)$ for some polynomial p.

89. If f and g are normal linear transformations on a finite dimensional complex inner product space and $fg = gf$, show that $f^*g = gf^*$. (Harder) Prove that the same result holds assuming only that f is normal.

90. Let f be a linear transformation on a finite dimensional complex inner product space. Suppose that f commutes with f^*f; that is, that $f(f^*f) = (f^*f)f$. We aim to show that f is normal.

(a) Show that f^*f is normal.

(b) Choose an orthonormal basis so that the matrix of f^*f takes the block diagonal form $\operatorname{diag}(A_1, \ldots, A_m)$ where $A_i = \lambda_i I_{m_i}$ and $\lambda_i = \lambda_j$ only if $i = j$.

(c) Show that f has matrix, with respect to this basis, of the block diagonal form $\operatorname{diag}(B_1, \ldots, B_m)$ for some $m_i \times m_i$ matrices B_i.

(d) Deduce that $B_i^* B_i = A_i$ and so that $B_i^* B_i = B_i B_i^*$.

(e) Deduce that f is normal.

91. The following is a question (unedited) submitted to an Internet news group:

```
Hello,
I have a question hopefully any of you can help.

As you all know :

If we have a square matrix A, we can always find another
square matrix X such that

   X(-1) * A * X = J

where J is the matrix with Jordan normal form.  Column
vectors of X are called principal vectors of A.
```

(If J is a diagonal matrix, then the diagonal members are the eigenvalues and column vectors of X are eigenvectors.)

It is also known that if A is real and symmetric matrix, then we can find X such that X is "orthogonal" and J is diagonal.

The question :

Are there any less strict conditions of A so that we can guarantee X orthogonal, with J not necessarily a diagonal ?

I would appreciate any answers and/or pointers to any references.

Can you help?

3 Groups

3.1 Symmetries

Warning! This section attempts to motivate the topic of this chapter: groups. As a consequence, you will find it vague and you may find it confusing. If you find it too confusing, ignore it. It is not *necessary* to anything that follows. The same comment applies to the exercises at the end.

We will start with a question:

What are the finite 2-dimensional symmetrical objects?

If you think about this for a while you will probably come to the conclusion that it is a bad question. Let us try to list some of the problems with it:

- What is a finite object? For example, is a square finite? Is a circle finite?
- What does symmetrical mean?
- Assuming we could gain some agreement about the answer to the first two points in this list, how could we list all of the answers?

  Let us assume that a square fits into our ideas of a symmetrical object.
- Are two squares of different side length to be listed separately?
- Are two squares of the same side length but with different centres to be listed separately?
- Are two squares of the same side length and the same centre but with different orientations to be listed separately?

The answer to the latter points in the list above are probably no, otherwise our list would be far too complicated to be of any use or interest. So let us try to re-formulate our question. Perhaps our interest should be in the symmetries themselves rather than the objects of which they are symmetries. That would answer most, if not all, of the questions above.

So we try again, with a different question.

What are the finite sets of symmetries of 2-dimensional objects?

This looks more possible; most of our previous problems disappear with this way of formulating the question. There is a very reasonable objection that we have changed the question and we aren't really answering the same question anymore. But this is at least a major step towards providing the sort of information that the first question was seeking.

Let us see what we can say about the new version of the question. Some sorts of finite sets of symmetries of 2-dimensional objects spring to mind fairly quickly.

Start with a square. It has 4 reflectional symmetries and at least 3, possibly 4, rotational symmetries. The uncertainty there is because we have to decide whether to call 'doing nothing' a symmetry. It turns out to be much more convenient if we do. So we have an example. More should come to mind fairly quickly.

- The set of four reflections and four rotations which are symmetries of a square.
- The set of three reflections and three rotations which are symmetries of an equilateral triangle.
- The set of five reflections and five rotations which are symmetries of a regular pentagon.

- The set of six reflections and six rotations which are symmetries of a regular hexagon.
- ...
- ...
- ...
- Keep going until you get tired, or work out a general statement.

So we have a large, but fairly easily described collection of symmetries of 2-dimensional objects. If you think harder about this, you may see that we can get some other examples by going 'down' rather than 'up' from an equilateral triangle. We can also get the set of two reflections and two rotations which are symmetries of (for example) a non-square rectangle. Finally, we can get the set of one reflection and one rotation which are symmetries of (for example) an isosceles but non-equilateral triangle.

There is another way we can produce new sets of symmetries from the ones we already have. We can try to 'break' some, but not all, of the symmetry of the object, such as the regular polygon, used to represent the set of symmetries. For example, with all of the previous examples of polygons, we can put arrowheads, at the centre of each side of the polygon and all pointing in a clockwise direction. The effect of this is to remove the reflectional symmetries and leave only the rotational symmetries. There are other ways of 'breaking symmetry' but you will find that the set of symmetries you are left with is the same as some other set you have constructed.

So let us summarise the possibilities we have found as an answer to the second question.

Answer (but maybe not complete yet) The following can be finite sets of symmetries of 2-dimensional objects:

- Any set of n rotations and n reflections as described above, where $n = 1, 2, 3, \ldots$.
- Any set of n rotations as described above, where $n = 1, 2, 3, \ldots$.

It is far from clear that these are the only possibilities. But they are, and we will eventually be able to prove it.

If you found all of this rather easy, what about the same problem for symmetries of three-dimensional objects?

3.1.1 Exercises

92. Describe the rotational symmetries of a cube. There are 24 in all. It will probably help to have a cube (or something your imagination will allow you to believe is a cube) near at hand. Are there any other symmetries besides these rotations?

93. Describe the 12 rotational symmetries of a regular tetrahedron.

94. Describe some rotational symmetries of a 4-dimensional cube. (Of course you will first have to work out what a 4-dimensional cube is.)

95. What letters in the Roman alphabet display symmetry?

3.2 What groups are, and some examples

I hope that you have been convinced in the last section that there is some point in studying sets of symmetries separately from the objects of which they are symmetries. Let us look at the properties which a set of symmetries of some geometric object must have. Firstly, by a *composition* of two symmetries, we mean the effect of applying one after the other. So we can see that the composition of two symmetries of the object is a third symmetry of the object. Also, symmetries are reversible; that is we can find a second symmetry so that the composition leads back to where we started. By agreement, 'doing nothing' is a symmetry.

There are many other collections of things, such as the real numbers with 'addition' replacing 'composition' which have similar properties. This leads us into an abstract definition.

3.2.1 Definition of a group

In the following definition, a binary operation on a set G is simply a function of two variables, from G, which takes its values in G. If we used f for this function and if $g, h \in G$, we could therefore write $f(g, h)$ for the value, in G of this function. But the examples of such functions in practice are very often functions such as addition or multiplication and most people find it strange to write $+(a, b)$ or $\times(a, b)$. We shall therefore denote the general function by $*$ (or something similar) and use $g * h$ rather than $*(g, h)$ for the value of the function. Later, we will often abbreviate $g * h$ to gh.

Definition 3.2.1. A ***group*** consists of a set G together with a binary operation $*$ that satisfies

(1) $g * (h * k) = (g * h) * k$ for all $g, h, k \in G$

(2) there is an element e_G in G which satisfies $g * e_G = e_G * g = g$ for all $g \in G$
(we call e_G the ***identity*** of G)

(3) for each $g \in G$ there is an element $g^{-1} \in G$ satisfying $g * g^{-1} = g^{-1} * g = e_G$
(we call g^{-1} the ***inverse*** of g)

We said nothing before the definition about rule (1). That was largely because it was too obvious. This will frequently, but not always, be the case. You need to watch out, however, for cases when the property in rule (1) is not a well-known fact.

Note that when we specify a group, we must specify an operation. Identities and inverses are then all taken with respect to this operation. It is possible that an individual element belongs to a different group with a different operation where its 'inverse' is quite different.

Before we go on to examples, let us work a little with the definitions.

Lemma 3.2.2. *(1) Each group has only one identity.*

(2) Each element has only one inverse.

(3) The inverse of $g * h$ is $h^{-1} * g^{-1}$.

Proof. (1) Suppose e_G and f_G both satisfy the properties required of an identity for the group G. Then, as e_G is an identity, $e_G * f_G = f_G$. Also, as f_G is an identity, $e_G * f_G = e_G$. So $e_G = f_G$.

(2) Exercise 102.

(3)

$$\begin{aligned}(g * h) * \left(h^{-1} * g^{-1}\right) &= g * \left(h * \left(h^{-1} * g^{-1}\right)\right) = g * \left(\left(h * h^{-1}\right) * g^{-1}\right) \\ &= g * \left(e_G * g^{-1}\right) = g * g^{-1} = e_G.\end{aligned}$$

A similar calculation shows that $\left(h^{-1} * g^{-1}\right) * (g * h) = e_G$ and so $h^{-1} * g^{-1}$ is the inverse of $g * h$. □

Remark: Other common symbols for group operations include: $g \cdot h, g \circ h, gh, g + h$.

3.2.2 Examples of groups

(1) The set of all symmetries of a 2-dimensional geometrical figure together with the operation of composition of symmetries. The same will apply in 3 or more dimensions once we formulate more clearly what we mean by a symmetry.

(2) To give a more precise version of the previous examples, denote by D_n—called the *dihedral* group of order n—the set of all symmetries of a regular n-gon together with the operation of composition. For $n \geq 3$ there is no problem with this definition. For $n = 1$ or $n = 2$ we have, for example, to 'interpret' a regular 2-gon as a non-square rectangle and a regular 1-gon as an isosceles but non-equilateral triangle. Then D_n has $2n$ elements and comprises n rotations and n reflections.

(3) The group $(\mathbb{Z}, +)$ of integers together with the operation of addition.

(4) The groups $(\mathbb{Q}, +)$ of rational numbers, $(\mathbb{R}, +)$ of real numbers and $(\mathbb{C}, +)$ of complex numbers furnished, in each case, with the operation of addition.

(5) $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are **not** groups with multiplication as the operation, since 0 has no inverse.

(6) The set $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$ of non-zero rational numbers furnished with the operation of multiplication is a group. We can similarly form groups $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ and $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$.

(7) $\mathbb{Z} \setminus \{0\}$ is **not** a group under multiplication since only ± 1 have inverses.

(8) The group of all $n \times n$ matrices with entries from the real numbers and the operation of matrix addition.

(9) The group of all invertible (non-singular) matrices with entries from the real numbers and the operation of matrix multiplication. This is an important example and we shall devote a sub-section to it.

3.2.3 Matrix groups

Let F be a field; for example F could be the rational numbers, the real numbers, the complex numbers or the integers modulo a prime number p.

We list here the names of some important groups. In every case the operation is matrix multiplication; we do not list this explicitly each time.

$GL(n, F)$	the collection of all $n \times n$ invertible matrices with entries from F
$O(n)$	the collection of all $n \times n$ orthogonal matrices (real matrices A such that $A^T A = I$)
$U(n)$	the collection of all $n \times n$ unitary matrices (complex matrices U such that $U^* U = I$)
$SL(n, F)$	the collection of all $n \times n$ matrices of determinant 1 with entries from F
$SO(n)$	the collection of all $n \times n$ orthogonal matrices of determinant 1
$SU(n)$	the collection of all $n \times n$ unitary matrices of determinant 1

There are also many other groups of matrices which have important special properties and corresponding names.

3.2.4 Groups with at most 4 elements

First we note some other simple properties of groups.

Lemma 3.2.3. *(1) In any group G we have* cancellation laws*:*

(a) $g * x = g * y$ implies $x = y$, *(b) $x * h = y * h$ implies $x = y$.*

(2) We can also solve equations*: Given any g, h in a group G there are unique $x, y \in G$ such that*

(a) $g * x = h$, *(b) $y * g = h$.*

Proof. 1(a) If $g * x = g * y$, then $g^{-1} * (g * x) = g^{-1} * (g * y)$, so $(g^{-1} * g) * x = (g^{-1} * g) * y$ by associativity. Hence $e * x = e * y$ by the property of inverses, so $x = y$ by the property of the identity element e. 1(b) is similar.

We leave part (2) as Exercise 99. □

Consequence: In a group multiplication table each element occurs *exactly once* in each row and each column. For example, 1(a) says that all entries gx in the row containing g are different; 2(a) says that every group element occurs in this row.

Examples: Using this observation, we see

(1) A group with 2 elements $\{e, a\}$ has a multiplication table:

$*$	e	a
e	e	a
a	a	e

So there is essentially only one possible group with 2 elements; $(\mathbb{Z}/2\mathbb{Z}, +)$ is such a group. (Two groups are "essentially the same" or *isomorphic* if their multiplication tables are the same after suitable *renaming* of elements.)

(2) For a group with 3 elements $\{e, a, b\}$ we must have:

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

Again, there is only one way to fill in the missing entries, so there is essentially only one possible group with 3 elements; $(\mathbb{Z}/3\mathbb{Z}, +)$ is such a group.

Exercise 96. Find two "different" multiplication tables for groups with 4 elements.

3.2.5 Permutations

Most of us have an intuitive idea of what a permutation is; it is a re-arrangement of some sort. We shall need a somewhat more precise version, however, so that we can work in more detail with them.

Definition 3.2.4. A ***permutation*** of a set S is a function $f : S \to S$ which is bijective, i.e. f is one-to-one and onto.

Thus the function does the 're-arrangement' for us. For example suppose that S has two elements a and b. There are two permutations of S:

$$\begin{array}{c} a \to a \\ b \to b \end{array} \quad \text{and} \quad \begin{array}{c} a \to b \\ b \to a \end{array}.$$

Now observe that the composition of two bijective functions is bijective and so the composition of two permutations is a permutation. Also the inverse of a bijective function is bijective and so the inverse of a permutation is a permutation. Thus we can see that the set of all permutations of a set S, together with the operation of composition of functions, gives a group which we will denote $\mathrm{Sym}(S)$.

A little thought should convince you that the number of permutations of a set and how they combine together does not depend on the names we give to the elements of the set. So if we want to consider the permutations of a finite set of size n, we will usually take the set to be $S = \{1, \ldots, n\}$.

The permutations of $\{1, 2, 3\}$ are

$$\begin{array}{cccccc} 1 \to 1 & 1 \to 2 & 1 \to 3 & 1 \to 2 & 1 \to 3 & 1 \to 1 \\ 2 \to 2 & 2 \to 3 & 2 \to 1 & 2 \to 1 & 2 \to 2 & 2 \to 3 \\ 3 \to 3 & 3 \to 1 & 3 \to 2 & 3 \to 3 & 3 \to 1 & 3 \to 2 \end{array}$$

We can write a permutation f without the arrows by writing it in the form

$$\begin{pmatrix} 1 & 2 & 3 & \ldots\ldots & n \\ f(1) & f(2) & f(3) & \ldots\ldots & f(n) \end{pmatrix}$$

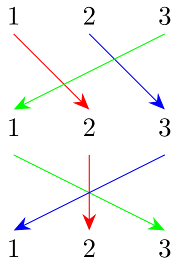
Thus the second permutation above can be written as

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

and the fifth as

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

We can see how to multiply these permutations by drawing a diagram:



We then deduce that

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

Note that we write the permutation that is to be applied first on the **right**. This is because we have thought of permutations as functions and by a composite fg of functions f and g, we usually mean 'first apply g then apply f'.

Similarly, we can calculate that

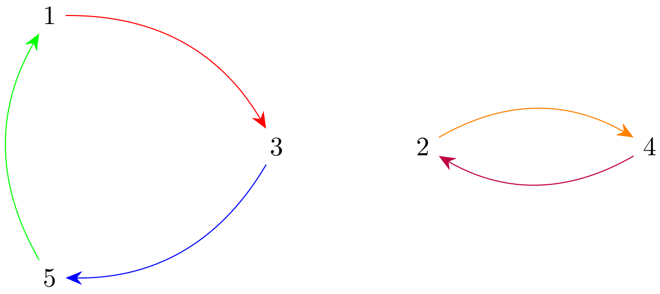
$$\begin{pmatrix}1 & 2 & 3 & 4\\ 3 & 1 & 2 & 4\end{pmatrix}\begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 3 & 4 & 1\end{pmatrix} = \begin{pmatrix}1 & 2 & 3 & 4\\ 1 & 2 & 4 & 3\end{pmatrix}.$$

As you can see, this is not a very efficient way to describe permutations. We are writing the top row of the permutation in the same way every time. So we adopt a new notation, called **cycle notation**:

Example: Let $S = \{1, 2, 3, 4, 5\}$. The permutation

$$f = \begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 3 & 4 & 5 & 2 & 1\end{pmatrix}$$

acts on S as shown below:



So S breaks up into two "cycles" when f is applied: $1 \to 3 \to 5 \to 1$ and $2 \to 4 \to 2$. In *cycle notation* we write:

$$f = (135)(24).$$

In general we describe each permutation as follows:

(1) Open a left parenthesis;

(2) write any element of the set;

(3) after each element write its image under the permutation;

(4) when you would write an element that has been written before, don't—but write a right parenthesis;

(5) if there are any elements of the set left unwritten, start again from step 1.

So, for example, $\begin{pmatrix}1 & 2 & 3 & 4\\ 3 & 1 & 2 & 4\end{pmatrix}$ is written as $(132)(4)$. We can recover the long form easily enough from the new short form if we need it. Note that the last element before a right parenthesis must be sent to the first element after the previous left parenthesis; in the above example, 2 is sent to 1, for example.

Some more examples

$$\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 3 & 1 & 4 & 2 & 7 & 5 & 6\end{pmatrix} \quad \text{is represented by } (1342)(576)$$

$$\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 2 & 3 & 4 & 5 & 6 & 7 & 1\end{pmatrix} \quad \text{is represented by } (1234567)$$

$$\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 2 & 1 & 4 & 3 & 6 & 5 & 7\end{pmatrix} \quad \text{is represented by } (12)(34)(56)(7)$$

Each string inside a set of parentheses is known as a cycle. In the representation of a permutation, no element of the set can appear in more than one of these cycles. This way of writing permutations is known as expressing them as a *product of disjoint cycles.*

Note: Cycle notation for a permutation is *not* unique, for example $(123) = (231) = (312)$ all represent the permutation taking $1 \to 2$, $2 \to 3$ and $3 \to 1$.

It is not hard to see how to multiply permutations when they are expressed in this form. The main thing is to remember that we must work *from the right.* We then just trace through the images of successive elements and write them according to the rules above. For example

$$(1324)(567) * (143)(5)(267) = (1)(274)(3)(56) \text{ and } (123) * (321) = (1)(2)(3).$$

There is yet another convention that is often used to save space. There is really little difference between (123) and (123)(4) so we often omit the letters that are fixed. With this convention the first product above would be (274)(56). The second product above yields an evident problem which we solve by calling it (1) (rather than (‘ ’)).

Definition 3.2.5. The set of all permutations of the set $\{1, \ldots, n\}$, together with the operation of composition of permutations, is called the ***symmetric group*** on n letters and is written S_n.

Note that S_n has $n!$ elements. In the new notation, S_3 has the elements

$$(1), (123), (132), (12), (13), (23).$$

3.2.6 Exercises

97. Decide whether the following are groups:

(a) the set of positive real numbers with the operation of addition;

(b) the set of all $n \times n$ matrices over the real numbers with the operation of addition;

(c) the set of all $n \times n$ matrices over the real numbers with the operation of multiplication;

98. Show that the set of all rotations of the plane about a fixed centre P, together with the operation of composition of symmetries, form a group. What about all of the reflections for which the axis (or mirror) passes through P?

99. Suppose that x and y are elements of a group. Show that there are elements w and z so that $wx = y$ and $xz = y$. Show that w and z are unique. Must w be equal to z?

100. Let n be a fixed natural number. Show that the set of complex numbers z which are nth roots of unity, that is which satisfy $z^n = 1$, together with multiplication of complex numbers, forms a group.

101. Show that the set of complex numbers z which are nth roots of unity for some (variable) natural number n, together with multiplication of complex numbers, forms a group.

102. Prove part (2) of Lemma 3.2.2.

103. Find the product of the following permutations:

(a) $(123)(456) * (134)(25)(6)$;

(b) $(12345) * (1234567)$;

(c) $(123456) * (123) * (123) * (1)$.

104. Set $X = \mathbb{R} \setminus \{0, 1\}$. Show the following set of functions $X \to X$, together with the operation of composition, form a group.

$$\begin{array}{ccc} f(x) = \frac{1}{1-x} & g(x) = \frac{x-1}{x} & h(x) = \frac{1}{x} \\ i(x) = x & j(x) = 1 - x & k(x) = \frac{x}{x-1} \end{array} .$$

105. (Harder) Describe the product of a rotation of the plane with a translation. Describe the product of two (planar) rotations about different axes.

3.3 Group discussion; some terminology for groups

3.3.1 Subgroups

In a number of the examples of groups given above, the underlying set of one group is a subset of the other and they use the same operation. Such groups are clearly very closely allied.

Definition 3.3.1. A subset H of a group G is a ***subgroup*** if it is a group in its own right, using the operation of G restricted to H. We often write this $H \leq G$.

If we know that G is a group, then the checking that H is a subgroup is made somewhat easier than usual. For example, we do not need to check that the operation is associative. In fact, we have the following.

Lemma 3.3.2. *Let $(G, *)$ be a group and let H be a non-empty subset of G. Then the following are equivalent:*

(1) H is a subgroup;

(2) H is closed under $*$ and inversion: for all $h_1, h_2 \in H$ we have $h_1 * h_2 \in H$ and $h_1^{-1} \in H$;

(3) for all $h_1, h_2 \in H$ we have $h_1 * h_2^{-1} \in H$.

Proof. We shall prove this by showing that (1) implies (2), that (2) implies (3) and that (3) implies (1).

It is clear from the definition of subgroup that (1) implies (2). It is not too hard to show that (2) implies (3). The real work is to show that (3) implies (1).

Suppose that (3) is true. Since H is non-empty it contains some element h. Setting $h_1 = h_2 = h$ we deduce that $e_G = h * h^{-1} \in H$. Now, if $k \in H$, then setting $h_1 = e_G$ and $h_2 = k$ we deduce that $k^{-1} \in H$. Finally, if $k_1, k_2 \in H$ then $k_2^{-1} \in H$ and setting $h_1 = k_1$ and $h_2 = k_2^{-1}$, we deduce that $k_1 * k_2 \in H$.

Thus the operation $*$, when restricted to H, gives a well-defined operation. The operation is associative since it is the same as that used for G. The identity element e_G of G lies in H and is an identity element for the restriction of $*$ to H. Finally each element of H has an inverse, in H, with respect to the restriction of $*$ to H. Thus H, together with the restriction of $*$ to H, forms a group. That is, H is a subgroup of G. □

Examples:

(1) The set $2\mathbb{Z}$ of even integers is a subgroup of the group $\mathbb{Z}$.

(2) The set $\{e_G\}$ is always a subgroup of G; G is always a subgroup of G.

(3) The subset $\{(1), (123), (132)\}$ is a subgroup of S_3.

(4) $SL(n, F)$ is a subgroup of $GL(n, F)$; in fact all the groups defined in the sub-section on matrix groups are subgroups of the appropriate $GL(n, F)$.

(5) Take any group of symmetries of a geometrical object, for example the group D_6 of a regular hexagon. Colour the edges of the hexagon in some way. Then the set of symmetries which preserve the colours of the edges is a subgroup of D_6.

(6) The set of negative integers is **not** a subgroup of $\mathbb{Z}$.

(7) The set $\{(1), (12), (23), (13)\}$ is **not** a subgroup of S_3.

(8) The set of all rotations form a subgroup of D_n; the set of all reflections do **not**.

(9) Let G be any group, $g \in G$ any element of G. Then the set $\{g^n : n \in \mathbb{Z}\}$ is a subgroup of G.

The last example is sufficiently important that we devote a whole sub-section to it.

3.3.2 Cyclic subgroups

Using the associative law and induction it can be shown that the product of group elements $g_1, g_2, \dots, g_n$ (in this order) does not depend on how parentheses are inserted. So we can write the product as $g_1 g_2 \dots g_n$.

In particular for any element g of a group, we can define g^n to be the product of n copies of g for $n = 1, 2, 3, \dots$. We also define $g^0 = e$ and $g^{-n} = (g^{-1})^n$ for $n = 1, 2, 3, \dots$.

Properties of powers in a group: For all $n, m \in \mathbb{Z}$:

(a) $g^{-n} = (g^n)^{-1}$

(b) $g^n g^m = g^{n+m}$

(c) $(g^m)^n = g^{mn}$

We leave this as an exercise.

Lemma 3.3.3. *Let G be a group and $g \in G$. The set $\langle g \rangle = \{g^n : n \in \mathbb{Z}\}$ is a subgroup, called the* cyclic subgroup generated by g.

Proof. $\langle g \rangle$ is closed under the group operation by property (b), and is closed under inversion by property (a) above. □

For groups with operation $+$, we write mg instead of g^m. Then $\langle g \rangle = \{mg : m \in \mathbb{Z}\}$.

Definition 3.3.4. A group G is called ***cyclic*** if it is equal to one of its cyclic subgroups, i.e. $G = \langle g \rangle$ for some element $g \in G$.

Examples:

(1) $\mathbb{Z}$ is cyclic; $\mathbb{Z} = \langle 1 \rangle = \langle -1 \rangle$.

(2) $\mathbb{Z}/n\mathbb{Z}$ is cyclic; $\mathbb{Z}/n\mathbb{Z} = \langle \overline{1} \rangle$.

(3) The subgroup $\{(1), (123), (132)\}$ is cyclic; it equals $\langle (132) \rangle$.

(4) The group S_3 is **not** cyclic.

Definition 3.3.5. The ***order*** of a group G is the number of elements in G. This is written $|G|$.

Examples:

(1) $|\mathbb{Z}/n\mathbb{Z}| = n$, $|S_n| = n!$, $|D_n| = 2n$.

(2) $|\mathbb{Z}| = \infty$ or "$\mathbb{Z}$ has infinite order".

Definition 3.3.6. The ***order*** of an element g of a group G is the number of elements in the cyclic subgroup $\langle g \rangle$ that it generates; that is, $|\langle g \rangle|$. It is usually written $|g|$.

If $\langle g \rangle$ is infinite, as is the case with $1 \in \mathbb{Z}$ for example, we say that g has infinite order.

Examples:

(1) In $\mathbb{Z}/4\mathbb{Z} = \{[0]_4, [1]_4, [2]_4, [3]_4\}$, $[0]_4$ has order 1, $[1]_4$ has order 4, $[2]_4$ has order 2 and $[3]_4$ has order 4.

(2) In $S_3 = \{(1), (123), (132), (12), (13), (23)\}$, (1) has order 1; (123) and (132) have order 3; (12), (13), (23) have order 2. In particular, S_3 contains no element of order 6, so is not a cyclic group.

Lemma 3.3.7. *Let g be an element of a group G.*

(1) If $|g|$ is infinite then $g^m = g^n$ only if $m = n$.

(2) If $|g|$ is finite, say $|g| = k$, then $g^m = g^n$ if and only if $m \equiv n \pmod{k}$. In particular, $|g|$ is the least positive integer so that $g^k = e_G$.

Proof. Suppose that $g^m = g^n$ for some $m \neq n$. Then $g^{m-n} = g^m(g^n)^{-1} = e_G$ and so some non-zero power of g is the identity. Let k be the least positive integer such that $g^k = e_G$. Using the usual division with remainder of integers, we can write $(m - n) = qk + r$ with $q, r \in \mathbb{Z}$ and $0 \leq r < k$. Then

$$g^r = g^{(m-n)-qk} = g^{m-n}(g^k)^{-q} = e_G(e_G)^{-q} = e_G.$$

This will contradict the choice of k as the least positive power of g which is equal to e_G unless $r = 0$. But then $m - n$ is a multiple of k.

In summary, we have shown that if $g^m = g^n$ for some $m \neq n$ then $m \equiv n \pmod{k}$. It is easy to check that the converse is true; that is, if k is least such that $g^k = 1$ and if $m \equiv n \pmod{k}$ then $g^m = g^n$. It follows that the order of g is k.

If $|g|$ is infinite, it follows that $g^m \neq g^n$ whenever $m \neq n$ which proves (1). If $|g|$ is finite, then the proof of (2) follows immediately from what we have done above. □

Note: From part (2) above, the order of an element g is equal to the least positive integer k so that $g^k = e_G$.

Subgroups of cyclic groups are particularly easy to understand:

Lemma 3.3.8. *Every subgroup of a cyclic group is again cyclic.*

Proof. Let $G = \langle g \rangle$ be a cyclic group and let H be a subgroup of G. If $H = \langle e_G \rangle$ then there is nothing to prove. Suppose therefore that H contains some element other than the identity. Since G consists of powers of g, so also must H. Let g^m be the smallest positive power of g which lies in H; we claim that $H = \langle g^m \rangle$.

Suppose that $g^n \in H$. Use the Euclidean algorithm to write $n = qm + r$ with $0 \leq r < m$. Then $g^m \in H$ implies $(g^m)^q = g^{qm} \in H$. But $g^n \in H$ and so $g^n(g^{qm})^{-1} \in H$. But $g^n(g^{qm})^{-1} = g^{n-qm} = g^r$ and so $g^r \in H$. But r is positive or zero and g^m was the smallest positive power of g lying in H. Thus r must be zero. So $n = qm$ and so $g^n = (g^m)^q \in \langle g^m \rangle$. Since g^n was an arbitrary element of H, it follows that $H = \langle g^m \rangle$. □

Definition 3.3.9. A group G is ***commutative*** (or ***abelian***) if $gh = hg$ for all elements g, h of G.

For example, $\mathbb{Z}$ is commutative but $GL(n, F)$ is not when $n > 1$. Cyclic groups are commutative.

Definition 3.3.10. Suppose that $g_1, \ldots, g_k$ are elements of a group G. Then $\langle g_1, \ldots, g_k \rangle$ denotes the smallest subgroup of G containing $g_1, \ldots, g_k$. We also say that $\langle g_1, \ldots, g_k \rangle$ is the ***subgroup generated by*** $g_1, \ldots, g_k$. (In fact this consists of *all possible products* $g_{i_1}^{n_1} g_{i_2}^{n_2} \ldots g_{i_m}^{n_m}$ of powers of $g_1, \ldots, g_k$.)

Note that ‘subgroup generated by’ is similar to ‘subspace spanned by’ in vector spaces.

3.3.3 Isomorphism

In discussing groups, we are interested only in the operation on the set, not in the names given to the elements. There are many groups of order 2, for example $\mathbb{Z}/2\mathbb{Z}$, the subgroup $\langle(12)\rangle$ of S_3, the subgroup $\langle r\rangle$ generated by a reflection r in D_6, the subgroup $\{\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}\}$ of $GL(2,\mathbb{R})$ are all of order 2. However we saw in section 3.2.4 that they all have essentially the same multiplication table:

$*$	e_G	g
e_G	e_G	g
g	g	e_G

We do not want to distinguish between groups which have, after relabelling, the same multiplication table. We can make this formal as follows.

Definition 3.3.11. An ***isomorphism*** between groups G and H is a bijection (1-1 and onto function) $f : G \to H$ such that

$$f(g_1 *_G g_2) = f(g_1) *_H f(g_2)$$

for all elements $g_1, g_2 \in G$. If there is an isomorphism between G and H we say that G and H are ***isomorphic*** and write $G \cong H$.

Literally, isomorphic groups have 'the same shape'. After possible relabelling, they will have the same multiplication table.

Example: In section 3.2.4 we showed that:

(a) all groups of order 2 are isomorphic, and

(b) all groups of order 3 are isomorphic.

Lemma 3.3.12. *A cyclic group is isomorphic to either $\mathbb{Z}$ or to $\mathbb{Z}/n\mathbb{Z}$ for some $n \geq 1$.*

Proof. Let $G = \langle g\rangle$ be a cyclic group. If G has infinite order, define a function $f\colon \mathbb{Z} \to G$ by $f(m) = g^m$. Part (1) of Lemma 3.3.7 shows that f is injective. The definition of cyclic group and the notation $G = \langle g\rangle$ show that f is surjective. Finally, the fact that f is an isomorphism simply comes from the fact that $g^{m+n} = g^m * g^n$.

If G has finite order k the argument is very similar. This time we define f by $f\colon \mathbb{Z}/k\mathbb{Z} \to G$ by $f([m]_k) = g^k$. The extra step in this argument is to show that f is well-defined. That is, we must show that if $[m]_k = [n]_k$ then $g^m = g^n$. But this follows from (2) of 3.3.7. □

The following lemma gives some useful some properties of isomorphisms.

Lemma 3.3.13. *Let $f : G \to H$ be an isomorphism between groups. Then*

(1) If e_G is the identity in G, then $f(e_G) = e_H$ is the identity in H.

(2) If $g \in G$, then $f(g^{-1}) = f(g)^{-1}$.

(3) If $g \in G$, then $|g| = |f(g)|$.

Proof. Exercise 117. □

Isomorphic groups must clearly have the same order. They must also have the same 'properties of multiplication'. For example, if one of a pair of isomorphic groups is commutative, so must be the other. If one has an element of order 29, so must the other. This is usually the first thing to consider when trying to show that two groups are non-isomorphic.

For example, S_3 is not isomorphic to $\mathbb{Z}/6\mathbb{Z}$ since the latter is commutative but the former is not. The group D_2 is not isomorphic to $\mathbb{Z}_4$; both are commutative but the latter is cyclic whereas the former is not.

Examples:

(1) What about D_3 and S_3? Both are non-commutative and non-cyclic. In fact they are isomorphic. We can set up the isomorphism geometrically. Consider D_3 as the symmetry group of an equilateral triangle $\mathcal{T}$ and number the vertices with $\{1, 2, 3\}$. Each symmetry σ of $\mathcal{T}$ permutes the vertices and so can be associated with a permutation $f(\sigma)$ of $\{1, 2, 3\}$, that is an element of S_3. Once we know what σ does to the vertices of $\mathcal{T}$, we know what it does to all of $\mathcal{T}$. Thus f is an injective function. Since $|D_3| = |S_3|$, f must be bijective. It is not difficult now to see that f must be an isomorphism.

In fact we can try the same sort of thing with D_n and S_n but can then only conclude that D_n is isomorphic to a subgroup of S_n. We cannot expect D_n and S_n to be isomorphic for $n > 3$ since they have different orders.

(2) Consider the group $\mathbb{R}$ of real numbers under addition and the group $\mathbb{R}_{>0}$ of positive real numbers with multiplication. We claim that they are isomorphic.

The function
$$\exp : \mathbb{R} \to \mathbb{R}_{>0}$$
is bijective and $\exp(a + b) = \exp(a)\exp(b)$ shows that exp is an isomorphism.

In general it can be very difficult to decide whether two groups are isomorphic or not; see Exercise 120 for some harder examples.

3.3.4 Products of groups

There is an easy way to combine groups to produce new groups.

Definition 3.3.14. Let G_1, G_2 be groups. Then the ***direct product*** $G_1 \times G_2$ is the group of ordered pairs
$$\{(g_1, g_2) : g_1 \in G_1, g_2 \in G_2\}$$
with operation:
$$(g_1, g_2) * (h_1, h_2) = (g_1 *_{G_1} h_1, g_2 *_{G_2} h_2)$$
for all $g_1, h_1 \in G_1$ and $g_2, h_2 \in G_2$.

It is easy to check that this defines a *group*:

(0) $G_1 \times G_2$ is closed under $*$ since G_1, G_2 are closed under their operations.

(1) (e_{G_1}, e_{G_2}) is an identity element

(2) $(g_1, g_2)^{-1} = ((g_1^{-1}, g_2^{-1})$

(3) the associative law in $G_1 \times G_2$ follows from the fact that it holds in G_1 and in G_2.

Note that:

(1) $|G_1 \times G_2| = |G_1| \cdot |G_2|$

(2) $G_1 \times G_2$ is abelian if G_1, G_2 are abelian.

Similarly we can define the product
$$G_1 \times G_2 \times \ldots \times G_n$$
of n groups.

Examples:

(1) If C_2 and C_3 are cyclic groups of order 2 and 3, then $C_2 \times C_3$ is a cyclic group of order 6.

(2) $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ is a non-cyclic abelian group of order 4. It is isomorphic to the dihedral group D_2.

(3) $\mathbb{Z}/8\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, and $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ are *non-isomorphic* abelian groups of order 8.

3.3.5 Exercises

106. Find the order of the following elements

(a) $(123)(4567)(89)$ in S_{10};

(b) $(14)(23567)$ in S_7;

(c) a reflection;

(d) a translation in the group of symmetries of a plane pattern;

(e) the elements $[6]_{20}, [12]_{20}, [11]_{20}, [14]_{20}$ in the additive group of $\mathbb{Z}/20\mathbb{Z}$;

(f) the elements $[2]_{13}, [12]_{13}, [8]_{13}$ in the multiplicative group of non-zero elements of $\mathbb{Z}/13\mathbb{Z}$.

107. If g is an element of a group G, prove that the orders of g and g^{-1} are equal.

108. Show that, in a *commutative* group, the product of two elements of finite order again has finite order.

109. Can you find an example of two symmetries of finite order where the product is of infinite order?

110. Set

$$A = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Show that A has order 3, that B has order 4 and that AB has infinite order.

111. Determine the possible orders of elements in the dihedral group D_n.

112. If G is a group and $(gh)^2 = g^2h^2$ for all $g, h \in G$, prove that G is commutative.

113. Decide whether the following are subgroups:

(a) the positive integers in the additive group of the integers;

(b) the set of all rotations in the group of symmetries of a plane tesselation;

(c) the set of all permutations which fix 1 in S_n.

114. List all of the subgroups of $\mathbb{Z}/12\mathbb{Z}$.

115. If H is a subgroup of a group G and if $g \in G$, show that $gHg^{-1} = \{ghg^{-1} : h \in H\}$ is a subgroup of G.

116. If G is a group and $g \in G$, show that the function $f : G \to G$ given by

$$f : h \mapsto ghg^{-1}$$

is an isomorphism from G onto itself.

117. Prove Lemma 3.3.13.

118. Show that the matrix group $SO(2)$ is isomorphic to the multiplicative group $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ of complex number with modulus 1.

119. (Harder) Show that D_m is isomorphic to a subgroup of D_n when m divides n.

120. (Harder) Show that

(a) $(\mathbb{R}, +)$ and $(\mathbb{R}^*, \times)$ are not isomorphic.

(b) $(\mathbb{Z}, +)$ and $(\mathbb{Q}, +)$ are not isomorphic.

(c) The additive group of rational numbers $(\mathbb{Q}, +)$ is not isomorphic to the multiplicative group of positive rationals $(\mathbb{Q}_{>0}, \times)$.

3.4 Lagrange's Theorem

When we construct the 'integers modulo n', $\mathbb{Z}/n\mathbb{Z}$, we form the elements of $\mathbb{Z}/n\mathbb{Z}$ by identifying all integers which differ by a multiple of n. We can state this in more group-theoretical language. The set of multiples of n is a subgroup $n\mathbb{Z}$ and we identify k and l if $k-l \in n\mathbb{Z}$. Alternatively, we identify all of the integers in $k+n\mathbb{Z} = \{k+ns : s \in \mathbb{Z}\}$. We are eventually going to make a similar construction with groups in general but firstly we are going to investigate subsets like $k+n\mathbb{Z}$.

3.4.1 Cosets

Definition 3.4.1. Let G be a group and let H be a subgroup of G. A ***right coset*** of H in G is any set of the form $Hg = \{hg : h \in H\}$ for some $g \in G$. Similarly a ***left coset*** of H in G is a set of the form $gH = \{gh : h \in H\}$ for some $g \in G$.

Note that $H = H.e_G$ is always a coset of itself.

Examples:

(1) $G = \mathbb{Z}$; $H = 2\mathbb{Z}$. The cosets are $2\mathbb{Z}$, the even integers and $1+2\mathbb{Z}$, the odd integers.

(2) $G = D_n$, $H = C_n$ the subgroup of rotations in D_n. The cosets are H itself and Hg where g is any reflection. Thus Hg is just the set of reflections in G.

(3) $G = S_3$, $H = \langle(123)\rangle$. The cosets are
$\{(1),(123),(132)\}$, $\{(12),(23),(13)\}$.

(4) $G = S_3$, $H = \langle(12)\rangle$. The cosets are
$\{(1),(12)\}$, $\{(123),(23)\}$, $\{(132),(13)\}$.

(5) $G = GL(2,\mathbb{R})$, $H = SL(2,\mathbb{R})$. The coset $H\begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}$ is just the set of all matrices of determinant a in $GL(2,\mathbb{R})$.

Notice that in these examples, the distinct cosets of H fill up G, are disjoint, and have the same number of elements as H. In fact this always happens.

Lemma 3.4.2. *Let G be a group and H be a subgroup of G.*

(1) If $a, b \in G$, then $Ha = Hb$ if and only if $ab^{-1} \in H$.

(2) Each element of G lies in exactly one coset of H.

(3) The function $Ha \to Hb$ given by $ha \mapsto hb$ for $h \in H$ is a bijection between Ha and Hb.

(Similarly for left cosets.)

Proof. (1) If $Ha = Hb$ then $a = e_G a \in Ha = Hb$ and so $a = hb$ for some $h \in H$. Thus $h = ab^{-1} \in H$. Conversely, if $ab^{-1} = h \in H$ then $a = hb$ and so, if $k \in H$, then $ka = khb \in Hb$ as $kh \in H$. So $Ha \subseteq Hb$. For the reverse inclusion (that $Hb \subseteq Ha$) note that $ba^{-1} = \left(ab^{-1}\right)^{-1} \in H$ and so we can repeat the previous argument, reversing the roles of a and b.

(2) Since $g \in Hg$ it is clear that every element of G lies in at least one coset of H. Suppose that $g \in Ha$ and $g \in Hb$. Then $g = h_1 a$ and $g = h_2 b$ for some h_1 and h_2 in H. So $ga^{-1} \in H$ and $gb^{-1} \in H$. Thus $\left(ga^{-1}\right)^{-1} gb^{-1} = ab^{-1} \in H$. By the first part of the lemma, $Ha = Hb$ and so the 'two' cosets in which g lies are in fact the same.

(3) Firstly note that each element of Ha can be expressed *uniquely* in the form ha with $h \in H$. So the function is well defined. The fact that it has an inverse given by $hb \mapsto ha$ shows that it is a bijection. □

Definition 3.4.3. Let G be a group and H a subgroup of G. The number of different cosets of H in G is called the ***index*** of H in G. It is often written $|G : H|$.

Examples: In the previous examples:

(1) $|\mathbb{Z} : 2\mathbb{Z}| = 2$,

(2) $|D_n : C_n| = 2$,

(3) $|S_3 : \langle(123)\rangle| = 2$,

(4) $|S_3 : \langle(12)\rangle| = 3$,

(5) $|GL(2,\mathbb{R}) : SL(2,\mathbb{R})| = \infty$.

3.4.2 The Theorem

Next we will look at one of the oldest and most important results about finite groups.

Theorem 3.4.4 (Lagrange's Theorem)**.** *Let G be a group of finite order and let H be a subgroup of G. Then $|H|$ divides $|G|$. If $g \in G$ then $|g|$ divides $|G|$.*

Proof. Firstly note that the second sentence follows easily from the first because $|g| = |\langle g\rangle|$ and $\langle g\rangle$ is a subgroup of G.

By part (2) of Lemma 3.4.2, we can write G as a non-overlapping union of cosets of H. By part (3) of Lemma 3.4.2 all cosets have the same order which is therefore also the order of H. So

$$|G| = |G : H||H|$$

and the theorem follows. □

An immediate consequence is the following.

Corollary 3.4.5. *If G is a finite group with $|G| = n$, then $g^n = e$ for all $g \in G$.*

Proof. Let $|g| = k$. Then k divides $|G| = n$, so $n = km$ for some integer m. Thus

$$g^n = g^{km} = (g^k)^m = e^m = e.$$

□

3.4.3 Some applications

One easy application is the following famous result from Number Theory.

Theorem 3.4.6 (Fermat's Little Theorem)**.** *Let p be a prime number. If a is any integer which is not a multiple of p, then*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Proof. The non-zero elements of $\mathbb{Z}/p\mathbb{Z}$ form a group, $(\mathbb{Z}/p\mathbb{Z})^\times = \mathbb{Z}/p\mathbb{Z} \setminus \{[0]_p\}$, under multiplication. This follows from the fact, proved at the beginning of the Linear Algebra chapter, that $\mathbb{Z}/p\mathbb{Z}$ is a field.

The order of the group $(\mathbb{Z}/p\mathbb{Z})^\times$ is $p-1$. If a is any integer which is not a multiple of p then $[a]_p \neq [0]_p$ and so $[a]_p \in (\mathbb{Z}/p\mathbb{Z})^\times$. Hence, by the last Corollary, $[a^{p-1}]_p = [a]_p^{p-1} = [1]_p$ in $(\mathbb{Z}/p\mathbb{Z})^\times$. Thus $a^{p-1} \equiv 1 \pmod{p}$, as required. □

Remark: One surprising application of Fermat's little theorem is to *cryptography* — it is the basis for the "RSA" public key cryptosystem (see Section 3.4.5).

Lagrange's theorem is also important for classifying finite groups. For example:

Theorem 3.4.7. *Let p be a prime number. Then every group of order p is cyclic, so isomorphic to $(\mathbb{Z}/p\mathbb{Z}, +)$.*

Proof. Let G be a group of order p and choose $g \in G$ with $g \neq e_G$. Then $\langle g \rangle$ is a subgroup of G and so $|\langle g \rangle|$ divides $|G| = p$. Since p is prime, $|\langle g \rangle| = 1$ or $|\langle g \rangle| = p$. Since $g \neq e_G$ we cannot have only one element in $\langle g \rangle$ and so $|\langle g \rangle| = p$. But then every element of G is in $\langle g \rangle$ and so $G = \langle g \rangle$; that is, G is cyclic. □

3.4.4 Exercises

121. If H and K are subgroups of a group G and if $|H| = 7$ and $|K| = 29$, show that $H \cap K = \{e_G\}$.

122. Let G be the subgroup of $GL(2, \mathbb{R})$ of the form

$$G = \{ \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} ; x, y \in \mathbb{R}, x > 0\}.$$

Let H be the subgroup of G defined by

$$H = \{ \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} ; z \in \mathbb{R}, z > 0\}.$$

Each element of G can be identified with a point (x, y) of the (x, y)-plane. Use this to describe the right cosets of H in G geometrically. Do the same for the left cosets of H in G.

123. Consider the set $AX = B$ of linear equations, where X and B are column matrices, X is the matrix of unknowns and A the matrix of coefficients. Let W be the subspace (and so additive subgroup) of $\mathbb{R}^n$ which is the set of solutions of the homogeneous equations $AX = 0$. Show that the set of solutions of $AX = B$ is either empty or is a coset of W in the group $\mathbb{R}^n$ with addition.

124. Let H be a subgroup of index 2 in a group G. If $a, b \in G$ and $a \notin H$ and $b \notin H$, show that $ab \in H$.

125. (Harder) Let H be a subgroup of a group G with the property that if $a, b \in G$ and $a \notin H$ and $b \notin H$, then $ab \in H$. Show that H has index 2 in G.

126. Let D_n denote the group of all symmetries of a regular n-gon. Let a denote a rotation through $2\pi/n$ and let b denote a reflection. Show that

$$a^n = e, \quad b^2 = e, \quad bab^{-1} = a^{-1}.$$

Show that every element of D_n has a unique expression of the form a^i or $a^i b$ where $i \in \{0, 1, 2, \ldots, n-1\}$. (This exercise is designed to help with future questions which involve dihedral groups.)

127. Determine all subgroups of the dihedral group D_5 (of order 10).

128. Determine all subgroups of the dihedral group D_4 (of order 8) as follows:

(a) List the elements of D_4 and hence find all of the cyclic subgroups.

(b) Find two non-cylic subgroups of order 4 in D_4.

(c) Explain why any non-cylic subgroup of D_4, other than D_4 itself, must be of order 4 and, in fact, must be one of the two subgroups you have listed in the previous part.

129. Let G denote the group of rotational symmetries of a regular tetrahedron so that $|G| = 12$. Show that G has subgroups of order 1,2,3,4 and 12. (Harder) Show that G has no subgroup of order 6.

130. Let G be a group of order 841 (which is $(29)^2$). If G is not cyclic, show that every element g of G satisfies $g^{29} = 1$.

3.4.5 RSA cryptography

Cryptography is the science of keeping messages secret, by coding the messages so only the intended recipient can read them.

Example: simple substitution. The sender and recipient agree on a permutation of the 26 letters $a, b, \ldots, z$. The *key* (this permutation) must be exchanged and kept secret.

In a *public key cryptosystem*, the key for encyrption can be made public, but decryption is not possible (in a reasonable amount of time) except by the intended recipient.

The *RSA cryptosystem* developed in 1977 by Rivest, Shamir and Adleman is a public key system very widely used today (for example, for transactions over the internet and in ATM machines). It relies on the difficulty of factoring large numbers (typically more than 200 decimal digits) in a practical amount of time. (Currently, it takes many months of computing time to factor most numbers of 120-130 digits.) In contrast, large primes can be found efficiently using known primality tests.

Setting up an RSA cryptosystem

(1) Choose two large primes $p \neq q$ (typically more than 150 decimal digits).

(2) Compute $m = pq$.

(3) Compute $n = (p-1)(q-1)$.

(4) Choose an integer e with $1 < e < n$ such that $\gcd(e, n) = 1$.

(5) Compute d such that $ed \equiv 1 \pmod{n}$ (using Euclid's algorithm).

We represent our message units by elements of $\mathbb{Z}/m\mathbb{Z} = \{0, 1, \ldots, m-1\}$.

Then:

- the *public key* is: m, e. (These are made public and used to encrypt.)
- the *private key* is d. (This is kept secret and used to decrypt.)
- *encryption* is given by
$$X \mapsto X^e \pmod{m}$$
- *decryption* is given by
$$Y \mapsto Y^d \pmod{m}$$
(These powers can be computed efficiently by "repeated squaring".)

The original message is recovered since

$$(X^e)^d \equiv X^{ed} \equiv X \pmod{m}$$

if $ed \equiv 1 \pmod{n}$. This follows from

Theorem 3.4.8 (Euler)**.** *Let $m = pq$ where p, q are distinct primes. If*

$$N \equiv 1 \pmod{(p-1)(q-1)},$$

then

$$a^N \equiv a \pmod{m}$$

for all integers a.

Proof. Assume $N = 1 + k(p-1)(q-1)$ where $k \in \mathbb{Z}$. By Fermat's Little Theorem, if $p \nmid a$ then $a^{p-1} \equiv 1 \pmod{p}$. Thus

$$a^{k(p-1)(q-1)} \equiv (a^{p-1})^{k(q-1)} \equiv 1 \pmod{p}$$

and

$$a^N \equiv a \pmod{p}.$$

This equation also holds if $p \mid a$, since both sides are then $0 \pmod{p}$, so

$$a^N \equiv a \pmod{p} \text{ for all } a \in \mathbb{Z} \qquad (1).$$

Similarly,

$$a^N \equiv a \pmod{q} \text{ for all } a \in \mathbb{Z} \qquad (2).$$

Since $p \mid a^N - a$, $q \mid a^n - a$ and $\gcd(p, q) = 1$ it follows that $pq \mid a^N - a$ (by Exercise 9). Hence

$$a^N \equiv a \pmod{pq}$$

for all $a \in \mathbb{Z}$. □

Why is the RSA cryptosystem secure? To decrypt a message efficiently need to find $n = (p-1)(q-1)$ or equivalently p and q. But factoring $m = pq$ is not computationally feasible with current algorithms and technology if m is large (e.g. 300-400 decimal digits).

A very small example:

- Choose $p = 7, q = 13$.
- Then $m = 7 \times 13 = 91$ and $n = 6 \times 72$.
- Choose $e = 5$. (This is OK since $\gcd(5, 72) = 1$.)
- Then $d = 29$ since $5 \times 29 \equiv 145 \equiv 1 \pmod{72}$.
- Public key is: $m = 91$, $e = 5$.

If someone wants to send us the message:

$$23 \quad 85$$

they calculate

$$23^5 \equiv 4 \pmod{91}, \quad 85^5 \equiv 50 \pmod{91}$$

and send:

$$4 \quad 50.$$

To decrypt, we calculate

$$4^{29} \equiv 23 \pmod{91}, \quad 50^{29} \equiv 85 \pmod{91}$$

and recover the original message:

$$23 \quad 85.$$

3.4.6 Exercises

131. We set up an RSA cryptosystem using primes $p = 3$ and $q = 19$.

(a) Write down $m = pq$ and $n = (p-1)(q-1)$.

(b) Show that $e = 5$ is a suitable choice of encrypting key.

(c) With this encrypting key, encrypt the message '2 3 6 18'.

(d) Calculate the decrypting key d (for $e = 5$).

(e) With this decrypting key, decrypt the message '7 50'.

132. In this question we suppose that it has been agreed that the letters of the alphabet are encoded as

$$a = 1, b = 2, \ldots, z = 26 \quad \text{and} \quad \text{'space'} = 27$$

with no distinction made between upper and lower case. Messages are to be broken down into single letters which are then encrypted and sent in sequence.

Anne wants to be able to receive encrypted messages from her friends. She chooses two prime numbers: $p = 5$ and $q = 11$. The first part of her public key is then $m = 55$. She then calculates $n = \phi(m) = (p-1)(q-1)$. Knowing that $e = 3$ satisfies $\gcd(e, n) = 1$, she tells all her friends to encrypt messages for her using the numbers 55 and 3.

(a) Calculate n. (Note that, in practice, m would be chosen large enough that calculating n without knowing the prime factorisation of m would be impractical.)

(b) Irene wants to send Anne the message: 'hi there'. What is the encrypted sequence Irene should send?

(c) Anne receives the encrypted message 2 20 39 15 8 21 9. By first calculating d such that $ed \equiv 1 \pmod{n}$, decrypt the message.

133. (Harder) Let $(\mathbb{Z}/m\mathbb{Z})^\times$ be the set of elements of $\mathbb{Z}/m\mathbb{Z}$ with multiplicative inverses. Then the Euler phi function $\phi(m)$ is the number of elements of $(\mathbb{Z}/m\mathbb{Z})^\times$.

(a) Show that $(\mathbb{Z}/m\mathbb{Z})^\times$ is a group under multiplication modulo m.

(b) Show that $x^{\phi(m)} \equiv 1 \pmod{m}$ for all integers x such that $\gcd(x, m) = 1$. (This is a theorem of Euler.)

(c) Show that $\phi(p^k) = p^{k-1}(p-1)$ is p is a prime and $k \in \mathbb{N}$.

(d) Show that $\phi(pq) = (p-1)(q-1)$ if p and q are distinct primes.

Remark: If m has prime factorization $m = p_1^{k_1} \cdots p_n^{k_n}$ then

$$\phi(m) = p_1^{k_1-1}(p_1 - 1) \cdots p_n^{k_n-1}(p_n - 1).$$

3.5 Quotient groups

When we introduced cosets, it was to mimic the process of 'equivalence modulo n'. This leads to a set $\mathbb{Z}/n\mathbb{Z}$. But $\mathbb{Z}/n\mathbb{Z}$ has a natural addition inherited from $\mathbb{Z}$. We now want to see how we can mimic this in general.

Let G be a group and H a subgroup of G. A natural way to attempt to form a product of cosets Ha and Hb of H is to take their representatives and multiply them and then form the corresponding coset; that is, the product of Ha and Hb should be Hab. This is the way the addition of the integers is carried over to the addition of $\mathbb{Z}/n\mathbb{Z}$.

There is a possible problem, however. Suppose that $Ha = Hb$. Then the product of Hc with Ha should equal the product of Hc with Hb. That is, Hca should equal Hcb. Using (1) of Lemma 3.4.2, we can re-interpret this as saying that

$$\text{if } ab^{-1} \in H \text{ then } (ca)(cb)^{-1} = c\left(ab^{-1}\right)c^{-1} \in H.$$

For example, if $b = e_G$ in the above, then we require that $a \in H$ implies $cac^{-1} \in H$ for all $c \in G$. But this is not true for every subgroup H. For example, take $G = S_3$ and $H = \{(1), (12)\}$. Then $(12) \in H$ but

$$(123)(12)(123)^{-1} = (123)(12)(132) = (123)(13) = (23) \notin H.$$

So we cannot expect to make sense of multiplying all cosets of H in this way. We need to restrict the sort of subgroups we are dealing with.

3.5.1 Normal subgroups

Definition 3.5.1. Let G be a group and H a subgroup of G. We say that H is a ***normal*** subgroup of G if

$$h \in H \text{ and } g \in G \text{ implies } ghg^{-1} \in H.$$

There are two other simple variations of this condition. We introduce some notation. If S and T are subgroups of a group G then ST denotes $\{st : s \in S, t \in T\}$. If $S = \{s\}$ has only one element, we write sT rather than $\{s\}T$, and similarly. Thus the earlier notation for cosets fits into this notation.

Lemma 3.5.2. *Let G be a group and H a subgroup of G. The following are equivalent:*

(1) H is a normal subgroup of G.

(2) for every $g \in G$, $Hg = gH$.

(3) for every $g \in G$, $gHg^{-1} = H$.

Proof. Suppose that (1) is true. Let $k \in Hg$. Then $k = hg$ for some $h \in H$. Thus

$$k = hg = g\left(g^{-1}hg\right) = g\left(g^{-1}h(g^{-1})^{-1}\right) \in gH$$

as $g^{-1}h(g^{-1})^{-1} \in H$ because H is a normal subgroup. Thus $Hg \subseteq gH$. Similarly, $gH \subseteq Hg$ and so $gH = Hg$, proving that (2) holds.

Suppose now that (2) holds. Then

$$gHg^{-1} = (gH)g^{-1} = (Hg)g^{-1} = H$$

as $gH = Hg$. Thus (3) holds.

If (3) holds, it is easy to show that (1) holds.

□

Examples:

(1) The subgroup C_n of rotations in D_n is normal.

(2) The subgroup $\{(1), (123), (132)\}$ of S_3 is a normal subgroup. The subgroup $\{(1), (12)\}$ is not normal.

(3) $SL(n, F)$ is a normal subgroup of $GL(n, F)$.

(4) $\langle e_G \rangle$ and G are normal subgroups of G.

(5) Every subgroup of a commutative group is normal.

Restricting ourselves to normal subgroups gets rid of at least some of the problems associated with taking products of cosets. To see that it works in all cases, we have the following result.

Lemma 3.5.3. *Let H be a normal subgroup of a group G. Then*

$$HaHb = Hab.$$

Proof.

$$HaHb = H(aHa^{-1})ab = HHab = Hab.$$

We have used the facts that $aHa^{-1} = H$, because H is normal, and that $HH = H$ because H is a subgroup. □

3.5.2 Trivialising subgroups

We now want to complete the process of forming a new group of which the elements are the cosets of some subgroup.

Definition 3.5.4. Let G be a group and H a normal subgroup of G. Let G/H denote the set of right cosets of H in G. Define an operation $\diamond$ on G/H by $Ha \diamond Hb = HaHb$ (where $HaHb$ is interpreted as a product of sets within G).

Theorem 3.5.5. *If H is a normal subgroup of G, then the set G/H with the operation '$\diamond$' gives a group, called 'the* **quotient group** *of G by H'. The identity $e_{G/H}$ is the coset H and the inverse of the coset Hg is Hg^{-1}.*

Proof. By Lemma 3.5.3 $Ha \diamond Hb = HaHb$ is a coset and so the operation is well-defined. (Note that we did not *need* to use coset representatives to define it). The fact that it is associative is an easy deduction from the fact that the operation $*$ of G is associative.

Observe that $Ha \diamond H = H \diamond Ha = Ha$ and that $Ha \diamond Ha^{-1} = Ha^{-1} \diamond Ha = H$ and this proves the claims of the last sentence. □

Examples:

(1) $G = D_4 = \langle a, b : a^4 = 1, b^2 = 1, bab = a^{-1}\rangle$, $H = \langle a\rangle$

Since $|G| = 8$ and $|H| = 4$, there are two cosets of H and so two elements in G/H. We can quickly identify these as H and Hb. The multiplication table is

	H	Hb
H	H	Hb
Hb	Hb	H

(2) $G = D_4 = \langle a, b : a^4 = 1, b^2 = 1, bab = a^{-1}\rangle$, $H = \langle a^2\rangle$

In this case $|H| = 2$ and so there are 4 cosets of H. They are

$$H = \{1, a^2\}, \quad Ha = \{a, a^3\}, \quad Hb = \{b, a^2b\}, \quad Hab = \{ab, a^3b\}.$$

Then we can calculate the multiplication table as

	H	Ha	Hb	Hab
H	H	Ha	Hb	Hab
Ha	Ha	H	Hab	Hb
Hb	Hb	Hab	H	Ha
Hab	Hab	Hb	Ha	H

The group is isomorphic to D_2.

(3) A coset of $SL(n, F)$ in $GL(n, F)$ is the set of all matrices which have a given fixed determinant. When we multiply two cosets we obtain a coset corresponding to the product of these determinants. Thus there is one coset for each non-zero element of F and the product of cosets corresponds to the product of elements of F. Thus

$$GL(n, F)/SL(n, F) \cong (F \setminus \{0\}, \times)$$

where $(F \setminus \{0\}, \times)$ is the group of all non-zero elements of F under the operation of multiplication. We will soon see an easier way to establish this.

3.5.3 Homomorphisms

There is an obvious function f from a group G to a quotient group G/H given by

$$f : g \longrightarrow Hg.$$

This function satisfies $f(g * h) = f(g) \diamond f(h)$ and so looks rather like an isomorphism. But it is certainly not bijective because, for example, every element of H maps to a single element of G/H. We need a new definition.

Definition 3.5.6. A ***homomorphism*** between groups G and H is a function $f : G \to H$ such that

$$f(a *_G b) = f(a) *_H f(b) \qquad \text{for all } a, b \in G.$$

Thus homomorphisms are the functions between groups compatible with the group operations; these are analogous to linear transformations between vector spaces.

In particular, an isomorphism is just a bijective homomorphism.

Examples:

(1) Let G be a group with a normal subgroup H. The function $G \to G/H$ given by $g \mapsto Hg$ is a homomorphism.

(2) The function $\det : GL(n, F) \to F \setminus \{0\}$ given by taking the determinant of a matrix is a homomorphism.

(3) The function $\mathbb{C} \setminus \{0\} \to \mathbb{R} \setminus \{0\}$ given by $z \mapsto |z|$ is a homomorphism.

Associated with any homomorphism there are two very natural subgroups.

Definition 3.5.7. Let $f : G \to H$ be a homomorphism. Then

(1) $\{g \in G : f(g) = e_H\}$ is called the ***kernel*** of f, written $\ker f$.

(2) $\{f(g) : g \in G\}$ is called the ***image*** of f, written $\operatorname{im} f$.

Lemma 3.5.8. *Let $f : G \to H$ be a homomorphism. Then $\ker f$ is a normal subgroup of G and $\operatorname{im} f$ is a subgroup of H. Further, f is injective if and only if $\ker f = \{e_G\}$.*

Proof. We leave the first part of this as Exercise 140. For the second part, observe firstly that if f is injective then at most one element of G can be taken, by f, to e_H. Since we know that $f(e_G) = e_H$, it follows that $\ker f = \{e_G\}$.

Suppose that $\ker f = \{e_G\}$; we will show that f is injective. Let g_1 and g_2 be elements of G so that $f(g_1) = f(g_2)$. Then

$$f(g_1 g_2^{-1}) = f(g_1) f(g_2^{-1}) = f(g_1) f(g_2)^{-1} = e_H.$$

Thus $g_1 g_2^{-1} \in \ker f = \{e_G\}$ and so $g_1 g_2^{-1} = e_G$. Thus $g_1 = g_2$ and so f is injective.

□

Lemma 3.5.9. *Let $f : G \to G/H$ be the natural homomorphism given by $f(g) = Hg$. Then $\ker f = H$ and $\operatorname{im} f = G/H$.*

Proof. We leave this as Exercise 141. □

In this section so far we have started with normal subgroups, formed quotient groups, discovered a corresponding homomorphism and then re-discovered the original normal subgroup as the kernel of the homomorphism.

What if we start with a homomorphism from G? It has a kernel which is a normal subgroup of G, we can form a quotient by this normal subgroup and then form a second homomorphism from G to the quotient group. How does this compare with the original homomorphism? In fact they are essentially the same.

Theorem 3.5.10 (First Isomorphism Theorem)**.** *Let $f : G \to H$ be a homomorphism. Then*

$$G/\ker f \cong \operatorname{im} f.$$

In particular, if f is surjective then $G/\ker f \cong H$.

Proof. For brevity, set $\ker f = K$. We attempt to define a function $F : G/K \to \operatorname{im} f$ by $F(Kg) = f(g)$. The possible problem with this attempt is that we may have $Ka = Kb$ with $a \neq b$. It will not, in fact, be a problem if we know that $Ka = Kb$ implies that $f(a) = f(b)$.

So suppose that $Ka = Kb$. By Lemma 3.4.2, $ab^{-1} \in K$ and so $f(ab^{-1}) = e_H$. But then $f(a)f(b)^{-1} = e_H$ and so $f(a) = f(b)$. Thus F is a well-defined function.

Now

$$F(Ka \diamond Kb) = F(Kab) = f(ab) = f(a)f(b) = F(Ka)F(Kb).$$

so F is a homomorphism.

Now, $F(Kg) = e_H$ if and only if $f(g) = e_H$ if and only if $g \in K$ if and only if $Kg = e_{G/K}$. By Lemma 3.5.8, F is injective.

So F will be an isomorphism of G/K with the image of F. But we can easily check that $\operatorname{im} F = \operatorname{im} f$ and so the proof is complete. □

Examples:

(1) $\det : GL(n,F) \to F\setminus\{0\}$ is a homomorphism with kernel $SL(n,F)$ (see Example 3 of Section 2.5.2). So $SL(n,F)$ is a normal subgroup of $GL(n,F)$ and

$$GL(n,F)/SL(n,F) \cong F\setminus\{0\}.$$

(2) The function $\mathbb{C}\setminus\{0\} \to \mathbb{R}\setminus\{0\}$ given by $z \mapsto |z|$ is a homomorphism with image $\mathbb{R}_{>0}$ and kernel $\{z : |z| = 1\}$. Call the latter $\mathbb{S}^1$ (because it is a 'one-dimensional sphere'). Then

$$\frac{\mathbb{C}\setminus\{0\}}{\mathbb{S}^1} \cong \mathbb{R}_{>0}.$$

(3) The function $\mathbb{R} \to SO(2,\mathbb{R})$ which maps each real number a to the matrix representing the rotation, centered at the origin, through an angle of a can be checked to be a homomorphism. The function is surjective and its kernel is $2\pi\mathbb{Z}$, the subgroup of all integral multiples of 2π. Thus we have

$$\mathbb{R}/2\pi\mathbb{Z} \cong SO(2,\mathbb{R}).$$

3.5.4 Exercises

134. Show that the set of matrices

$$H = \left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} : ad \neq 0 \right\}$$

forms a subgroup of $GL(2,\mathbb{R})$. Show that the set of matrices

$$K = \left\{ \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} : b \in \mathbb{R} \right\}$$

forms a normal subgroup of H.

135. If H is a subgroup, prove that $HH = H$.

136. If K and L are normal subgroups of a group G, show that $K \cap L$ is also a normal subgroup of G.

137. Let G be a group and n a positive integer. If H is the only subgroup of G which has order n, show that H is a normal subgroup of G. (Hint: Use Exercise 115.)

138. Find all of the normal subgroups of D_4. (It will be a great help if you have done Exercise 128 of the previous section first.)

139. The *quaternion* group Q_8 is the subgroup of $GL(2,\mathbb{C})$ consisting of the matrices $\{\pm U, \pm I, \pm J, \pm K\}$ where

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad I = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad K = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

(a) Verify that

$$I^2 = J^2 = K^2 = -U, \quad IJ = K, JK = I, KI = J$$

and so that these 8 elements do give a subgroup.

(b) Find all of the cyclic subgroups of Q_8.

(c) Show that every subgroup of Q_8, except Q_8 itself, is cyclic.

(d) Are Q_8 and D_4 isomorphic?

140. Prove the first part of Lemma 3.5.8.

141. Prove Lemma 3.5.9.

142. If G is an abelian group, show that any quotient G/N is also abelian.

143. If G is a cyclic group, show that any quotient G/N is also cyclic.

144. Show that the cyclic group of order 8 has as homomorphic images the cyclic groups of order 2 and 4 as well as itself and the trivial group.

145. Let $\mathbb{R}$ denote the group of real numbers with the operation of addition and let $\mathbb{Q}$ and $\mathbb{Z}$ denote the subgroups of rational numbers and integers, respectively. Show that it is possible to regard $\mathbb{Q}/\mathbb{Z}$ as a subgroup of $\mathbb{R}/\mathbb{Z}$ and show that this subgroup consists exactly of the elements of finite order in $\mathbb{R}/\mathbb{Z}$.

146. Let H denote the subgroup of $D_8 =< a, b >$ generated by a^4. Write out the multiplication table of D_8/H.

3.6 Groups that can act

The importance of groups stems from the fact that they occur as 'generalised symmetries' of objects. This may occur in many ways apart from the more obvious geometrical ways. The common denominator is that the group will act as *permutations* of some set.

3.6.1 Definition and examples

Definition 3.6.1. An ***action of a group*** G ***on a set*** X is a function $G\times X \to X$, written $(g, x) \mapsto g\cdot x$, which combines elements $g \in G$, $x \in X$ to give an element $g \cdot x \in X$ satisfying the properties:

(a) $e \cdot x = x$ for all $x \in X$ where e is the identity in G,

(b) $(gh) \cdot x = g \cdot (h \cdot x)$ for all $g, h \in G$, $x \in X$.

We also call this a "G-action on X", and say that "G acts on X".

Examples:

(1) Let G be *any* group of bijections $g : X \to X$ with composition of functions as the operation. Then G acts on X by defining

$$g \cdot x = g(x) \text{ for all } g \in G, x \in X.$$

Many interesting examples arise in this way!

(2) S_n acts on $\{1, \dots, n\}$.

(3) The group of symmetries of a polygon P acts on P.

(4) $GL(n, F)$ acts on the vector space F^n.

(5) $GL(n, F)$ acts on the set of bases of the vector space F^n.

(6) $GL(n, F)$ acts on the set of all subspaces of the vector space F^n.

Another viewpoint: Given any G-action on X, if we fix $g \in G$ and let $x \in X$ vary, then we obtain a function $\phi(g) : X \to X$, taking $x \mapsto g \cdot x = \phi(g)(x)$.

Then $\phi(g)$ is a *permutation* of X (i.e. a bijection $X \to X$) since it has an inverse $\phi(g^{-1})$: for all $x \in X$

$$\phi(g^{-1})(\phi(g)(x)) = g^{-1} \cdot (g \cdot x) = (g^{-1} \cdot g) \cdot x = e \cdot x = x$$

using axioms (a) and (b) for a group action. Similarly $\phi(g)(\phi(g^{-1})(x)) = x$ for all $x \in X$.

This gives a function $\phi : G \to \operatorname{Sym} X$ where $\operatorname{Sym} X$ is the group of all permutations of X furnished with the operation of composition. Further, ϕ is a *homomorphism*, since

$$\phi(g_1g_2)(x) = (g_1g_2) \cdot x = g_1 \cdot (g_2 \cdot x) = \phi(g_1)(\phi(g_2)(x))$$

for all $x \in X$, i.e. $\phi(g_1g_2) = \phi(g_1)\phi(g_2)$.

Conversely, any homomorphism $\phi : G \to \operatorname{Sym} X$ gives a group action as defined above, with $g\cdot x = \phi(g)(x)$. (Exercise: check this.) Sometimes this is taken as the definition of a group action.

3.6.2 Orbits and stabilizers

Definition 3.6.2. Suppose that a group G acts on a set X.

(1) The ***orbit*** of $x \in X$ is
$$G \cdot x = \{g \cdot x : g \in G\}.$$
It is a subset of X.

(2) The ***stabilizer*** of $x \in X$ is
$$\operatorname{Stab} x = G_x = \{g \in G : g \cdot x = x\}.$$
It is a subset of G, in fact a subgroup of G.

Lemma 3.6.3. *Suppose that the group G acts on the set X. Then the stabilizer of $x \in X$ is a subgroup of G.*

Proof. Suppose that $g, h \in \operatorname{Stab} x$ so that $g \cdot x = x$ and $h \cdot x = x$. Then $h^{-1} \cdot h \cdot x = h^{-1} \cdot x$ and so $x = e_G \cdot x = h^{-1} \cdot x$. Thus
$$(gh^{-1}) \cdot x = g \cdot (h^{-1} \cdot x) = g \cdot x = x$$
and so $gh^{-1} \in \operatorname{Stab} x$. Thus $\operatorname{Stab} x$ is a subgroup of G. □

Examples:

(1) Consider $G = S_3$ acting on $\{1, 2, 3\}$. The stabilizer of 2 is $\langle(13)\rangle$; the orbit of 2 is $\{1, 2, 3\}$.

(2) Set $G = \langle(12)\rangle$, a subgroup of S_3. Then G acts on $\{1, 2, 3\}$. The orbit of 1 (or of 2) is $\{1, 2\}$. The orbit of 3 is $\{3\}$. The stabilizer of 1 or of 2 is the identity subgroup. The stabilizer of 3 is G itself.

(3) Consider $SO(2, \mathbb{R})$ acting on the plane $\mathbb{R}^2$. The orbit of any point is the circle, centered at the origin, which passes through that point. The stabilizer of any point, other than the origin, is the identity subgroup; the stabilizer of the origin is the whole group.

(4) Consider $SO(3, \mathbb{R})$ acting on $\mathbb{R}^3$. The orbit of any point is the sphere, centered at the origin, which passes through that point. The stabilizer of any point, other than the origin, is the set of rotations which have as axis the line passing through the origin and the chosen point.

Lemma 3.6.4. *Suppose that the group G acts on the set X. Then every element of X lies in one and only one orbit.*

Proof. Since $x \in G \cdot x$, every element of X lies in at least one orbit. Suppose that $z \in G \cdot x$ and $z \in G \cdot y$; we need to show that $G \cdot x = G \cdot y$; it will follow that different orbits can contain no element in common.

Since $z \in G \cdot x$ and $z \in G \cdot y$, we can write $z = a \cdot x$ and $z = b \cdot y$ for some $a, b \in G$. Let w be any element of $G \cdot x$, say $w = g \cdot x$. Then
$$w = g \cdot x = g \cdot (a^{-1}) \cdot z = (ga^{-1}) \cdot z = (ga^{-1}) \cdot (b \cdot y) = (ga^{-1}b) \cdot y \in G \cdot y.$$
Thus $G \cdot x \subseteq G \cdot y$. To show that $G \cdot y \subseteq G \cdot x$ and so that $G \cdot x = G \cdot y$, just reverse the roles of x and y in this argument. □

Next we come to the key relationship between orbits and stabilzers.

Theorem 3.6.5 (The Orbit-Stabilizer Relation)**.** *Suppose that the group G acts on the set X. Then, for each $x \in X$, there is a bijective correspondence between the (left) cosets of* $\operatorname{Stab} x$ *and the elements of the orbit $G \cdot x$. Thus the size of any orbit $|G \cdot x|$ is equal to the index $|G : Stab(x)|$ of the stabilizer $Stab(x)$ in G.*

In particular, if G is finite then
$$|G| = |G \cdot x| \cdot |\operatorname{Stab} x|,$$
and so $|G \cdot x|$ divides $|G|$.

Proof. Set $H = \operatorname{Stab} x$. Let $\mathcal{C}$ denote the set of left cosets of H in G. We will prove the theorem by setting up mutually inverse correspondences between $\mathcal{C}$ and $G \cdot x$.

Define $\sigma : \mathcal{C} \to G \cdot x$ by $\sigma(gH) = g \cdot x$. Since σ is defined by choosing a representative of the coset gH we must check that its value is the same for all possible choices of representative. So suppose that $aH = bH$. Then $b^{-1}a \in H$ and so $b^{-1}a \cdot x = x$. But then $a \cdot x = b \cdot x$ and so $\sigma(aH)$ does not depend on the choice of representative.

Define $\tau : G \cdot x \to \mathcal{C}$ by $\tau(g \cdot x) = gH$. Again, if $a \cdot x = b \cdot x$, then $b^{-1}a \in \operatorname{Stab} x = H$ and so $aH = bH$. Thus the function τ is well-defined.

Now σ and τ are easily seen to be inverse functions. So they are both bijective. Hence $|G \cdot x| = |G : H|$ is the number of cosets of H in G.

But G is the disjoint union of these cosets and each has $|H|$ elements. Hence $|G| = |G \cdot x| \cdot |H|$ if $|G|$ is finite (as in the proof of Lagrange's theorem). □

Examples:

Consider D_6 as the symmetry group of a regular hexagon $ABCDEFA$. Let G be that subgroup of D_6 which consists of the identity together with a rotation through π together with reflections in the line AD and the line joining the midpoint of EF to the midpoint of BC. (Alternatively, colour two opposite edges blue and the other four edges red. Then G is the group of colour symmetries).

Then G is a group of order 4 acting on the vertices $\{A, B, C, D, E, F\}$. The stabilizer of A is the subgroup generated by the reflection in the line AD. The orbit of A is $\{A, D\}$. The stabilizer of B is the identity subgroup. The orbit of B is $\{B, C, E, F\}$.

3.6.3 Exercises

147. Any subgroup G of S_4 acts on the set $\{1, 2, 3, 4\}$ in a natural way. For each choice of G given below, describe the orbits of the action and the stabilizer of each point.

(a) $G = \langle (123) \rangle$　(b) $G = \langle (1234) \rangle$
(c) $G = \langle (12), (34) \rangle$　(d) $G = S_4$
(e) $G = D_4 (= \langle (1234), (13) \rangle)$.

148. Let $X = \mathbb{R}^3$ and let $v \neq 0$ be a fixed element of X. Show that

$$\alpha \cdot x = x + \alpha v \quad (x \in X, \alpha \in \mathbb{R})$$

defines an action of the additive group of the real numbers on X. Give a geometrical description of the orbits.

149. Let G be the subgroup of S_{15} generated by the three permutations

$$(1,12)(3,10)(5,13)(11,15) \quad (2,7)(4,14)(6,10)(9,13)$$
$$(4,8)(6,10)(7,12)(9,11).$$

Find the orbits in $S = \{1, \ldots, 15\}$ under the action of G. Deduce that G has order which is a multiple of 60.

150. If a group G of order 5 acts on a set X with 11 elements, must there be an element of the set X which is left fixed by every element of the group G? What if G has order 15 and X has 8 elements?

3.7 Groups acting on themselves

One of the most powerful techniques of finite group theory involves the consideration of groups acting on themselves. That is, the set X on which the group G acts is just the set of elements of G. We shall consider two examples and apply them to prove results about the groups themselves.

3.7.1 Left multiplication

Any group G acts on $X = G$ by *left multiplication*:

$$g \cdot x = gx \text{ for all } g \in G, x \in X = G,$$

where the right hand side is group multiplication. This gives a group action since

(a) $e \cdot x = ex = x$

(b) $(gh) \cdot x = (gh)x = g(hx) = g \cdot (h \cdot x)$.

Similarly G acts on itself by *right multiplication*: $g \cdot x = xg^{-1}$. (Exercise.)

As usual, this group action gives us a homomorphism $\alpha : G \to \operatorname{Sym} G$. This describes how left multiplication permutes the rows of the multiplication table for the group G:

Example: Consider the group $D_2 = \langle a, b : a^2 = e, b^2 = e, ab = ba \rangle$ and the action by left multiplication. The group D_2 has 4 elements $\{e, a, b, ab\}$ and its multiplication table is:

	e	a	b	ab
e	e	a	b	ab
a	a	e	ab	b
b	b	ab	e	a
ab	ab	b	a	e

Left multiplication permutes the group elements as follows:

$$\begin{array}{cc}
\alpha(e) & \text{the identity permutation} \\
\alpha(a) & \begin{pmatrix} e & a & b & ab \\ a & e & ab & b \end{pmatrix} \\
\alpha(b) & \begin{pmatrix} e & a & b & ab \\ b & ab & e & a \end{pmatrix} \\
\alpha(ab) & \begin{pmatrix} e & a & b & ab \\ ab & b & a & e \end{pmatrix}
\end{array}$$

If we wish, we could re-name the elements of D_2 via $(e, a, b, ab) \to (1, 2, 3, 4)$ and we would then have an isomorphism between D_2 and a subgroup of S_4:

$$\begin{array}{rr}
e \to (1) & a \to (12)(34) \\
b \to (13)(24) & ab \to (14)(23)
\end{array}$$

The same method gives the following general result.

Theorem 3.7.1 (Cayley's Theorem)**.** *Every group G is isomorphic to a subgroup of a permutation group. If G has finite order n then G is isomorphic to a subgroup of S_n.*

Proof. The action of G on itself by left multiplication gives a homomorphism $\alpha : G \to \operatorname{Sym} G$ where $\alpha(g)(x) = g \cdot x = gx$. We show that α is injective (or 1-1) so that G is isomorphic to $\operatorname{im} \alpha$ which is a subgroup of the permutation group $\operatorname{Sym} G$. If G has n elements then $\operatorname{Sym} G$ will be isomorphic to S_n so that the second sentence follows.

We need to show that α is injective. Suppose that, for some $a, b \in G$, $\alpha(a) = \alpha(b)$. Then, for any $x \in G$, $\alpha(a)(x) = \alpha(b)(x)$; that is, $ax = bx$. But then $a = b$ and so α is injective. □

3.7.2 Conjugation

Each group G acts on itself by *conjugation*:

$$g \cdot x = gxg^{-1} \text{ for all } g \in G, x \in G.$$

This is a group action since

(a) $e \cdot x = exe^{-1} = x$

(b) $(gh) \cdot x = (gh)x(gh)^{-1} = (gh)x(h^{-1}g^{-1}) = g(hxh^{-1})g^{-1} = g \cdot (h \cdot x)$.

The case where a group acts on itself by conjugation has particular importance and acquires its own notation rather than the general notation of 'orbit' and 'stabilizer'.

Definition 3.7.2. Let G be a group, and let $x \in G$. Then

(1) a ***conjugate*** of x is any element of the form gxg^{-1} where $g \in G$.

(2) the ***conjugacy class*** of x is the set of all elements of G of the form gxg^{-1} for $g \in G$.

(3) the ***centralizer*** of x, written $C_G(x)$, is the subgroup of G of all elements g satisfying $gxg^{-1} = x$ or, equivalently, $gx = xg$.

Observe that a conjugacy class is just an orbit and a centralizer is just a stabilizer when we consider G acting on itself by conjugation. Thus we can translate the orbit-stabilizer relation into this context.

Theorem 3.7.3. *Let G be a finite group. Then the number of elements in a conjugacy class is equal to the number of cosets of the centralizer of any element of the conjugacy class. Thus the number of conjugates of $g \in G$ is $|G|/|C_G(g)|$, so divides $|G|$.*

Examples: We shall work out the conjugacy classes of $G = D_4 = \langle a, b : a^4 = b^2 = e, bab^{-1} = a^{-1} \rangle$.

Before we proceed with the details we mention briefly the general line of approach. We pick an element that we have not yet assigned to a conjugacy class. We note any obvious conjugates and so get a lower bound on the size of the conjugacy class. We then try to find elements in the centralizer of the element. This will give us a lower bound on the size of the centralizer and so, using Theorem 3.7.3, an upper bound on the size of the conjugacy class. With luck the lower and upper bounds on the size of the conjugacy class are the same; if not, then we need to find more conjugates or more elements in the centralizer to refine our bounds.

Firstly, $\{e\}$ is a conjugacy class.

Let us calculate the conjugacy class containing a. Observe that $a \in C_G(a)$ and so $\langle a \rangle \leq C_G(a)$. Thus, as $|\langle a \rangle| = 4$, $C_G(a)$ has at least 4 elements. Thus it has at most $8/4 = 2$ cosets. So the conjugacy class of a has at most 2 elements. The relation $bab^{-1} = a^{-1}$ above tells us that a^{-1} is a conjugate of a. So this conjugacy class is $\{a, a^{-1}\}$. Note that $a^{-1} = a^3$.

We will now calculate the conjugacy class containing a^2. Observe that $a \in C_G(a)$ and so $\langle a \rangle \leq C_G(a)$. Also, $ba^2b^{-1} = (bab^{-1})(bab^{-1}) = (a^{-1})^2 = a^2$ and so $b \in C_G(a)$. Hence $C_G(a^2) = G$ and the conjugacy class containing a^2 has one element. So this conjugacy class is $\{a^2\}$.

We will now calculate the conjugacy class containing b. Observe that $b \in C_G(b)$. In the previous paragraph we showed that $ba^2b^{-1} = a^2$ and so that $ba^2 = a^2b$. This also implies that $a^2b(a^2)^{-1} = b$ and so $a^2 \in C_G(b)$. Thus $C_G(b)$ contains at least two different non-identity elements and so must have order at least 4 (recall that its order must divide 8, by Lagrange's theorem). So the conjugacy class containing b can have at most 2 elements. Since

$$aba^{-1} = b^2aba^{-1}b^2 = b(bab^{-1})a^{-1} = b(a^{-1})a^{-1} = ba^{-2} = ba^2 = a^2b$$

we have that a^2b is a conjugate of b. So this conjugacy class is $\{b, a^2b\}$.

A similar argument shows that the final class is $\{ab, a^3b\}$.

Thus the conjugacy classes of D_4 are

$$\{e\}, \{a^2\}, \{a, a^3\}, \{b, a^2b\}, \{ab, a^3b\}.$$

Note how little hard calculation we had to do in order to find these conjugacy classes.

3.7.3 Some consequences for group theory

The elements of the group which form a conjugacy class on their own clearly play a special role.

Definition 3.7.4. Let G be a group. The ***centre*** of G, written $Z(G)$, is the set of elements $x \in G$ such that $gx = xg$ for all $g \in G$.

Examples:

(1) The centre of D_4 is $\{e, a^2\}$ (this needs some checking).

(2) The element e_G always lies in the centre of G.

(3) The centre of S_3 is just $\{(1)\}$ (this needs some checking).

(4) The centre of $GL(n, F)$ is the subgroup of scalar matrices $\{aI : a \in F^*\}$ (this needs a lot of checking).

Lemma 3.7.5. *The centre of a group G is a normal subgroup of G.*

Proof. We leave this as Exercise 156. □

As we have seen, the centre of a group may contain only the identity element. But there is one case where we can guarantee that it contains more than this.

Theorem 3.7.6. *Let p be a prime and let G be a group of order p^n for some integer $n \geq 1$. Then the centre of G contains more than the identity element.*

Proof. Recall that in studying the centre, we are looking at elements which form a conjugacy class on their own.

Write G as a disjoint union of conjugacy classes:

$$G = C_1 \cup C_2 \cup \cdots \cup C_k$$

and recall that the size of each conjugacy class has order dividing $|G|$. Now group the one element conjugacy classes together to form the centre Z. Thus we have

$$G = Z \cup C_l \cup C_{l+1} \cup \cdots \cup C_k$$

where we have renamed the classes so that $C_l, \ldots, C_k$ are exactly the classes with more than one element. Thus, looking at sizes,

$$|G| = p^n = |Z| + |C_l| + |C_{l+1}| + \cdots + |C_k|.$$

But $|C_l|, \ldots, |C_k|$ are all divisors of $|G|$ and are all bigger than one. Thus they are all powers of p and, in particular, are all multiples of p. Thus the equation becomes: $p^n = |Z| + pm$ where $|C_l| + \cdots + |C_k| = pm$. So p divides $|Z|$ and so $|Z|$ must be bigger than 1. □

One application of this result is the following.

Theorem 3.7.7. *If p is prime, then every group of order p^2 is isomorphic to $\mathbb{Z}/p^2\mathbb{Z}$ or $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.*

A proof is outlined in Exercises 157, 159 and 160.

We will finish this section with a partial converse to Lagrange's theorem. We know that the order of any element in a group G divides $|G|$, but if n divides $|G|$ there may be no element of order n in G. However we have:

Theorem 3.7.8 (Cauchy's Theorem)**.** *Let G be a finite group of order divisible by a prime number p. Then G has an element of order p.*

Proof. We need to find $x \in G$, with $x \neq e$, such that $x^p = e$.

Consider the set

$$X = \{(x_1, x_2, \ldots, x_p) : x_i \in G, x_1 x_2 \ldots x_p = e\},$$

i.e. all ordered p-tuples of elements in G whose product is the identity.

First we determine the size of X. We can choose *arbitrary* elements $x_1, \dots, x_{p-1}$ from G. Then $(x_1, \dots, x_{p-1}, x_p) \in X$ if and only if $x_p = (x_1 \dots x_{p-1})^{-1}$. Hence $|X| = |G|^{p-1}$. In particular,

$$p \text{ divides } |X|, \tag{1}$$

since p divides $|G|$ and $p - 1 \geq 1$.

Now the group $\mathbb{Z}/p\mathbb{Z}$ acts on X by cyclically permuting the p-tuples: for $m = 0, 1, \dots, p-1$, define

$$[m] \cdot (x_1, x_2, \dots, x_p) = (x_{m+1}, x_{m+2}, \dots, x_p, x_1, \dots, x_m).$$

Each orbit for this action has 1 or p elements, since the size of each orbit divides $|\mathbb{Z}/p\mathbb{Z}| = p$. Clearly, the orbit of $(e, e, \dots, e)$ has 1 element. If every other orbit has p elements then

$$|X| = \text{ sum of orbit sizes}$$

would not be divisible by p contradicting (1).

Hence there is $(x_1, x-2, \dots, x_p) \in X$ other than $(e, e, \dots, e)$ which is *fixed* by every element of $\mathbb{Z}/p\mathbb{Z}$. Then $x_1 = x_2 = \ldots = x_p = x \neq e$, and $x^p = e$. □

This can be used to prove:

Theorem 3.7.9. *If p is an odd prime, then each group of order $2p$ is isomorphic to the cyclic group C_{2p} or the dihedral group D_p.*

We leave this as Exercise 162.

3.7.4 Exercises

151. Find the conjugacy classes in the quaternion group described in Exercise 139.

152. Find the conjugates of

(a) (123) in S_3;

(b) (123) in S_4;

(c) (1234) in S_4;

(d) (1234) in S_n where $n \geq 4$;

(e) $(12 \dots m)$ in S_n where $n \geq m$.

153. (Harder) Let τ be a permutation in S_n. Suppose that $\sigma = (12 \dots k)$. Show that $\tau\sigma\tau^{-1} = (\tau(1)\tau(2) \dots \tau(k))$. What is the result if σ is replaced by a general element of S_n? Use this to describe the conjugacy classes of S_n.

154. Suppose that g and h are conjugate elements of a group G. Show that $C_G(g)$ and $C_G(h)$ are conjugate subgroups of G.

155. Determine the centralizer in $GL(3, \mathbb{R})$ of the following matrices:

(a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

156. Prove Lemma 3.7.5.

157. Suppose that G is a group with centre Z and so that G/Z is a cyclic group. Show that there must be an element $h \in G$ so that every element of G can be written in the form $g = h^n z$ with $n \in \mathbb{Z}$ and $z \in Z$. Deduce that G is commutative.

158. Describe the finite groups with exactly one or exactly two or exactly three conjugacy classes (the last of these is harder).

159. If p is a prime, use Theorem 3.7.6 and Exercise 157 to show that a group of order p^2 is commutative.

160. (Harder) If p is a prime, use the previous exercise to help show that each group of order p^2 is isomorphic to either $\mathbb{Z}/p^2\mathbb{Z}$ or $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.

161. If p is a prime number, show that any group of order $2p$ must have a subgroup of order p and that this subgroup must be normal.

162. Show that, up to isomorphism, there are two groups of order $2p$ for every prime number p.

3.8 Keeping our distance; isometries

3.8.1 Definition and basic properties

We want to consider ‘symmetries’ of the standard Euclidean plane. To avoid confusion with the vector space $\mathbb{R}^2$, we shall denote the Euclidean plane by $\mathbb{E}^2$. The points of $\mathbb{R}^2$ and of $\mathbb{E}^2$ are the same but $\mathbb{E}^2$ is a *geometric* object, equipped with the usual notions of distance and angle in Euclidean geometry. Further all points of $\mathbb{E}^2$ are considered identical; there is no ‘origin’.

We begin with a precise version of what we mean by a ‘symmetry’.

Definition 3.8.1. A function $f : \mathbb{E}^2 \to \mathbb{E}^2$ is an *isometry* if, for any pair of points $P, Q \in \mathbb{E}^2$, we have $|PQ| = |f(P)f(Q)|$.

Here $|PQ|$ denotes the distance $d(P, Q)$ between the points P and Q.

Examples:

(1) A reflection in a line;

(2) a rotation about a point;

(3) a translation;

(4) a glide reflection.

Since many of you may not know what the last is, here is a definition.

Definition 3.8.2. A ***glide reflection*** is the combination of a translation followed by a reflection in an axis parallel to the direction of translation.

Lemma 3.8.3. *If an isometry fixes two points, it fixes all points of the line on which they lie. If an isometry fixes three points which do not all lie on a line, then it fixes all of $\mathbb{E}^2$.*

Proof. The proof involves some elementary Euclidean geometry. For example, if a point lies on the line between two fixed points then it is uniquely determined by its distances from those two points, and so must also be fixed.. We leave the details to Exercise 169. □

We need some information about multiplication of different kinds of isometries. Once we have used this to establish Theorem 3.8.5, it will become easier to calculate the result of other combinations.

Lemma 3.8.4. *(1) Let σ_1 and σ_2 be reflections in axes L_1 and L_2. The product $\sigma_1\sigma_2$ is either,*

a) a rotation about the point of intersection of L_1 and L_2, with angle of rotation twice the angle between L_1 and L_2 , if L_1 and L_2 intersect;

b) a translation in a direction perpendicular to L_i with a magnitude equal to twice the distance between L_1 and L_2, if L_1 and L_2 are parallel.

(2) The product of three reflections in parallel axes is a further reflection.

(3) The product of three reflections in axes which are not parallel and which do not intersect in a point is a glide reflection.

Proof. Perhaps the easiest (but not the most elegant) way to do (1) and (2) is to choose co-ordinate axes suitably and to represent the reflections by specific functions of the co-ordinates. For example, to consider reflections in two parallel axes, choose co-ordinate axes so that one of the axes of reflection is the x-axis and the other is the line $y = a$. Then the two reflections are given by the functions

$$f : (x, y) \mapsto (x, -y) \qquad \text{and} \qquad g : (x, y) \mapsto (x, a/2 - y).$$

The composite $g \circ f$ of these two functions is then the function $(x, y) \mapsto (x, a/2 + y)$ which is easily seen to be a translation. We leave the rest of (1) and (2) as Exercise 170.

We turn to (3); call the product $\rho_1\rho_2\rho_3$. Consider the reflections ρ_2, ρ_3. If their axes are parallel, then this product is a translation τ and we must consider $\rho_1\tau$. If their axes intersect, then the product of these two reflections is a rotation about the point of intersection P of the axes. But this rotation can be represented by any product of reflections which both have axes passing through P and lying at an angle which is the same as the angle between the two original axes. Thus we can replace $\rho_2\rho_3$ by a product $\rho_2'\rho_3'$ of reflections so that the axis of ρ_2' is parallel to the axis of ρ_1. Then $\rho_1\rho_2'$ is a translation τ' and we must consider $\tau'\rho_3$.

Thus we must consider either a translation followed by a reflection or a reflection followed by a translation. The two cases are, not surprisingly, very similar, and we shall consider only the second. Choose axes so that the axis of the reflection is the y-axis. In co-ordinates the reflection is then given by $(x, y) \mapsto (-x, y)$. The translation will be of the form $(x, y) \mapsto (a + x, b + y)$. The composite is therefore

$$(x, y) \mapsto (-x, y) \mapsto (a - x, b + y).$$

But we can represent this composite in the form

$$(x, y) \mapsto (a - x, y) \mapsto (a - x, y) + (0, b)$$

which shows that it is the combination of a reflection $(x, y) \mapsto (a - x, y)$ in the line $x = a/2$ with a translation parallel to the y axis and so to the line of reflection. It is therefore a glide reflection.

□

3.8.2 What isometries are there?

It is surprisingly easy to describe all isometries of the plane. We do it by looking at successively smaller sets of fixed points.

Theorem 3.8.5. *The set of fixed points of an isometry is one of the following:*

(1) All of $\mathbb{E}^2$; in this case, the isometry is the identity.

(2) A line in $\mathbb{E}^2$; in this case, the isometry is the reflection in that line.

(3) A single point; in this case, the isometry is a rotation about that point and can be expressed as the product of two reflections.

(4) empty; in this case, the isometry is either a translation and can be expressed as the product of two reflections or a glide reflection and can be expressed as the product of three reflections.

Proof. There is nothing to prove if the isometry ϕ fixes all of $\mathbb{E}^2$. So suppose that it does not. Then, by Lemma 3.8.3, it must fix no more than a line of points. Suppose that it fixes exactly a line; call this line L. Let P, Q be points on L and let R be a point which is not on L. Then $|RP| = |\phi(R)\phi(P)| = |\phi(R)P|$ and, similarly, $|RQ| = |\phi(R)\phi(P)| = |\phi(R)Q|$. A little school geometry will now tell you that $\phi(R)$ is either R or the reflection of R in L. We have assumed, however, that ϕ fixes points only on L and so $\phi(R) \neq R$. Let ρ be the reflection in L. Then $\phi(R) = \rho(R)$. Thus $\rho^{-1} \circ \phi$ fixes all of P, Q and R. By Lemma 3.8.3, $\rho^{-1} \circ \phi$ fixes $\mathbb{E}^2$ and so is the identity. Hence $\phi = \rho$ and ϕ is a reflection, as claimed.

Suppose now that ϕ does not fix any line. By Lemma 3.8.3, it can fix no more than a point. Suppose that it fixes a point, P say. Let Q be any point different from P and let L be the perpendicular bisector of Q and $\phi(Q)$. Since Q and $\phi(Q)$ must have the same distance from the fixed point P, it follows that P must lie on L. Also, the reflection in L, call it ρ say, must send Q to $\phi(Q)$. Thus $\rho^{-1} \circ \phi$ fixes both P and Q and so, by what we have done so far, is either the identity or a reflection. We can easily see that it cannot be the identity if ϕ is to fix no more than one point. Thus $\rho^{-1} \circ \phi$ is a reflection, say ρ_1, and so $\phi = \rho \circ \rho_1$. Since the axes of ρ and ρ_1 intersect at P, it follows from Lemma 3.8.4 that ϕ is a rotation about P.

Suppose finally that ϕ fixes no points. Then if P is any point, we have $P \neq \phi(P)$. Let L be the perpendicular bisector of P and $\phi(P)$ and let ρ be the reflection in L. Then $\rho^{-1} \circ \phi$ fixes P and so $\rho^{-1} \circ \phi$ is either a rotation about P or a reflection fixing P. Thus, using the result of the last case, ϕ is a product of at most three reflections in this case.

It remains to show that ϕ is either a translation or a glide reflection. If $\rho^{-1} \circ \phi$ is a reflection fixing P then ϕ is a product of two reflections and so must be a translation, by Lemma 3.8.3. If $\rho^{-1} \circ \phi$ is a rotation about P then ϕ is the product of three reflections. If the axes of all three of these reflections are parallel then, by Lemma 3.8.3, ϕ is a reflection, which is impossible. If not, then the axes cannot meet in a point, since that point would then be fixed by ϕ. Thus the axes are not all parallel and do not meet in a point and so, by Lemma 3.8.3, ϕ is a glide reflection. □

So we have described *all* isometries on $\mathbb{E}^2$. We can use this to help us describe the group of all isometries.

Definition 3.8.6. The group of all isometries of $\mathbb{E}^2$ will be denoted by $\mathcal{I}$.

Before we go on, we make a few observations about conjugates in $\mathcal{I}$. If ϕ is an isometry, and X is the set of fixed points of ϕ then it is not too hard to check that, for some other isometry σ, the set of fixed points of $\sigma\phi\sigma^{-1}$ is $\sigma(X)$. This shows us immediately that if ϕ is a reflection then any conjugate of ϕ is also a reflection because the set of fixed points of the conjugate must also be a line. Similarly a conjugate of a rotation is another rotation. By this method we know that a conjugate of a translation is either a translation or a glide reflection. But a translation is a product of two reflections and then its conjugate will be the product of two conjugates of reflections and so itself the product of two reflections. Thus its conjugate is another translation. Similarly, a conjugate of a glide reflection is a glide reflection. We leave to Exercise 168 the question of determining all of the conjugacy classes exactly.

Theorem 3.8.7. *(1) The set of translations forms a normal subgroup $\mathcal{T}$ of $\mathcal{I}$.*

(2) Let P be a point of $\mathbb{E}^2$; the set of isometries of $\mathcal{I}$ which fix P forms a subgroup $\mathcal{O}_P$.

(3) If $P, Q \in \mathbb{E}^2$ then $\mathcal{O}_P$ and $\mathcal{O}_Q$ are conjugate subgroups of $\mathcal{I}$. In particular, they are isomorphic.

(4) Let P be a point of $\mathbb{E}^2$; every element of $\mathcal{I}$ can be uniquely expressed as a product of a translation and an isometry fixing P.

(5) Let P be a point of $\mathbb{E}^2$; there is a surjective homomorphism $\pi_P : \mathcal{I} \to \mathcal{O}_P$.

Proof. (1) It is easy to check that $\mathcal{T}$ is a subgroup of $\mathcal{I}$. We have just shown, in the discussion preceding the statement of the theorem, that a conjugate of a translation is another translation. Thus $\mathcal{T}$ is a normal subgroup.

(2) is an easy check.

(3) Choose any $\phi \in \mathcal{I}$ satisfying $\phi(P) = Q$. For example, ϕ could be a translation. We claim that $\phi\mathcal{O}_P\phi^{-1} = \mathcal{O}_Q$. Let $\psi \in \phi\mathcal{O}_P\phi^{-1}$; say $\psi = \phi\nu\phi^{-1}$ with $\nu \in \mathcal{O}_P$. Then

$$\psi(Q) = \phi\nu\phi^{-1}(Q) = \phi\nu(P) = \phi(P) = Q.$$

Thus $\psi \in Q$ and so $\phi\mathcal{O}_P\phi^{-1} \subseteq \mathcal{O}_Q$; the reverse inclusion is similar.

(4) Suppose that $\phi \in \mathcal{I}$. Set $Q = \phi(P)$. Then there is a translation τ which takes P to Q. Also $\tau^{-1}\phi(P) = \tau^{-1}(Q) = P$. If we set $\tau^{-1}\phi = \mu$, then $\mu \in I$ and $\phi = \tau\mu$ with $\tau \in \mathcal{T}$ and μ an isometry fixing P. It remains to show the uniqueness. Suppose that $\phi = \tau_1\mu_1$ with $\tau_1 \in T$ and $\mu_1 \in \mathcal{O}_P$. Then $\tau\mu = \tau_1\mu_1$ and so $\tau_1^{-1}\tau = \mu_1\mu^{-1}$. But $\tau_1^{-1}\tau$ is a translation and $\mu_1\mu^{-1}$ fixes the point P. Thus they must both be the identity. That is, $\tau = \tau_1$ and $\mu = \mu_1$.

(5) If $\phi \in \mathcal{I}$, we can write $\phi = \tau\mu$ with $\tau \in \mathcal{T}$ and $\mu \in \mathcal{O}_P$. Define $\pi_P(\phi) = \mu$. This is a well-defined function by the previous part. We must show that it is a homomorphism. Suppose that $\phi_1 = \tau_1\mu_1$ and $\phi_2 = \tau_1\mu_1$. Then

$$\phi_1\phi_2 = \tau_1\mu_1\tau_1\mu_1 = \tau_1 \left(\mu_1\tau_2\mu_1^{-1}\right) \mu_1\mu_2$$

and $\mu_1\mu_2 \in \mathcal{O}_P$ and $\tau_1 \left(\mu_1\tau_2\mu_1^{-1}\right) \in \mathcal{T}$ (using part (1)). Thus

$$\pi_P(\phi_1\phi_1) = \mu_1\mu_2 = \pi_P(\phi_1)\pi_P(\phi_2)$$

and π_P is a homomorphism. □

We have shown that $\mathcal{I}$ has a normal subgroup $\mathcal{T}$ and a subgroup $\mathcal{O}_P$ (not unique) which satisfy $\mathcal{T} \cap \mathcal{O}_P = \{e_I\}$ and $\mathcal{T}\mathcal{O}_P = \mathcal{I}$. This kind of structure is quite common; we say that $\mathcal{I}$ is a *semi-direct product* of $\mathcal{T}$ by $\mathcal{O}_P$.

The fact that $\mathcal{O}_P$ and $\mathcal{O}_Q$ are conjugate via a translation shows that the homomorphism π_P need not depend on P. Thus we shall simply regard it as a homomorphism

$$\pi : \mathcal{I} \longrightarrow \mathcal{O}$$

where $\mathcal{O}$ is the set of elements in $\mathcal{I}$ which fix a given (unnamed) point. Note that, if we identify $\mathbb{E}^2$ with $\mathbb{R}^2$ in such a way that the point P becomes the origin of $\mathbb{R}^2$ then all of these elements are linear transformations of $\mathbb{R}^2$ and so $\mathcal{O} \cong O(2, \mathbb{R})$.

3.8.3 Classification of finite symmetry groups

Theorem 3.8.8. *The only finite groups of isometries of* $\mathbb{E}^2$ *are the cyclic groups and the dihedral groups.*

Proof. We have already seen that the groups described are finite groups of isometries; we must show that there are no more. Suppose that G is a finite group of isometries of $\mathbb{E}^2$.

We claim first that G has a fixed point. Suppose that $G = \{g_1, \dots, g_n\}$ and let P be any point of $\mathbb{E}^2$. Consider the set of points $S = \{g_1(P), \dots, g_n(P)\}$. If $g \in G$ then

$$g(S) = g\big(\{g_1(P), \dots g_n(P)\big) = \{gg_1(P), \dots, gg_n(P)\} = S$$

as $\{gg_1, \dots, gg_n\}$ is just G, but listed in a different order. Thus the set S of points is left fixed by each element of G and so the centroid (centre of gravity) of S is also left fixed by each element of G. Thus we have the required fixed point; call it O.

By Theorem 3.8.5, G consists of reflections and rotations. It is clear that the rotations form a subgroup H.

Let $k \in H$ be the rotation in H of least possible positive angle θ and let $h \in H$ be an arbitrary element of H, with angle η. Then we can write $\eta = n\theta + \zeta$ with $n \in \mathbb{Z}$ and $0 \le \zeta < \theta$. So hk^{-n} will have angle $\eta - n\theta = \zeta$. But $hk^{-n} \in H$ and this will contradict the choice of θ as the least possible angle of any element of H unless $\zeta = 0$. Thus $\zeta = 0$ and so $hk^{-n} = e_H$. Hence $h = k^n$ and $H = \langle k \rangle$, showing that H is cyclic. If $G = H$, G is thus the cyclic group.

Otherwise, G must contain reflections. By Lemma 3.8.4, the product of any two reflections is a rotation. Thus, if we take two reflection ρ_1 and ρ_2 in G, then $\rho_2\rho_1^{-1} \in H$. Hence $\rho_2 \in H\rho_1$. Hence all cosets containing a reflection coincide; in fact they are just the complement of H in G. That is, the only coset of H, apart from H itself, is the complement of H in G and consists of all of the reflections. Also, if ρ is any reflection and σ is a rotation through θ, it is an easy exercise to check that $\rho\sigma\rho^{-1}$ is a rotation through $-\theta$. Thus $\rho\sigma\rho^{-1} = \sigma^{-1}$. It is now relatively easy to check that G is a dihedral group.

□

Thus the natural symmetry groups we discussed at the beginning of this section are the only possibilities.

3.8.4 Isometries of Euclidean space $\mathbb{E}^n$

Let $\mathbb{E}^n$ denote n-dimensional Euclidean space, that is $\mathbb{R}^n$ together with its usual Euclidean notion of distance: the distance between points x and y in $\mathbb{R}^n$ is $d(x,y) = ||x-y|| = \sqrt{(x_1-y_1)^2 + \ldots + (x_n-y_n)^2}$. This is just the usual distance defined in terms of the dot product on $\mathbb{R}^n$ since $||x|| = \sqrt{x \cdot x}$.

An **isometry** of $\mathbb{E}^n$ is a distance preserving function, i.e. a function $f : \mathbb{R}^n \to \mathbb{R}^n$ such that $d(x,y) = d(f(x), f(y))$ for all x, y in $\mathbb{R}^n$.

Examples:

(1) If $A \in O(n)$ is an orthogonal matrix (i.e. $A^TA = I$), then the linear transformation $f_A : \mathbb{R}^n \to \mathbb{R}^n$ defined by $f_A(x) = Ax$ is an isometry.

(2) If $b \in \mathbb{R}^n$, then *translation by* b is an isometry $t_b : \mathbb{R}^n \to \mathbb{R}^n$, $t_b(x) = x + b$. Further the inverse of t_b is t_{-b}.

(3) Compositions of isometries are isometries.

In fact, every isometry of $\mathbb{E}^n$ is a composition of a translation and an orthogonal transformation.

Lemma 3.8.9. *Let f be an isometry of $\mathbb{E}^n$ such that $f(\mathbf{0}) = \mathbf{0}$. Then f is a linear transformation, $f(x) = Ax$ where $A \in O(n)$ is an orthogonal matrix, and $x \in \mathbb{R}^n$ is a column vector.*

Proof. We have a function $f : \mathbb{R}^n \to \mathbb{R}^n$ such that
(i) $f(\mathbf{0}) = \mathbf{0}$ and (ii) $||f(x) - f(y)|| = ||x - y||$ for all x, y in $\mathbb{R}^n$.

Step 1. f preserves dot products.

For all x, y we have

$$||x-y||^2 = (x-y)\cdot(x-y) = x\cdot x - 2x\cdot y + y\cdot y = ||x||^2 - 2x\cdot y + ||y||^2. \tag{1}$$

Hence we also have

$$||f(x)-f(y)||^2 = ||f(x)||^2 - 2f(x)\cdot f(y) + ||f(y)||^2. \tag{2}$$

Since f preserves distances, we have $||x-y|| = ||f(x)-f(y)||$ for all x, y. Taking $y = \mathbf{0}$ in this gives $||x|| = ||x - \mathbf{0}|| = ||f(x) - f(\mathbf{0})|| = ||f(x) - \mathbf{0}|| = ||f(x)||$ for all x, and similarly $||y|| = ||f(y)||$ for all y. Using these facts together with (1) and (2) then gives $x \cdot y = f(x)\cdot f(y)$ for all x, y.

Step 2. If $\{e_1, \ldots, e_n\}$ is the standard basis for $\mathbb{R}^n$ then $\{f(e_1), \ldots, f(e_n)\}$ is an orthonormal basis for $\mathbb{R}^n$.

Since $e_1, \ldots, e_n$ are orthonormal vectors, step 1 implies that $f(e_1), \ldots f(e_n)$ are also orthonormal vectors. Since orthonormal vectors are linearly independent and $\dim \mathbb{R}^n = n$ it follows that $\{f(e_1), \ldots, f(e_n)\}$ is an orthonormal basis for $\mathbb{R}^n$.

Step 3. f is a linear transformation.

Let x be any vector in $\mathbb{R}^n$. Then we can write

$$x = x_1e_1 + \ldots + x_ne_n$$

where $x_i = x \cdot e_i$, since $\{e_1, \ldots, e_n\}$ is an orthonormal basis. Similarly

$$f(x) = (f(x)\cdot f(e_1))f(e_1) + \ldots + (f(x)\cdot f(e_n))f(e_n)$$

since $\{f(e_1), \ldots, f(e_n)\}$ is an orthonormal basis. But $f(x)\cdot f(e_i) = x \cdot e_i = x_i$ by step 1. Hence we have

$$f(x_1e_1 + \ldots + x_ne_n) = x_1f(e_1) + \ldots + x_nf(e_n)$$

for all $x_1, \ldots x_n$. It follows f is a linear transformation.

Step 4. The matrix A of f with respect to the standard basis satisfies $A^TA = I$, so A is orthogonal.

The standard matrix of A has columns $f(e_1), \ldots, f(e_n)$ (we regard these as column vectors). Then the (i,j) entry in the matrix A^TA is

$$f(e_i)^Tf(e_j) = f(e_i)\cdot f(e_j) = e_i \cdot e_j = \delta_{ij}.$$

So A^TA is the identity matrix. □

Corollary 3.8.10. *Every isometry f of $\mathbb{E}^n$ has the form $f(x) = Ax + b$ where $A \in O(n)$ and $b \in \mathbb{R}^n$. (We write $f = (A, b)$ for short.)*

Proof. If $f(\mathbf{0}) = b$, then $t_b^{-1} \circ f = t_{-b} \circ f$ is an isometry fixing $\mathbf{0}$. Hence, by the previous lemma, $f(x) - b = t_{-b} \circ f(x) = Ax$ for some $A \in O(n)$. □

3.8.5 Exercises

163. Let $\mathcal{I}_+$ denote the subset of $\mathcal{I}$ consisting of all translations together with all rotations. Show that $\mathcal{I}_+$ is a subgroup of $\mathcal{I}$.

164. Show that $\mathcal{I}_+$ has index 2 in $\mathcal{I}$; deduce that $\mathcal{I}_+$ is normal. The isometries in $\mathcal{I}_+$ are called *orientation preserving.*

165. Show that an isometry is orientation preserving precisely if it is a product of an even number of reflections.

166. Identify $\mathbb{E}^2$ with the complex plane. Then each point can be represented by a complex number. Show that every isometry can be represented in the form $z \mapsto e^{i\theta}z + u$ or the form $z \mapsto e^{i\theta}\overline{z} + u$ for some real number θ and some complex number u. Show that the former type correspond to orientation-preserving isometries.

167. Let D_∞ be the set of isometries consisting of all translations of $\mathbb{R}^2$ which are parallel to the x-axis and through an integer distance together with all reflections in a line $x = n/2$ for n an integer. Show that D_∞ is a subgroup of the group of all isometries. Show that D_∞ acts on the x-axis and find the orbit and stabilizer of $(1, 0), (\frac{1}{2}, 0), (\frac{1}{3}, 0)$.

168. Describe the conjugacy classes in the group $\mathcal{I}$.

169. Prove Lemma 3.8.3.

170. Complete the proof of parts (1) and (2) of Lemma 3.8.4.

4 Hints and answers to Exercises

4.1 Modular Arithmetic

1. (a) No, for instance $\mathbb{Z}$ itself is nonempty but has no smallest element.

(b) No: the interval $(0, 1] \subset \mathbb{R}$ is bounded below, but has no smallest element.

(c) No: the set $\{x \in \mathbb{Q} \mid x > 0\} \subset \mathbb{Q}$ is bounded below, but has no smallest element.

2. (a) $q = 8$, $r = 1$

(b) $q = 9$, $r = 5$

(c) $q = -5$, $r = 2$

3. If $a \mid b$ then $b = ka$ for some integer $k \in \mathbb{Z}$. If $c \mid d$ then $d = cl$ for some integer $l \in \mathbb{Z}$. Then $bd = (ka)(lc) = (kl) \times (ac)$ and $lk \in \mathbb{Z}$. Hence $ac \mid bd$.

4.

5. If k, ℓ are positive integers with $k\ell = 1$ what can you say about k and ℓ?

6. Assume $a = q_1 d + r_1 = q_2 d + r_2$ where $0 \leq r_1 < d$ and $0 \leq r_2 < d$.

7. (a) 7 (b) 15 (c) 143

(d) 8 (e) 1

8. (a) $\gcd(27, 33) = 3 = 5 \times 27 + (-4) \times 33$.

(b) $\gcd(27, 32) = 1 = 11 \times 32 + (-13) \times 27$.

(c) $\gcd(312, 317) = 13 = 5 \times 377 - 6 \times 312$.

9. We can find integers x, y such that $ax + by = 1$.

10.

11.

12. (a) $3 \equiv 42 \pmod{13}$

(b) $2 \equiv -20 \pmod{11}$

(c) $26 \not\equiv 482 \pmod{14}$

(d) $-2 \equiv 933 \pmod{5}$ as 935 is a multiple of 5.

(e) $-2 \equiv 933 \pmod{11}$ as 935 is a multiple of 11.

(f) As 935 is a multiple of 5 and 11, it is a multiple of 55, hence $-2 \equiv 933 \pmod{55}$.

13. (a) $6 \pmod{14}$

(b) $7 \pmod{9}$

(c) $0 \pmod{11}$

(d) $933 \equiv -2 \equiv 53 \pmod{55}$

(e) $5 \pmod{10}$

(f) $57102725 \equiv 5 + 7 + 1 + 0 + 2 + 7 + 2 + 5 \equiv 29 \equiv 2 \pmod{9}$

14. (a) $24 \times 25 \equiv 3 \times 4 \equiv 12 \pmod{21}$

(b) $0 \pmod{210}$

(c) $7 \pmod 9$

(d) $5 \pmod{11}$

(e) $1 \times (2 \times 3) \times (4 \times 5) \times 6 \equiv -1 \times -1 \times -1 \equiv -1 \equiv 6 \pmod 7$

(f) $1 \times 2 \times 3 \times \ldots \times 20 \times 21 \equiv (2 \times 11) \times (3 \times \ldots \times 10) \times (12 \times \ldots \times 21) \equiv 0 \times (\ldots) \times (\ldots) \equiv 0 \pmod{22}$

15.

16.

17. We have that $326 \equiv (3+2+6) \equiv 11 \equiv (1+1) \equiv 2 \pmod 9$, and $4471 \equiv (4+4+7+1) \equiv (16) \equiv 7 \pmod 9$. Therefore $(326 \times 4471) \equiv (2 \times 7) \equiv 14 \equiv 5 \pmod 9$. But $1357546 \equiv (1+3+5+7+5+4+6) \equiv 31 \equiv 4 \pmod 9$. Therefore $326 \times 4471 \neq 1357546$.

18. Consider the equation modulo 8.

19.

20. (a) $\mathbb{Z}/7\mathbb{Z}$ has the set of multiplicative units $\{1,2,3,4,5,6\}$

(b) $\mathbb{Z}/8\mathbb{Z}$ has the set of multiplicative units $\{1,3,5,7\}$

(c) $\mathbb{Z}/12\mathbb{Z}$ has the set of multiplicative units $\{1,5,7,11\}$

(d) $\mathbb{Z}/13\mathbb{Z}$ has the set of multiplicative units $\{1,2,\ldots,12\}$

(e) $\mathbb{Z}/15\mathbb{Z}$ has the set of multiplicative units $\{1,2,4,7,8,11,13,14\}$

21. (a) 32 in $\mathbb{Z}/27\mathbb{Z}$ has inverse 11 as $1 \equiv 11 \times 32 - 13 \times 27 \equiv 11 \times 32 \pmod{27}$.

(b) 32 in $\mathbb{Z}/39\mathbb{Z}$ has inverse 11.

(c) 17 in $\mathbb{Z}/41\mathbb{Z}$ has inverse $-12 \equiv 29 \pmod{41}$.

(d) 18 in $\mathbb{Z}/33\mathbb{Z}$ has no inverse as $3 = \gcd(18, 33)$.

(e) 200 has inverse 41 in $\mathbb{Z}/911\mathbb{Z}$.

22. 52

23.

24.

25. The hardest part is to show that it is closed under multiplication and division.

26. For example, $(\sqrt[3]{2})^2$ is not in the set. Set $\alpha = \sqrt[3]{2}$ and suppose that $\alpha^2 = a + b\alpha$ for some $a, b \in \mathbb{Q}$. Then

$$2 = \alpha^3 = a\alpha + b\alpha^2 = a\alpha + b(a + b\alpha) = ab + (a + b^2)\alpha.$$

It would follow that $\alpha = (2 - ab)/(a + b^2)$ and so that α is rational, which we know to be false.

If we take the set of all $a + b\sqrt[3]{2} + c(\sqrt[3]{2})^2$ with $a, b, c \in \mathbb{Q}$ then we do obtain a field.

27. The first part is routine. Using this, you can obtain every element as a power of $[3]_7$, for example.

$$[1]_7 = [3]_7^6, [2]_7 = [3]_7^2, [3]_7 = [3]_7^1, [4]_7 = [3]_7^4, [5]_7 = [3]_7^5, [6]_7 = [3]_7^3.$$

28. x does not have an inverse.

29. Showing closure under addition and subtraction is relatively straightforward. It is reasonably easy to convince yourself of closure under multiplication. The problem is with division. (I do not recommend attempting detailed proofs of all of the axioms.)

To see that

$$c_{-k}t^{-k} + c_{-k+1}t^{-k+1} + \cdots + c_0 + c_1t + \cdots + c_st^s + \ldots$$

has an inverse, assume that $c_{-k} \neq 0$ and write the above as $c_kt^{-k}g$ where g is a power series involving only non-negative powers of t and with constant term 1. Now show g has an inverse in this set of power series.

30. For example, $x^2 + 1$ has no root in the field.

31. We need to show that, whatever the value of p, there is a polynomial with coefficients in $\mathbb{F}_p$ but no root in $\mathbb{F}_p$. Try

$$x(x - [1]_p)(x - [2]_p) \dots (x - [p - 1]_p) + 1.$$

4.2 Linear Algebra

32. A typical element of $U + W$ has the form $u_1 + w_1$ where $u_1 \in U$ and $w_1 \in W$. If α is a scalar, then $\alpha(u_1 + w_1) = \alpha u_1 + \alpha w_1$. Since U and W are subspaces, $\alpha u_1 \in U$ and $\alpha w_1 \in W$. Hence $\alpha(u_1 + w_1) \in U + W$. We have shown that $U + W$ is closed under scalar multiplication. The argument that it is closed under addition is similar.

33. Use the definition of subspace, as in the previous question.

34. If neither U_1 nor U_2 is V, then neither can lie inside the other. Consider an element of V of the form $u_1 + u_2$ with $u_1 \in U_1$ but $u_1 \notin U_2$ and $u_2 \in U_2$ but $u_2 \notin U_1$. Does it lie in U_1 or in U_2?.

35. (a) linearly independent, not a basis;

(b) linearly independent, not a basis;

(c) linearly dependent, not a basis.

36. (a) yes;

(b) yes;

(c) no;

37. Use the definition of linear independence.

38. Show that every element of $M_{m\times n}(F)$ can be written uniquely as a linear combination of the suggested basis elements.

39. $\{1, x, x^2, x^3, \dots\}$ is an infinite linearly independent set.

40. 9.

41. For example, $g(A_1 + A_2) = (A_1 + A_2)B = A_1B + A_2B = g(A_1) + g(A_2)$.

42. $\begin{bmatrix} 2 & 3 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$.

43. The matrix is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. The matrix with respect to the new basis is

$$\frac{1}{ad - bc}\begin{bmatrix} ad + bc & 2cd \\ -2ab & -(ad + bc) \end{bmatrix}.$$

44. The nullity is 1; the rank is 2.

45. The nullity is 1; the rank is 2.

46. Briefly, f is surjective if and only if the range of f equals V if and only if the rank of f equals the dimension of V if and only if the nullity of f is zero.

47. Use the outline given in the lemma.

48. $(X - 2)(X + 1)$, $X^2 + X - 1$, $X^3 - 1$, $(X - 1)^3$.

49. The minimal polynomial of either matrix is $(X - 1)(X - 2)$. The characteristic polynomials are $(X - 1)^2(X - 2)^2$ and $(X - 1)^3(X - 2)$, respectively.

50. Check by direct computation that $A^2 - 2A - 8I_3 = 0$. Since this polynomial has distinct roots, it must be the minimal polynomial of the matrix. If the inverse a^{-1} exists, we can multiply the equation above by A^{-1} to obtain $A - 2I_3 - 8A^{-1} = 0$. This shows that the inverse does exist and how to calculate it.

51. If the minimal polynomial has non-zero constant term, use the idea of the last question to show there is an inverse. If the minimal polynomial has zero constant term then it is of the form $m(X) = Xp(X)$ for some polynomial $p(X)$. Since $p(f) \neq 0$, there is a vector v such that $w = p(f)(v) \neq 0$. But $f(w) = f(p(f)(v)) = m(f)(v) = 0$. If f had an inverse f^{-1} we could deduce that $w = f^{-1}(f(w)) = f^{-1}(0) = 0$ which is a contradiction.

52. You can do this by taking a power of an appropriate matrix. But the 'slick' way to do it is to use the linear transformation f which corresponds to A, using the standard basis $\{e_1, \ldots, e_n\}$. Note that $f(e_i)$ can be written as a linear combination of those e_j with $j < i$. Now show that $f^2(e_i)$ can be written as a linear combination of those e_j with $j < i - 1$ and then work out what happens for $f^n(e_i)$.

53. Write any $v \in V$ as $v = (1/2)(v_1 + v_2)$ where $v_1 = f(v) + v$ so that $f(v_1) = v_1$ and $v_2 = f(v) - v$ so that $f(v_2) = -v_2$. The diagonal matrix will have 1s (corresponding to the elements of a basis of the first space) and -1s (corresponding to the elements of a basis of the second space) on the diagonal.

54. Read the proof of Lemma 2.2.3 before doing this question.

55. Check the definitions of 'null-space' and 'f-invariant' before attempting the question.

56. Firstly show that 0 is the only eigenvalue of differentiation. If differentiation were represented by a diagonal matrix, what would that matrix be?

57. Show that the eigenvalues of f are 0 and 1. Write $v \in V$ as $v = f(v) + (v - f(v))$ to show that V is the sum of the two eigenspaces. Show that the corresponding eigenspaces are complementary and then pick an appropriate basis.

58. For the first part you need to show that if v is an eigenvector of f then $g(v)$ is also an eigenvector, with the same eigenvalue. In the second part, use the fact that F is algebraically closed to show that the restriction of g to the eigenspace of f has an eigenvalue and so an eigenvector.

59.

$$\begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}, \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

60. Recall that $J(a, n)$ represents the Jordan block with size n and diagonal entry a.

(a) one $J(0, 2)$ plus two $J(-1, 2)$ **or** one $J(0, 2)$ plus one $J(-1, 2)$ plus two $J(-1, 1)$;

(b) two $J(3, 2)$ plus one $J(3, 1)$ **or** one $J(3, 2)$ plus three $J(3, 1)$;

(c) two $J(0, 3)$ plus one $J(0, 1)$ **or** one $J(0, 3)$ plus two $J(0, 2)$ **or** one $J(0, 3)$ plus one $J(0, 2)$ plus two $J(0, 1)$ **or** one $J(0, 3)$ plus four $J(0, 1)$;

(d) two $J(1, 2)$ plus two $J(-1, 2)$ **or** two $J(1, 2)$ plus one $J(-1, 2)$ plus two $J(-1, 1)$ **or** one $J(1, 2)$ plus two $J(1, 1)$ plus two $J(-1, 2)$ **or** one $J(1, 2)$ plus two $J(1, 1)$ plus one $J(-1, 2)$ plus two $J(-1, 1)$;

61. (a) no;

(b) yes;

(c) yes.

62. This is similar to the 3×3 case given in the notes. Remember that the minimal polynomial divides the characteristic polynomial and has the same roots (possibly with different multiplicity). Then use Lemma 2.2.20. If the characteristic polynomial has the form $(X - a)^4$ and the minimal polynomial has the form $(X - a)^2$ then there are two possibilities which we cannot distinguish without more information.

63. Set D to be the diagonal part of J and set $N = J - D$. Then Exercise 5 shows that N is nilpotent. For the second part, choose a basis so that f is represented by a JCF matrix J. Write $J = D + N$ as in the first part. Then let d and n be the linear transformations corresponding to the matrices D and N.

64. Show that $JD = DJ$ first (you can easily reduce it to the case where J is just a single Jordan block.) Then $JN = NJ$ follows quickly. The last part is now immediate.

65.

66. (a) $\sqrt{19}$.

(b) $\sqrt{\frac{1}{5} - \frac{6}{4} + \frac{11}{3} - \frac{6}{2} + 1} = \sqrt{\frac{11}{30}}$.

(c) $\sqrt{30}$.

67. If $u = v$ then the claim gives $\|2u\| = 4\|u\|$, which is wrong. Try proving

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$$

by expanding into inner products.

68. We do the second part. The first part is similar but easier. We have, for all $u, v \in V$,

$$(fg(u), v) = (f\,(g(u))\,, v) = (g(u), f^*(v)) = (u, g^*\,(f^*(v))) = (u, g^*f^*(v)).$$

But also, by the definition of adjoint, we have

$$(fg(u), v) = (u, (fg)^*(v)).$$

Thus, for all $u, v \in V$,

$$(u, (fg)^*(v) - g^*f^*(v)) = 0$$

and so, for all $v \in V$, $(fg)^*(v) - g^*f^*(v) = 0$. That is, $(fg)^* = g^*f^*$.

69. Show firstly that, because the columns of A represent one orthonormal basis in terms of another, they are an orthonormal basis for the column space.

Suppose that $f(v) = \lambda v$ for some $v \in V$ and scalar λ. For the first part, $(f(v), v) = (v, f^*(v)) = (v, f(v))$ gives $\lambda(v, v) = \overline{\lambda}(v, v)$ and so $\lambda = \overline{\lambda}$.

The second part is similar but start with $(f(v), f(v))$.

70. For the first part, use the fact that if $f(v) = 0$ then $(v, f^*(u)) = (f(v), u) = 0$ for all $u \in V$. This shows that the nullspace of f is orthogonal to the range of f^*. You can now use Lemmas 1.2.8 and 1.4.4 to show the statements of the second and third sentences. For the final part, use the fact that the matrix of f^* is equal to the complex conjugate transpose of the matrix of f.

71.

72.

73. What is the adjoint of a triangular matrix? For the case of a unitary matrix, what does the inverse of a triangular matrix look like?

74. Suppose that $\{v_1, \ldots, v_n\}$ is an orthonormal basis of V. Suppose that

$$w = a_1v_1 + \ldots a_nv_n \text{ and that } f(v_i) = \sum_j b_{ji}v_j.$$

Then $(f(v_i), w) = \sum_j b_{ji}a_j$. Set $w_1 = \sum_k c_kv_k$ with

$$c_k = \sum_j b_{jk}v_j.$$

For the uniqueness, note that if w_2 also satisfies the conditions, then $(v, w_1) = (v, w_2)$ for all $v \in V$.

75. Apply the definition of linear transformation.

76. For all $u, v \in V$, we have,

$$(u, (f^*)^*(v)) = (f^*(u), v) = \overline{(v, f^*(u))} = \overline{(f(v), u)} = (u, f(v)).$$

Now deduce that $f = (f^*)^*$.

77. (a) Use the hint.

(b) Observe that $(g(w), u) = (w, g^*(u)) = (w, g(u))$ and that $(w, g(u)) = \overline{(g(u), w))} = (g(u), w))$ to show that $2(g(u), w) = 0$. Now deduce that g is zero.

(c) As in the previous part but deduce that the real part of $(g(u), w)$ is 0.

(d) If $(g(iu), w)$ is imaginary for all $u, w \in V$ then $(g(u), w) = i(g(u), w)$ is both real and imaginary and so zero.

(e) Not much left to do now.

78. Apply the definitions carefully. This is not a difficult question but *is* rather subtle.

79. The comments for the previous question also apply here. You also need to use the previous question.

80. You can solve this by writing the matrix as

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

Then multiply this by its transpose and equate the result to the identity. It will be useful to observe that if $a^2 + b^2 = 1$ then there is an angle θ so that $a = \cos\theta$ and $b = \sin\theta$.

81. Show that $AA^* = UDD^*U^* = UD^*DU^* = A^*A$.

82. Use the definition of normal and Lemma 2.3.10.

83. Find a diagonal matrix similar to A, take the square root of that and use that to find a square root of A.

84. Try firstly to find square roots for 2 × 2 Jordan matrices. In fact

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

has no square root.

85. (a) No; different eigenvalues.

(b) No; different eigenvalues.

(c) Yes.

86. No; they may have different eigenvalues.

87. Think of f represented by a diagonal matrix. If $f^2 = f^3$ then $a^2 = a^3$ for each eigenvalue a. Thus $a = 0$ or $a = 1$ and so $a = a^2$. Hence $f = f^2$. Since each eigenvalue is real, f is self-adjoint.

88. Choose a diagonal matrix A to represent f. Find a polynomial $p(X)$ so that $p(\lambda) = \overline{\lambda}$ for each eigenvalue λ of A (this may require a technique known as *interpolation*). Then $p(A) = A^*$ and so $p(f) = f^*$.

89. If f, g are normal, they can be simultaneously diagonalised to two matrices A and B say. Then the matrix of f^* is A^* and A^* is diagonal. Thus $A^*B = BA^*$ and so $f^*g = gf^*$.

90. (a) $(f^*f)(f^*f)^* = (f^*f)(f^*f)$ and $(f^*f)^*(f^*f) = (f^*f)(f^*f)$.

(b) Use the Spectral Theorem and group together equal eigenvalues.

(c) Let B denote the matrix for f with respect to the basis used in the previous part. Write B as an $m \times m$ block matrix and then use the fact that this matrix commutes with the matrix found in the previous part.

(d) The first part is immediate. For the second part, firstly recall that $A_i = \lambda_i I_{m_i}$. Thus, if $\lambda_i \neq 0$ then $B_i^* = \lambda_i B_i^{-1}$ and the result follows. If $\lambda_i = 0$, then $B_i^* B_i$ is the zero matrix. Check that this implies that $B_i = 0$ and so again the result follows.

(e) The matrix of f has the required property and hence so also does f.

91. Up to you.

4.3 Groups

92. (a) There are 6 rotations through π about the line joining midpoints of opposite edges.

(b) There are 3 rotations through π about the lines joining midpoints of opposite faces.

(c) There are 6 rotations through $\pm\pi/2$ about the lines joining midpoints of opposite faces.

(d) There are 8 rotations through $\pm 2\pi/3$ about the main diagonals.

(e) There is the identity rotation.

Yes. There are many reflections.

93. (a) There are 8 rotations through $\pm 2\pi/3$ about the lines joining a vertex to a midpoint of the opposite side.

(b) There are 3 rotations through π about the line joining midpoints of opposite edges.

(c) There is the identity rotation.

94. This is intended as a fairly open-ended question; I do not expect a complete answer.

We can give a complete answer, however, in a fairly slick fashion as follows. One possible n-dimensional cube is the set of all points $(x_1, \ldots, x_n)$ in $\mathbb{R}^n$ which satisfy $|x_i| \leq 1$. The vertices of the cube are then the 2^n points for which the entries are either 1 or -1. If we do this, then we can describe *all* of the symmetries via matrices which have exactly one non-zero entry, consisting of either 1 or -1 in each row or column. Such a matrix is *isometric* and so represents a rigid motion of the plane. It is also easy to check that it takes any vertex of the cube to some other vertex.

It is not too hard to argue that there are $2^n n!$ such matrices. Of these $2^{n-1}n!$ (those with determinant $+1$) will represent rotational symmetries. Compare this number with Question 1.

95. I leave this to you; the answer depends to some extent on how you draw your letters.

96.

97. (a) No: there are no 'inverses';

(b) yes;

(c) no: for example, the zero matrix has no multiplicative inverse.

98. For the first part, use the axioms. For the second part, note that the product of two reflections having the same centre, is a rotation. You can check this, for example by finding the product of reflections in the x and y axes.

99. Use $w = yx^{-1}$ and $z = x^{-1}y$. For uniqueness, suppose that, also, $w_1 x = y$. Then $wx = w_1 x$ and so $wx(x^{-1}) = w_1 x(x^{-1})$. Then $w(xx^{-1}) = w_1(xx^{-1})$ and so $we_G = w_1 e_G$. That is, $w = w_1$. A similar argument works to show the uniqueness of z. For the final sentence, choose x and y which do not commute to show that the answer is no.

100. Use the axioms.

101. As for the previous question but n is no longer fixed.

102. If X is another inverse of x then $xX = e_G$. Multiply this equation on the left by x^{-1} to show that $X = x^{-1}$.

103. (a) (1)(264)(35);

(b) (1356724);

(c) (1456)(2)(3).

104. For example,

$$h\left(g(x)\right) = \frac{1}{\frac{x-1}{x}} = \frac{x}{x-1} = k(x)$$

so that $hg = k$. The simplest way to do this question is to construct a multiplication table:

	f	g	h	i	j	k
f	g	i	k	f	h	j
g	i	f	j	g	k	h
h	j	k	i	h	f	g
i	f	g	h	i	j	k
j	k	h	g	j	i	f
k	h	j	f	k	g	i

The operations is associative because it is composition of functions. The identity is i and it is easy to check from the table that every element has an inverse.

105. The product of a rotation of the plane with a translation is another rotation, through the same angle but with a different centre. One slick way to see this is to identify the plane with the complex plane (Argand diagram). Then a rotation through angle θ and centre given by the complex number a is represented by $z \mapsto (z-a)e^{i\theta} + a$. If we combine the rotation about the origin with a translation which takes the origin to a point represented by b, the result has the form $z \mapsto ze^{i\theta} + b$. It is now not hard to check that the latter is a rotation through θ with centre represented by the complex number $b/(1-e^{i\theta})$.

The product of two rotations is a translation if the two rotations have angles which add to a multiple of 2π and otherwise is another rotation.

106. (a) 12;

(b) 10;

(c) 2;

(d) infinite order;

(e) 10, 5, 20, 10;

(f) 12, 2, 4.

107. Show that $g^m = e_G$ if and only if $(g^{-1})^m = e_G$.

108. If a has order p and b has order q, try the pq'th power of ab.

109. Try the product of reflections with parallel but distinct axes.

110. Checking the orders of A and B is straightforward computation.
$AB = \begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix}$. For the order of AB, show (by induction?) that

$$(AB)^n = (-1)^n \begin{bmatrix} 1 & 0 \\ -n & 1 \end{bmatrix}.$$

111. 2 and any divisor of n.

112. Use

$$g^2h^2 = (gh)^2 = ghgh$$

and cancel a g on the left and an h on the right.

113. (a) no: the subgroup is not closed under taking 'inverses';

(b) no: the product of rotations with different centres is not a rotation;

(c) yes.

114. $\langle[0]\rangle$, $\langle[1]\rangle = \langle[5]\rangle = \langle[7]\rangle = \langle[11]\rangle = \mathbb{Z}/12\mathbb{Z}$, $\langle[2]\rangle = \langle[10]\rangle$, $\langle[3]\rangle = \langle[9]\rangle$, $\langle[4]\rangle = \langle[8]\rangle$, $\langle[6]\rangle$.

115. Apply the check in Lemma 2.3.1.

116. Check firstly that $f(g_1 * g_2) = f(g_1) * f(g_2)$. To show that f is bijective, it is easier to find an inverse for f; try replacing g by g^{-1}.

117. For (1), (2) apply the homomorphism f to the group axioms. For (3), first show that $f(g^k) = f(g)^k$ for all integers $k > 0$.

118. Note that $SO(2)$ consists of the rotations of $\mathbb{R}^2$. Let $R_\theta \in SO(2)$ be the matrix for rotation by θ. Show that $R_\theta \mapsto e^{i\theta}$ gives an isomorphism from $SO(2) \to S^1$.

119. If $n = mk$, then we can draw a regular m-gon inside the regular n-gon. Use this to show that a subgroup of D_n can be identified with (that is, is isomorphic to) the symmetries of a regular m-gon.

120. For (a) look at orders of elements. For (b), (c) try looking at "kth roots" of group elements.

121. The order of $H \cap K$ must divide both 7 and 29.

122. Each right coset

$$H \begin{bmatrix} x_0 & y_0 \\ 0 & 1 \end{bmatrix} = \{ \begin{bmatrix} zx_0 & zy_0 \\ 0 & 1 \end{bmatrix} : z > 0 \}$$

can be identified with a half-line through a point (x_0, y_0) with $x_0 > 0$ and the origin (that is, with a non-vertical half-line through the origin). Each left coset

$$\begin{bmatrix} x_0 & y_0 \\ 0 & 1 \end{bmatrix} H = \{ \begin{bmatrix} x_0 z & y_0 \\ 0 & 1 \end{bmatrix} : z > 0 \}$$

can be identified with a horizontal half-line.

123. This is exactly the argument that each solution of the inhomogeneous set of equations is the sum of a solution of the corresponding homogeneous set and a fixed solution of the inhomogeneous set (a coset representative).

124. The cosets of H are H and Hb. If $ab \in Hb$, show that $a \in H$.

125. Consider cosets H, Hx, Hy with $H \neq Hx$ and $H \neq Hy$. So $x, y \notin H$ and so $x, y^{-1} \notin H$. Thus $xy^{-1} \in H$ and so $Hx = Hy$. Thus there can be at most one coset different from H.

126. The first question (in the third sentence) can be seen by direct calculation. For the second question, we first consider uniqueness. If two expressions of the kind given were apparently different but equal in the group then one can easily deduce either that some power a^k of a with $0 < k < n$ is the identity or that b is a power of a. Both these are impossible. To show that every element x of D_n has an expression of the given type, note that either x is a rotation in which case $x = a^k$ for $0 \leq k < n$ or x is a reflection, in which case xb is a rotation.

127. Suppose that a is a rotation through $2\pi/5$ and b is a reflection. Then the subgroups are

$$\{e_{D_5}\}, \langle a \rangle, \langle b \rangle, \langle ab \rangle, \langle a^2 b \rangle, \langle a^3 b \rangle, \langle a^4 b \rangle, D_5.$$

128. (a) $e, a, a^2, a^3, b, ab, a^2b, a^3b$; the cyclic subgroups are

$$\langle e \rangle, \langle a \rangle = \langle a^3 \rangle, \langle a^2 \rangle, \langle b \rangle, \langle ab \rangle, \langle a^2 b \rangle, \langle a^3 b \rangle.$$

(b) Colour two opposite sides of the square red and the other two opposite sides blue. The set of symmetries which preserve this colouring will be a subgroup isomorphic to D_2 and so will be a non-cyclic subgroup of order 4. If b represents a reflection about a line joining midpoints of opposite edges then this subgoup will be $\{e, a^2, b, a^2 b\}$. Now colour two opposite vertices of the square red and the other two opposite vertices blue. This gives another non-cylic subgroup of order 4.

(c) The subgroup of all rotations is cyclic and so any non-cylic subgroup must contain at least one reflection. Since groups of order 2 are cyclic, it must also contain at least one more non-identity element. If this is another reflection then the product of these two different reflections is a non-identity rotation. Thus the subgroup must contain a non-identity rotation. A little checking should now convince you that the subgroup is either one of the two subgroups above or the whole group.

129. For the subgroups of orders 1,2,3, take cyclic subgroups generated by rotations. For the subgroup of order 4, take the set of all rotations about axes connecting the midpoints of opposite sides (you need to show it is a subgroup). For the last part, first establish that there is no element of order 6 and so no cyclic subgroup of order 6.

130. Let $p = 29$ and note that p is prime. An element of G must have order dividing p^2 and so must have order p or p^2 if it is not the identity. If there is an element of order p^2, show that G is cyclic.

131. (a) $m = 57$, $n = 2 \times 18 = 36$

(b) $\gcd(5, 36) = 1$

(c) $2^5 \equiv 32 \pmod{57}$, $3^5 \equiv 15 \pmod{57}$, $6^5 \equiv 24 \pmod{57}$ and $18^5 \equiv 18 \pmod{57}$. So the encrypted message is '32 15 24 18'.

(d) Since $36 - 7 \times 5 = 1$, we have $-7 \times 5 \equiv 29 \times 5 \equiv 1 \pmod{36}$. So we can take $d = 29$.

(e) $7^{29} \equiv 49 \pmod{57}$ and $50^{29} \equiv 8 \pmod{57}$. So the decrypted message is '49 8'.

132. (a) 40

(b) 17 14 48 25 17 15 2 15

(c) rosebud

133.

134. Use the definitions.

135. This uses the fact that H must be closed under multiplication.

136. Use the definition of normal subgroup.

137. Show that, for any $g \in G$, gHg^{-1} has the same order as H by showing that the map $H \to gHg^{-1}$ given by $h \mapsto ghg^{-1}$ is a bijection. So, by assumption $gHg^{-1} = H$.

138. The normal subgroups of D_4 are

$$\langle e\rangle, \langle a^2\rangle, \langle a\rangle, \langle a^2, b\rangle, \langle a^2, ab\rangle, D_4.$$

139. (a) This is essentially straight calculation. To show that the elements form a group, we need only show closure under multiplication. You can do this via a multiplication table, but there are quicker ways.

(b) They are

$$\langle U\rangle, \langle -U\rangle, \langle I\rangle = \langle -I\rangle, \langle J\rangle = \langle -J\rangle, \langle K\rangle = \langle -K\rangle.$$

(c) If we are to find a non-cyclic subgroup of Q_8 then we must include at least two elements out of $\pm I, \pm J, \pm K$ and not two of the form $\{I, -I\}$ etc. But then it is not too hard to check that we can generate every element and so the subgroup is the whole group.

(d) No. *All* proper subgroups of Q_8 are cyclic but this is not true for D_4. (Or: all subgroups of Q_8 are normal but this is not true for D_4.)

140. Note that if $f(h) = e_H$ and $g \in G$ then

$$f(ghg^{-1}) = f(g)f(h)f(g)^{-1} = f(g)e_H f(g)^{-1} = e_H.$$

141. Use the definition of kernel.

142. If $a, b \in G$, then $NaNb = Nab = Nba = NbNa$.

143. The image, under the natural homomorphism, of the generator of G, will generate G/N.

144. Find suitable normal subgroups and find the quotients.

145.

$$\mathbb{Q}/\mathbb{Z} = \{a + \mathbb{Z} : a \in \mathbb{Q}\} \leq \{a + \mathbb{Z} : a \in \mathbb{R}\} = \mathbb{R}/\mathbb{Z}.$$

For the second part, $a + \mathbb{Z}$ has finite order if and only if $n(a + \mathbb{Z}) = 0 + \mathbb{Z}$ if and only if $na \in \mathbb{Z}$. That is, if and only if $a \in \mathbb{Q}$.

146.

	H	Ha	Ha^2	Ha^3	Hb	Hab	Ha^2b	Ha^3b
H	H	Ha	Ha^2	Ha^3	Hb	Hab	Ha^2b	Ha^3b
Ha	Ha	Ha^2	Ha^3	H	Hab	Ha^2b	Ha^3b	Hb
Ha^2	Ha^2	Ha^3	H	Ha	Ha^2b	Ha^3b	Hb	Hab
Ha^3	Ha^3	H	Ha	Ha^2	Ha^3b	Hb	Hab	Ha^2b
Hb	Hb	Ha^3b	Ha^2b	Hab	H	Ha^3	Ha^2	Ha
Hab	Hab	Hb	Ha^3b	Ha^2b	Ha	H	Ha^3	Ha^2
Ha^2b	Ha^2b	Hab	Hb	Ha^3b	Ha^2	Ha	H	Ha^3
Ha^3b	Ha^3b	Ha^2b	Hab	Hb	Ha^3	Ha^2	Ha	H

147. (a) Orbits: $\{1,2,3\},\{4\}$; stabilisers, $\mathrm{Stab}(1)=\mathrm{Stab}(2)=\mathrm{Stab}(3)=\{(1)\},\mathrm{Stab}(4)=G$.

(b) Orbits: $\{1,2,3,4\}$; stabilisers, $\mathrm{Stab}(1)=\mathrm{Stab}(2)=\mathrm{Stab}(3)=\mathrm{Stab}(4)=\{(1)\}$.

(c) Orbits: $\{1,2\},\{3,4\}$; stabilisers, $\mathrm{Stab}(1)=\mathrm{Stab}(2)=\langle(34)\rangle,\mathrm{Stab}(3)=\mathrm{Stab}(4)=\langle(12)\rangle$.

(d) Orbits: $\{1,2,3,4\}$; stabilisers, $\mathrm{Stab}(i)$ is the set of all permutations not involving i (isomorphic to S_3).

(e) Orbits: $\{1,2,3,4\}$; stabilisers, $\mathrm{Stab}(1)=\mathrm{Stab}(3)=\langle(24)\rangle,\mathrm{Stab}(2)=\mathrm{Stab}(4)=\langle(13)\rangle$. (The easiest way to see this is to think of $\{1,2,3,4\}$ as the vertices of a square.)

148. It may help to spell this out more explicitly as a homomorphism $\phi:\mathbb{R}\to\mathrm{Sym}(X)$ given by $\phi(\alpha)(x)=X+\alpha v$. You need to show first that $\phi(\alpha)$ is a bijection and then that ϕ is a homomorphism. The orbits are lines parallel to v.

149. The orbits are $\{1,2,7,12\},\{3,6,10\},\{4,8,14\},\{5,9,11,13,15\}$. The orbit-stabiliser relation implies that the order of the group is divisible by the size of the orbits. Thus $|G|$ is a multiple of 3 and 4 and 5 and so of 60.

150. Since G has order 5, each orbit has size 1 or 5. The size of X is the sum of the sizes of these orbits. So at least one orbit has size one; that is, some point of X is fixed by every element of G.

For the second part, consider $G=\langle(123)(45678)\rangle$ acting on
$X=\{1,2,3,4,5,6,7,8\}$. There is no element of X fixed by every element of G (the orbits have size 5 and 3).

151.

$$\{U\},\{-U\},\{I,-I\},\{J,-J\},\{K,-K\}.$$

152. (a) $(123),(132)$;

(b) $(123),(132),(124),(142),(134),(143),(234),(243)$;

(c) $(1234),(1243),(1324),(1342),(1423),(1432)$;

(d) all 4-cycles;

(e) all m-cycles.

153. Suppose that $\sigma(i)=j$. Then

$$\tau\sigma\tau^{-1}\left(\tau(i)\right)=\tau\sigma(i)=\tau(j).$$

Thus if j follows i in the cycle decomposition of σ then $\tau(j)$ follows $\tau(i)$ in the cycle decomposition of $\tau\sigma\tau^{-1}$. With suitable adaptations for the elements preceding a right parenthesis, this gives the general answer.

154. If $g=khk^{-1}$ show that $C_G(g)=kC_G(h)k^{-1}$.

155. This is done by direct computation. We do the first one as an example. Suppose that a matrix

$$\begin{bmatrix} a & b & c\\ d & e & f\\ g & h & i\end{bmatrix}$$

commutes with the matrix of part (a). Then we have

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}.$$

Thus

$$\begin{bmatrix} a & 2b & 3c \\ d & 2e & 3f \\ g & 2h & 3i \end{bmatrix} = \begin{bmatrix} a & b & c \\ 2d & 2e & 2f \\ 3g & 3h & 3i \end{bmatrix}$$

and so,comparing coefficients, we obtain $b = c = f = d = g = h = 0$; that is, the centraliser of the given matrix consists only of (invertible) diagonal matrices.

(a) The subgroup of diagonal matrices.

(b) All matrices of the form $\begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & e \end{bmatrix}$;

(c) All matrices of the form $\begin{bmatrix} a & b & 0 \\ 0 & a & 0 \\ 0 & 0 & e \end{bmatrix}$;

(d) All matrices of the form $\begin{bmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{bmatrix}$;

(e) All matrices of the form $\begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{bmatrix}$.

156. This is straightforward checking.

157. Choose h so that G/Z is generated by Zh. Then each element of G can be written in the form zh^i for some $z \in Z$ and some i.

158. Note that $\{e_G\}$ is always a conjugacy class. So if there is only one class then G is the identity group. If there are two classes $\{e_G\}$ and C, say, then $|G| = 1 + |C|$ and $|C|$ divides $|G|$. So $|C| = 1$ and $|G| = 2$. Thus G is the cyclic group of order 2.

If there are three classes, $\{e_G\}$, C and D say with $|C| \leq |D|$, then $|G| = 1 + |C| + |D|$ and both $|C|$ and $|D|$ divide $|G|$. Check that the only solutions to this equation are $|C| = |D| = 1$ or $|C| = 1, |D| = 2$ or $|C| = 2, |D| = 3$ (or vice-versa). The first possibility corresponds to the cyclic group of order 3. You can use the previous question to show that the second possibility does not occur. The third possibility corresponds to S_3. Ask if you have trouble proving this (or use the fact that any group of order 6 is either cyclic or S_3).

159. By Lemma 3.7.5 (and Lagrange's Theorem) the centre of G has order p or p^2. If the order is p then the quotient also has order p and so must be cyclic. Now use Exercise 157.

160. By the previous exercise, such a group G is abelian of order p^2. If G is not cyclic, choose two elements a, b of order p, with b not in the cyclic subgroup $\langle a \rangle$ generated by a. Then show that $([i]_p, [j]_p) \mapsto a^i b^j$ is an isomorphism from $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$ to G.

161. Use Cauchy's Theorem to show that the group has an element of order p and so a subgroup of order p. Since this subgroup has index $2p/p = 2$, it is normal.

162. Use the previous exercise to show that there is a normal subgroup of order p, generated by x say. Let y be an element of order 2. Then yxy^{-1} is again of order p and so must be a power of x. Show that $yxy^{-1} = x$ or $yxy^{-1} = x^{-1}$. Then show that the former case corresponds to the cyclic group of order 2p and the latter to D_p.

163. Various geometrical facts are needed: the product of two translations is a translation, the product of a translation and a rotation is a rotation and the product of two rotations is either a translation or a rotation. You also need to check inverses.

164. Use Exercise 2.4.4 and show that the product of two (glide) reflections is either a translation or a rotation.

165. Use Theorem 2.8.3.

166. A rotation about the origin is of the form $z \mapsto e^{i\theta}z$, a translation is of the form $z \mapsto z + u$ and the reflection in the real axis is of the form $z \mapsto \bar{z}$. Show that any isometry can be written as a combination of these.

167. The orbit of $(1, 0)$ is $\{(n, 0) : n \in \mathbb{Z}\}$. The stabiliser is the group generated by the reflection which passes through $(1, 0)$. The orbit of $(\frac{1}{2}, 0)$ is $\{(\frac{n}{2}, 0) : n \in \mathbb{Z}, n \text{ odd}\}$. The stabiliser is the group generated by the reflection which passes through $(\frac{1}{2}, 0)$. The orbit of $(\frac{1}{3}, 0)$ is $\{(\frac{n}{3}, 0) : n \in \mathbb{Z}, n \text{ not a multiple of } 3\}$. The stabiliser is the identity subgroup.

168. The set of all translations through a fixed distance form a conjugacy class. The set of all glide reflections through a fixed distance form a conjugacy class. The set of all rotations through a fixed angle form a conjugacy class. The set of all reflections form a conjugacy class.

169. Use the hint in the sketch of the proof.

170. This can be done by elementary (but not necessarily easy) co-ordinate geometry. Alternatively, use Question 166. For example, we can do part (1a) as follows. Choose axes so that L_1 is the x-axis (real axis) and so that L_1 and L_2 meet at the origin. Suppose that L_2 includes all points of the form $\{re^{i\phi} : r \geq 0\}$. then we can represent σ_1 and σ_2 as

$$\sigma_1 : z \mapsto \overline{z} \qquad \sigma_2 : z \mapsto \overline{z}e^{-2\phi i}.$$

Thus

$$\sigma_1\sigma_2 : z \mapsto ze^{2\phi i}$$

and it is easily checked that $\sigma_1\sigma_2$ represents a rotation through 2ϕ.

5 Old Examinations

5.1 The 1997 examination

The University of Melbourne
Semester 2 Assessment, 1997
Department of Mathematics
618-202 Linear and Abstract Algebra

Instructions to Students:
The examination paper is in two sections. The questions in Section A are shorter and more routine than those in Question B. It is recommended that candidates attempt the questions in Section A before trying those in Section B. It is possible to pass the examination on marks from Section A alone. Full marks may be obtained by answering correctly all questions in Section A and four of the questions in Section B. All questions may, however, be attempted.

Identical Examination Papers: nil

Common content examinations: nil

Reading time: 15 minutes

Duration of examination: Three hours

Length of this question paper: 5 pages

Authorized materials:
Numerical calculators, pens, rubbers, and rulers are authorized. No other materials are authorized. Candidates are reminded that no written or printed material related to this subject may be brought into the examination. If you have any such material in your possession, you should immediately surrender it to an invigilator.

Instructions to Invigilators:
Script books only are required. Candidates are permitted to take this question paper with them at the end of the examination. No written or printed material related to the subject may be brought into the examination.

Reproduction of question paper: After the examination, this question paper may be reproduced and lodged in the Baillieu Library.

Section A

(1) In the vector space $\mathcal{P}_2(\mathbb{R})$ of polynomials with real coefficients and degree at most 2, decide whether the following set of three vectors is linearly independent, giving reasons for your answer:

$$\{1+2x, 1-x, 1+x+x^2\}.$$

(6 marks)

(2) If a linear transformation f on a finite dimensional vector space V satisfies $f^2 = f$, describe the possibilities for the Jordan Normal Form of f.

(6 marks)

(3) Give an example of a 4×4 matrix over the real numbers which is not diagonalisable (that is, is not similar to a diagonal matrix). Give reasons for your answer.

(6 marks)

(4) Give an example of a 3×3 Hermitian matrix over the complex numbers in which not all entries are real.

(6 marks)

(5) Let f and g be linear transformations on a finite dimensional inner product space. If f^* and g^* denote the adjoints of f and g, respectively, show that

$$(fg)^* = g^*f^*$$

where $(fg)^*$ denotes the adjoint of fg.

(6 marks)

(6) Give the product of the following permutations in S_8 and the order of the result:

$$(12)(3456) \text{ and } (16483725).$$

(6 marks)

(7) Give an example, with explanation, of an infinite non-commutative group.

(6 marks)

(8) Let $\mathbb{Z}_7 \setminus \{0\}$ denote the group of non-zero elements of the integers modulo 7 together with the operation of multiplication. (You need not verify that this is a group.) Show that this group is cyclic.

(6 marks)

(9) Give an example of a a group G and a subgroup H of G which is **not** normal in G.

(6 marks)

(10) In the group D_4 of all symmetries of a square, let g be an element which represents a reflection in a diagonal of the square. Describe the conjugacy class containing g.

(6 marks)

Section B

(11) Given that the eigenvalues of the following matrix are 1,1 and 2, calculate its Jordan Normal Form:

$$\begin{bmatrix} -1 & 1 & -2 \\ 1 & 1 & 1 \\ 3 & -1 & 4 \end{bmatrix}.$$

(10 marks)

(12) Show that, if a linear transformation f on a complex inner product space V has a diagonal matrix with respect to some orthonormal basis of V then f is normal.

(10 marks)

(13) State the Spectral Theorem for a linear transformation on a complex vector space. Use it to deduce that any normal matrix A which satisfies $A^n = 0$ for some n must satisfy $A = 0$.

(10 marks)

(14) A regular pentagon has as its symmetry group the group D_5 with 10 elements (you need not prove this). Suppose that the edges of the pentagon are coloured in some way in which each edge may be coloured with one or several colours. Show that the set of elements of D_5 which map the pentagon to itself, with the same colouring, is a subgroup of D_5. Give the possible orders of such subgroups. For each such order, give a coloured pentagon which has a symmetry group of that order.

(10 marks)

(15) Let $\mathbb{R}$ denote the group of real numbers under the operation of addition. Let U denote the group of complex numbers with absolute value 1 under the operation of multiplication. Prove that the function $\mathbb{R} \to U$ given by $a \mapsto e^{2\pi i a}$ for all $a \in \mathbb{R}$ is a homomorphism. Deduce that the quotient group $\mathbb{R}/\mathbb{Z}$ is isomorphic to U.

(10 marks)

(16) Let $\mathcal{I}$ denote the group of all isometries of the plane. Let P be some fixed point in the plane. Show that any element of $\mathcal{I}$ can be uniquely expressed as a product of a translation and an isometry which fixes P.

(10 marks)

5.2 The 1998 examination

The University of Melbourne

Semester 2 Assessment, 1998

Department of Mathematics and Statistics

620-222 Linear and Abstract Algebra

Instructions to Students:
The examination paper is in two sections. The questions in Section A are shorter and more routine than those in Question B. It is recommended that candidates attempt the questions in Section A before trying those in Section B. It is possible to pass the examination on marks from Section A alone. Full marks may be obtained by answering correctly all questions in Section A and four of the questions in Section B. All questions may, however, be attempted.

Identical Examination Papers: nil
Common content examinations: nil
Reading time: 15 minutes
Duration of examination: Three hours
Length of this question paper: 4 pages

Authorized materials:
Numerical calculators, pens, rubbers, and rulers are authorized. No other materials are authorized. Candidates are reminded that no written or printed material related to this subject may be brought into the examination. If you have any such material in your possession, you should immediately surrender it to an invigilator.

Instructions to Invigilators:
Script books only are required. Candidates are permitted to take this question paper with them at the end of the examination. No written or printed material related to the subject may be brought into the examination.

Reproduction of question paper: After the examination, this question paper may be reproduced and lodged in the Baillieu Library.

Section A

(1) Decide whether the following matrices span the space $M_{2\times 2}(\mathbb{R})$ of all 2×2 matrices with real entries:

$$\begin{bmatrix}1 & 2\\ 0 & 0\end{bmatrix} \quad \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} \quad \begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix} \quad \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}.$$

Give brief reasons for your answer.

(6 marks)

(2) Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation represented by the matrix

$$\begin{bmatrix}1 & 2 & 0\\ 2 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Describe a 2-dimensional f-invariant subspace of $\mathbb{R}^3$.

(6 marks)

(3) Find the Jordan canonical (or normal) form of the matrix

$$\begin{bmatrix}4 & 1\\ -1 & 2\end{bmatrix}.$$

(6 marks)

(4) a) Give an example of a diagonal 3×3 matrix over the complex numbers which is unitary (that is, isometric) but not Hermitian (that is, self-adjoint).

b) Give an example of a diagonal 3×3 matrix over the complex numbers which is Hermitian but not unitary .

(6 marks)

(5) Let f be an isometric linear transformation on a finite dimensional complex inner product space V; thus f satisfies $f^*f = \mathrm{id}_V$. Explain briefly why $||f(v)|| = ||v||$ for all $v \in V$.

(6 marks)

(6) Find the order of the permutation (15749)(23)(68). What is the 12th power of this permutation?

(7) Give an example, with explanation, of a non-abelian group which contains an element of order 4.

(6 marks)

(8) Let G denote the group of all 3×3 diagonal matrices where each diagonal entry is 1 or -1, with the operation of matrix multiplication. (You need not show that G is a group). Decide, giving reasons, whether G is a cyclic group.

(6 marks)

(9) A group is known to contain an element of order 2, an element of order 3, and an element of order 5. What is the least possible order for this group?

(6 marks)

(10) How many conjugates does the permutation (123) have in the group S_3 of all permutations on 3 letters? Give brief reasons for your answer.

(6 marks)

Section B

(11) For any natural number n, $\mathbb{C}^n$ denotes (as usual) the set of n-tuples of complex numbers, furnished with the inner product given by

$$\big((a_1, \ldots, a_n), (b_1, \ldots, b_n)\big) = a_1\overline{b_1} + \cdots + a_n\overline{b_n}.$$

Let $f : \mathbb{C}^3 \to \mathbb{C}^1$ denote the linear transformation given by

$$f(a_1, a_2, a_3) = (a_1 + a_2 + ia_3).$$

Find a basis for the nullspace (or kernel) of f. Find also a basis for the orthogonal complement of the nullspace of f.

(10 marks)

(12) Let A be a 6×6 matrix with complex entries. Suppose that the characteristic polynomial of A is known to be as follows:

$$x(x-1)^2(x-2)^3.$$

Given this information, what are the possibilities for the Jordan canonical (or normal) form of A? What further computations could be used to establish which was the correct choice for the canonical form? Explain clearly how the outcome of your computations would enable you to determine the Jordan canonical form.

(10 marks)

(13) Let A be a square matrix, with complex entries, of finite order. That is, $A^n = I$ for some natural number n, where I represents the identity matrix of appropriate size. Show that the minimal polynomial of A has no repeated roots. Deduce that A is diagonalisable; that is, A is similar to a diagonal matrix. You may use any results from the notes or lectures but should then quote them carefully.

(10 marks)

(14) Let G denote the group of all matrices of the form

$$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \qquad a, b, c \in \mathbb{R}, \quad a \neq 0, c \neq 0$$

with the operation of multiplication and let H denote the group of all matrices of the form

$$\begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix} \qquad a, c \in \mathbb{R}, \quad a \neq 0, c \neq 0$$

also with the operation of multiplication.

Show that the function $f : G \to H$ given by

$$f\left(\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}\right) = \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix}$$

is a homomorphism. Find the kernel and image of this homomorphism. Hence show that H is isomorphic to a quotient group of G.

(10 marks)

(15) Let X denote the set of n axes of symmetry of a regular n-gon. Let D_n denote the group of all symmetries of a regular n-gon; thus D_n contains $2n$ elements. Explain briefly what it means to say that D_n acts on X. If n is even, use this action to show that D_n has a normal subgroup of order 2.

(10 marks)

(16) Consider the one dimensional pattern below

...EEEEEEEEEEEEE...

which is assumed to be repeated indefinitely in both horizontal directions.

Find the symmetry group of this pattern. Identify the translation subgroup and the point group. Is the symmetry group abelian?

(10 marks)

5.3 The 1999 examination

The University of Melbourne

Semester 2 Assessment, 1999

Department of Mathematics and Statistics

620-222 Linear and Abstract Algebra

Instructions to Students:
The examination paper is in two sections. The questions in Section A are shorter and more routine than those in Question B. It is recommended that candidates attempt the questions in Section A before trying those in Section B. It is possible to pass the examination on marks from Section A alone. All questions may be attempted.

Identical Examination Papers: nil
Common content examinations: nil
Reading time: 15 minutes
Duration of examination: Three hours
Length of this question paper: 4 pages

Authorized materials:
Numerical calculators, pens, rubbers, and rulers are authorized. No other materials are authorized. Candidates are reminded that no written or printed material related to this subject may be brought into the examination. If you have any such material in your possession, you should immediately surrender it to an invigilator.

Instructions to Invigilators:
Script books only are required. Candidates are permitted to take this question paper with them at the end of the examination. No written or printed material related to the subject may be brought into the examination.

Reproduction of question paper: After the examination, this question paper may be reproduced and lodged in the Baillieu Library.

Section A

(1) Decide whether the following matrices form a basis of the space $M_{2\times 2}(\mathbb{R})$ of all 2×2 matrices with real entries:

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Give brief reasons for your answer.

(6 marks)

(2) Decide, giving reasons, whether the matrix

$$\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$$

is diagonalisable (that is, is it similar to a diagonal matrix?)

(6 marks)

(3) A matrix is known to have minimal polynomial $(X-1)^2(X-2)(X-3)$ and characteristic polynomial $(X-1)^2(X-2)^2(X-3)$. Write down its Jordan Normal Form.

(6 marks)

(4) Give an example of an inner product on the space $\mathcal{P}_4(\mathbb{R})$ of all polynomials with real coefficients and degree at most 4. Explain briefly the checks needed to verify that this is an inner product.

(6 marks)

(5) If f is a linear transformation on an inner product space and if f^* denotes the adjoint of f, show briefly that $(f^*)^* = f$.

(6 marks)

(6) Find the order of the product of the two permutations

$$(167253)(48) \text{ with } (1645)(28).$$

(6 marks)

(7) Give an example, with explanation, of an abelian (that is, commutative) group which is not cyclic.

(6 marks)

(8) You are told that an icosahedron has a total of 120 symmetries. Explain how to deduce, *from this information alone*, that there is no rotational symmetry of order 9.

(6 marks)

(9) Suppose that G is an abelian group. If N is a normal subgroup of G, explain briefly why the quotient group G/N is also abelian.

(6 marks)

(10) Let $GL(2,\mathbb{R})$ denote the group of all invertible 2×2 matrices with real entries. Give an example of a matrix in $GL(2,\mathbb{R})$ which is conjugate to, but not equal to, the matrix

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

(6 marks)

Section B

(11) Let V be an inner product space with the inner product of vectors $u, v \in V$ denoted by (u, v). Suppose that $\{v_1, \ldots, v_n\}$ is an orthonormal basis of V. Show carefully that, for any $u \in V$,

$$u = (u, v_1)v_1 + (u, v_2)v_2 + \cdots + (u, v_n)v_n.$$

(10 marks)

(12) Let f be a non-zero linear transformation $f : \mathbb{R}^3 \to \mathbb{R}^3$. If f^2 is the zero transformation, show that there is only one possible Jordan Normal Form for f.

If $g : \mathbb{R}^4 \to \mathbb{R}^4$ is non-zero and g^2 is zero, show that there are two possible Jordan Normal Forms for g.

(For the purposes of this question, we do not regard two Jordan Normal Forms as being different if one can be obtained from the other by re-ordering the Jordan blocks.)

(10 marks)

(13) Let f be a normal linear transformation on an inner product space V. Let f^* denote the adjoint of f. If $v \in V$, show that

$$(f(v), f(v)) = (f^*(v), f^*(v))$$

and use this to deduce that the kernel of f is equal to the kernel of f^*.

(10 marks)

(14) Let $n = 2^k$ for some natural number k. Let U_n denote the elements $[a]_n \in \mathbb{Z}_n$ with *a odd.* When U_n is furnished with the operation of multiplication modulo n, then the result is a group.(You need not prove this).

What is the order of U_n? Use this to deduce that, if a is an odd integer, then $a^{2^{k-1}} - 1$ is a multiple of 2^k.

(10 marks)

(15) Let $\mathbb{C}^*$ denote the group of non-zero complex numbers with the operation of multiplication and let $\mathbb{R}^*$ denote the group of non-zero real numbers with the operation of multiplication. Consider the function $abs : \mathbb{C}^* \to \mathbb{R}^*$ given by $abs : z \mapsto |z|$. Identify the kernel and image of abs and use this to show that a quotient group of $\mathbb{C}^*$ is isomorphic to a subgroup of $\mathbb{R}^*$.

(10 marks)

(16) Suppose that a group of order 35 acts on a set X with 18 elements. Show that some element of X must be left fixed by every permutation corresponding to an element of G. Give an example of a group of order 35 which acts on a set with 12 elements in such a way that no element is left fixed by every permutation corresponding to an element of G.

(10 marks)

5.4 The 2000 examination

The University of Melbourne

Semester 2 Assessment, 2000
Department of Mathematics and Statistics

620-222 Linear and Abstract Algebra

Instructions to Students:
The examination paper is in two sections. The questions in Section A are shorter and more routine than those in Question B. It is recommended that candidates attempt the questions in Section A before trying those in Section B. It is possible to pass the examination on marks from Section A alone. All questions may be attempted.

Identical Examination Papers: nil
Common content examinations: nil
Reading time: 15 minutes
Duration of examination: Three hours
Length of this question paper: 4 pages

Authorized materials:
Numerical calculators, pens, rubbers, and rulers are authorized. No other materials are authorized. Candidates are reminded that no written or printed material related to this subject may be brought into the examination. If you have any such material in your possession, you should immediately surrender it to an invigilator.

Instructions to Invigilators:
Script books only are required. Candidates are permitted to take this question paper with them at the end of the examination. No written or printed material related to the subject may be brought into the examination.

Reproduction of question paper: After the examination, this question paper may be reproduced and lodged in the Baillieu Library.

Section A

(1) Decide whether the following polynomials form a basis of the space $\mathcal{P}_2(\mathbb{R})$ of all polynomials of degree at most 2 with real coefficients:

$$1 - x + x^2, 1 + x + x^2, 1 + x - x^2.$$

Give brief reasons for your answer.

(6 marks)

(2) Find the minimal polynomial for the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Is the matrix diagonalizable? Give brief reasons.

(6 marks)

(3) A complex matrix has minimal polynomial $X^2(X-1)^2$ and characteristic polynomial $X^2(X-1)^4$. Find all the possibilities for the Jordan form of the matrix.

(6 marks)

(4) Let V be the subspace of $\mathbb{R}^3$ spanned by the vectors $(1, 1, 0), (0, 1, 2)$. Find the orthogonal complement of V, using the dot product as inner product on $\mathbb{R}^3$.

(6 marks)

(5) Give an example of a 2×2 Hermitian (i.e. self-adjoint) matrix in which not all entries are real. Is your matrix normal? Give a brief explanation.

(6 marks)

(6) Does the set $P = \{x \in \mathbb{R} \colon x > 0\}$ of positive real numbers form a group using the operation of: (i) addition, (ii) multiplication? Give brief reasons.

(6 marks)

(7) Calculate the product of the following permutations, and find the order of the result:

$$(135)(2678) \text{ and } (14)(23578).$$

(6 marks)

(8) The set $H = \{[1]_9, [2]_9, [4]_9, [5]_9, [7]_9, [8]_9\}$ of elements from $\mathbb{Z}_9$ forms a group under multiplication modulo 9. (You do not need to prove this.) Show that this group is cyclic and find a generator for the group.

(6 marks)

(9) A finite group G has fewer than 100 elements and has subgroups of orders 10 and 25. What is the order of G? Give a brief explanation.

(6 marks)

(10) Let G be the group of rotational symmetries of a cube, and consider the action of G on the set X of all 6 faces of the cube. Describe the orbit and stabiliser of a face F. Use this to find the order of G. Give brief reasons.

(6 marks)

Section B

(11) $V = \{x + y\sqrt{2} \colon x, y \in \mathbb{Q}\}$ is a vector space over the rational numbers $\mathbb{Q}$, using the usual operations of addition and multiplication for real numbers. (You do not need to prove this.) Let $f : V \to V$ be multiplication by $a + b\sqrt{2}$, where $a, b \in \mathbb{Q}$ and a, b are not both zero.

a) Show that f is a linear transformation.

b) Find the matrix of f with respect to the basis $\{1, \sqrt{2}\}$ for V. (You do not need to prove that this is a basis.)

c) Find the nullspace (or kernel) of f. Hence find the rank of f.

d) Is f surjective (i.e. onto)? Explain your answer.

(10 marks)

(12) State the Spectral Theorem for a linear transformation on a complex vector space. Use it to show that every normal matrix A with complex entries has a cube root, i.e. there is a matrix B with complex entries such that $B^3 = A$.

(10 marks)

(13) Let f be an isometry on a complex inner product space V, (i.e. f is a linear transformation satisfying $f^*f = \text{identity}$).

a) Show that each eigenvalue λ of f satisfies $|\lambda| = 1$.

b) Show that if v_1, v_2 are eigenvectors of f corresponding to distinct eigenvalues $\lambda_1 \neq \lambda_2$, then v_1, v_2 are orthogonal.

(10 marks)

(14) Let G be a group in which every element satisfies $g^2 = e$.

a) Prove that G is abelian (i.e. commutative). [Hint: consider $(ab)^2$.]

b) What can you say about the order of G? Explain your answer.

(10 marks)

(15) Consider the function $f : \mathbb{C}^* \to \mathbb{C}^*$ defined by $f(z) = z^2$, where $\mathbb{C}^*$ denotes the group of non-zero complex numbers under multiplication.

a) Show that f is a homomorphism.

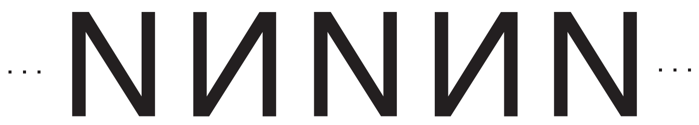
b) Find the kernel and image of f.

c) Describe the quotient of $\mathbb{C}^*$ by the kernel of f.

d) Is f an isomorphism? Give brief reasons.

(10 marks)

(16) Let G be the group of all symmetries of the frieze pattern shown below. (The pattern repeats to fill out an infinite strip in the plane.)



a) Copy the pattern into your exam book and mark:

i. the centres of rotations in G (as small circles),

ii. mirror lines of reflections in G (as solid lines),

iii. axes of glide reflections in G (as dotted lines).

b) Describe the translation subgroup and the point group for G.

c) Is the group G abelian? Give a brief explanation.

(10 marks)

5.5 Solutions to the 1997 Examination

Section A

(1) The given set of vectors is linearly independent. Suppose that

$$a(1+2x)+b(1-x)+c(1+x+x^2)=0.$$

Then $(a+b+c)+(2a-b+c)x+cx^2=0$ and so $a+b+c=0, 2a-b+c=0, c=0$. Hence we can easily deduce that $a=b=c=0$ and so the three vectors are linearly independent.

(2) Because $f^2=f$, the minimal polynomial is a factor of X^2-X and so is either X^2-X, X or $X-1$. In each case, the minimal polynomial has no repeated roots and so the Jordan Normal Form (or Jordan Canonical Form) is diagonal. Since the eigenvalues must be roots of the minimal polynomial, they must be 0 or 1. Hence the Jordan Normal Form is a diagonal matrix with 1s and 0s on the diagonal.

(3)

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

This is a matrix in Jordan Normal Form. If it were diagonalisable then the diagonal form would be another Jordan Normal Form for it. But Jordan Normal Form of a matrix is unique up to rearranging the Jordan blocks. This if one Jordan Normal Form is diagonal, then all are. Thus the matrix is not diagonalisable.

(4)

$$\begin{bmatrix} 1 & i & 0 \\ -i & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(5) Let V denote the inner product space. For any $v, w \in V$, we have

$$(fg(v), w) = (v, (fg)^*(w))$$

using the definition of $(fg)^*$. But, using the definition of f^* and g^* separately,

$$(fg(v), w) = (g(v), f^*(w)) = (v, g^*(f^*(w)))$$

and so $(v, (fg)^*(w)) = (v, g^*(f^*(w)))$ for all $v, w \in V$. It follows easily that $(fg)^* = g^*f^*$.

(6) (137)(265)(48); the order is 6.

(7) The group $GL(2,\mathbb{R})$ is a group which is clearly infinite (it contains, for example one scalar matrix for each non-zero real number). It is also non-commutative because, for example,

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \neq \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

(8) The elements of $\mathbb{Z}_7 \setminus \{0\}$ are $[1]_7, [2]_7, [3]_7, [4]_7, [5]_7, [6]_7$ and, after a little trial and error, we have

$$[3]_7^1 = [3]_7, [3]_7^2 = [2]_7, [3]_7^3 = [6]_7, [3]_7^4 = [4]_7, [3]_7^5 = [5]_7, [3]_7^6 = [1]_7.$$

Thus every element of $\mathbb{Z}_7 \setminus \{0\}$ is a power of $[3]_7$ and so $\mathbb{Z}_7 \setminus \{0\}$ is cyclic.

(9) Let $G = S_3$ and let $H = \{(1), (12)\}$. Since $(123)(12)(123)^{-1} = (23)$, H is not normal in G.

(10) The conjugacy class consists of exactly the (two) reflections in the diagonals of the square. (In more detail, the centraliser of g contains itself, the identity, and after a short computation, the rotation through π. Thus the centraliser has at least 4 elements and so the conjugacy class has at most 2 elements. A short computation shows that if h is the rotation through $\pi/2$, then the conjugate hgh^{-1} is the reflection in the other diagonal. Thus these two reflections make up the conjugacy class.)

Section B

(11) Since the eigenvalues of the matrix are 1,1 and 2, the characteristic polynomial is $(X-1)^2(X-2)$. Since the minimal polynomial divides the characteristic polynomial and has the same roots then either the minimal polynomial is $(X-1)(X-2)$ and the Jordan Normal Form is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

or else the minimal polynomial is $(X-1)^2(X-2)$ and the Jordan Normal Form is

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

To ascertain which case holds, we calculate $(A-I_3)(A-2I_3)$ (where A is the given matrix). This is

$$\begin{bmatrix} -2 & 1 & -2 \\ 1 & 0 & 1 \\ 3 & -1 & 3 \end{bmatrix} \begin{bmatrix} -3 & 1 & -2 \\ 1 & -1 & 1 \\ 3 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix}$$

ans so is, in particular, non-zero. Thus the minimal polynomial is $(X-1)^2(X-2)$ and the Jordan Normal Form is

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

(12) Let D be the diagonal matrix with respect to the orthonormal basis ($\mathcal{B}$ say). Because $\mathcal{B}$ is othonormal, the matrix of f^* with respect to $\mathcal{B}$ is then D^*, the complex conjugate transpose of D. Thus D^* is also diagonal and so $DD^* = D^*D$. As multiplication of matrices corresponds to composition of linear functions, we therefore have that $ff^* = f^*f$; that is, f is normal.

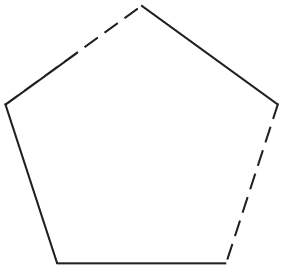
(13) The statement of the Spectral Theorem is bookwork. Suppose that A is $m \times m$ and let V denote the space $\mathbb{C}^m$. Choose an orthonormal basis of V and let f denote the linear transformation of V which corresponds to A. Then f is also normal. Since $A^n = 0$ we also have $f^n = 0$.

Thus, by the Spectral Theorem, there is an orthonormal basis $\mathcal{B}_1$ of V so that the matrix of f with respect to $\mathcal{B}_1$ is diagonal; call it D. Since $f^n = 0$ we have $D^n = 0$. Thus, for each diagonal entry d of D, we have $d^n = 0$. Hence $d = 0$ and so all of the diagonal entries of D are zero. Thus D is zero and so f is zero. Hence A is zero.

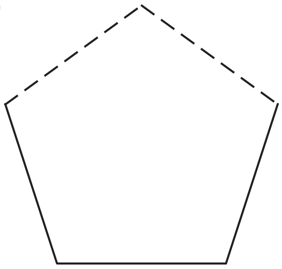
(14) To show that a subset H is a subgroup, we must show that, if h_1 and h_2 are elements of H then the product h_1h_2 is also an element of H and that if h is an element of H then h^{-1} is an element of H. In this case, that means showing that if we take two symmetries of the pentagon, each of which preserves the colouring then the effect of applying one after the other also preserves the colouring. This is clear (more detailed argument is possible but probably not useful). If a symmetry preserves the colouring, then its inverse must also preserve the colouring.

The possible orders of subgroups of D_5 are divisors of $|D_5| = 10$; that is, 1,2,5,10. We give explicit examples below for the cases where the subgroup has order 1,2,5,10 (corresponding to a),b),c),d) respectively).

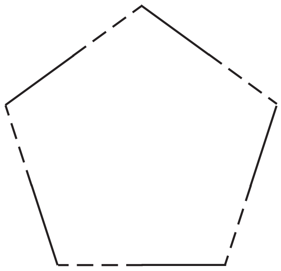
(a)



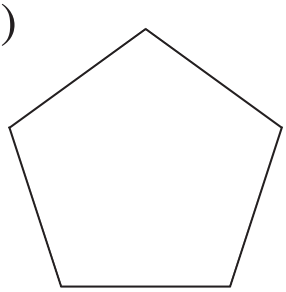
(b)



(c)



(d)



(15) Denote the function by μ. To prove that μ is a homomorphism, we must have that $\mu(a+b) = \mu(a)\mu(b)$ for all $a, b \in \mathbb{R}$. but,

$$\mu(a+b) = e^{2\pi i(a+b)} = e^{2\pi ia}e^{2\pi ib} = \mu(a)\mu(b)$$

and so μ is a homomorphism.

By the isomorphism theorem, $\mathbb{R}/\ker\mu \cong \operatorname{im}\mu$. The kernel of μ is the set

$$\{a \in \mathbb{R} : \mu(a) = 1\} = \{a \in \mathbb{R} : e^{2\pi ia} = 1\} = \{a \in \mathbb{R} : a \in \mathbb{Z}\} = \mathbb{Z}.$$

Each complex number is expressible in the form $re^{2\pi ia}$ where r is the absolute value. Thus each complex number of absolute value 1 is expressible as $e^{2\pi ia}$ and so is in the image of μ. Thus $\operatorname{im}\mu = U$. Hence $\mathbb{R}/\mathbb{Z} \cong U$, as required.

(16) Let σ be an element of $\mathcal{I}$. Suppose that $\sigma(P) = Q$. Let τ be the translation that takes P to Q. Then $\sigma(P) = \tau(P) = Q$ and so $\tau^{-1} \circ \sigma(P) = \tau^{-1}(Q) = P$. Set $\nu = \tau^{-1} \circ \sigma$. Then ν fixes P and $\sigma = \tau \circ \nu$. Thus every element of $\mathcal{I}$ can be written in the required form.

Suppose that σ could be written $\sigma = \tau_1 \circ \nu_1 = \tau_2 \circ \nu_2$ where τ_1, τ_2 are translations and ν_1, ν_2 are isometries fixing P. Then

$$\tau_2^{-1}\tau_1 = \tau_2^{-1} \circ \tau_1 \circ \nu_1 \circ \nu_1^{-1} = \tau_2^{-1} \circ \tau_2 \circ \nu_2 \circ \nu_1^{-1} = \nu_2 \circ \nu_1^{-1}.$$

But $\tau_2^{-1}\tau_1$ is a translation and $\nu_2 \circ \nu_1^{-1}$ fixes the point P. Since only the trivial translation fixes any point, we have that both $\tau_2^{-1}\tau_1$ and $\nu_2 \circ \nu_1^{-1}$ are the identity; that is $\tau_1 = \tau_2$ and $\nu_1 = \nu_2$. Thus the expression for σ is unique.

5.6 Solutions to the 1998 Examination

Section A

(1) **Either** because the $(2,2)$-entry of each matrix is zero, the $(2,2)$-entry of any linear combination of the matrices is zero. Thus not every matrix in $M_{2\times 2}(\mathbb{R})$ can be written as a linear combination of the given matrices. In particular, $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ cannot be written as a linear combination of the given matrices.

Or There are 4 matrices, and if they spanned $M_{2\times 2}(\mathbb{R})$ then they would have to be a basis, since $M_{2\times 2}(\mathbb{R})$ has dimension 4. Thus the given set would have to be linearly independent. But

$$\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = 2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

and so the set is not linearly independent.

(2) Let v_1, v_2, v_3 denote the basis of $\mathbb{R}^3$ with respect to which the matrix is written. Then

$$f(v_1) = v_1 + 2v_2, \qquad f(v_2) = 2v_1 + v_2, \qquad f(v_3) = v_3.$$

Set $W = \langle v_1, v_2 \rangle$. Then $f(v_1) \in W$ and $f(v_2) \in W$. Thus $f(\alpha v_1 + \beta v_2) \in W$ for all $\alpha, \beta \in \mathbb{R}$. Hence $f(W) \subseteq W$ and so W is f-invariant.

(3) The characteristic polynomial of the matrix is:

$$\begin{vmatrix} 4-X & 1 \\ -1 & 2-X \end{vmatrix} = (4-X)(2-X) - (1)(-1) = X^2 - 6X + 9 = (X-3)^2.$$

Thus the minimal polynomial is either $(X-3)^2$ or $(X-3)$. If it were the latter, then the matrix would be scalar. Thus it is the former and the JCF is

$$\begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}.$$

(4) a) For example,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

b) For example,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(5)

$$||f(v)||^2 = (f(v), f(v)) = (v, f^*f(v)) = (v, v) = ||v||^2.$$

(6) The fifth power of the permutation is $(15749)^5(23)^5(68)^5 = (23)(68)$. Thus the tenth power is the identity. The order is 10. Thus the twelth power equals the square and so is

$$(15749)^2(23)^2(68)^2 = (17954).$$

(7) An example is the group S_4 of all permutations of four things. This is non-abelian since, for example $(12)(13) = (132) \neq (13)(12) = (123)$. An element of order 4 is (1234).

(8) If $g \in G$, then we can write g in the form

$$g = \begin{bmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_2 & 0 \\ 0 & 0 & \epsilon_3 \end{bmatrix}$$

and $\epsilon_i = \pm 1$. But then

$$g^2 = \begin{bmatrix} \epsilon_1^2 & 0 & 0 \\ 0 & \epsilon_2^2 & 0 \\ 0 & 0 & \epsilon_3^2 \end{bmatrix} = I_3.$$

Since G has order 8 and every element of G has at most 2 different powers, G cannot be a cyclic group.

(9) By Lagrange's Theorem, the order of the group must be a multiple of the order of any element. Thus the order of this group must be a multiple of 2 and of 3 and of 5 and so a multiple of 30. (The cyclic group of order 30 is an example of a group of this order with elements of order 2,3 and 5.)

(10) The permutation (123) has 2 conjugates, itself and (132). We can see this in at least two ways.

Either, we can use the orbit-stabiliser relationship to see that the size of the conjugacy class of (123) is equal to the index of its stabiliser. But the stabiliser is a subgroup of S_3 which contains at least $\langle(123)\rangle$. If it were to contain more then it must be all of S_3 (by Lagrange's Theorem) and so contain, for example, (12). But $(12)(123)(12)^{-1} = (132)$ and so the centraliser is exactly $\langle(123)\rangle$. Since it has order 3, it also has index 2 and so (123) has 2 conjugates.

Or,if an element of S_3 is given as a product of disjoint cycles, then its conjugate by a permutation τ is obtained by replacing each entry i in each cycle by $\tau(i)$. Thus the conjugates are all of the other elements which are of the 'same cycle shape'. In the given case, this means all elements which are 3-cycles. There are exactly 2 of these.

Section B

(11) The nullspace of f is the set of elements v of $\mathbb{C}^3$ such that $f(v) = 0$. But if $v = (a_1, a_2, a_3)$ and $f(v) = 0$, then $a_1 + a_2 + ia_3 = 0$. Since $\text{im}(f) = \mathbb{C}^1$ we have $\text{rank}(f) = 1$. Thus $\text{nullity}(f) = \dim(\mathbb{C}^3) - 1 = 2$. Thus to find a basis for the nullspace of f, it will suffice to find two independent elements of the nullspace. By inspection, we can see that $(1, -1, 0)$ and $(1, 0, i)$ are elements of the nullspace and they are clearly independent. Thus $\{(1, -1, 0), (1, 0, i)\}$ is a basis for the nullspace of f.

To find the orthogonal complement of the nullspace, we need to find all vectors which are orthogonal to every vector in the nullspace. Since a vector which is orthogonal to every basis vector is orthogonal to the whole space, it suffices to work with the given basis. Then (a, b, c) lies in the orthogonal complement to the nullspace of f if and only if $\big((a, b, c), (1, -1, 0)\big) = 0$ and $\big((a, b, c), (1, 0, i)\big) = 0$; that is, if and only if $a - b = 0$ and $a - ci = 0$. Thus $b = a$ and $c = -ia$. Thus the orthogonal complement of the null-space is $\langle(1, 1, -i)\rangle$.

(12) The possibilities for the JCF J of A can be listed as follows:

$$J = \begin{bmatrix} J_0 & 0_{12} & 0_{13} \\ 0_{21} & J_1 & 0_{23} \\ 0_{31} & 0_{32} & J_2 \end{bmatrix}$$

where 0_{rs} is a $r \times s$ matrix of zeroes and

$$J_0 = [0], \qquad J_1 \in \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\}, J_2 \in \left\{ \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \right\}$$

Thus there are 6 possibilities in all.

To establish which possibility was correct, we would need to know the minimal polynomial of A. We know that this is a divisor of $x(x-1)^2(x-2)^3$ and a multiple of $x(x-1)(x-2)$ (since it must have the same roots as the characteristic polynomial). Thus we calculate successively,

$$\begin{aligned} B_1 &= A(A-I)(A-2I), B_2 = A(A-I)^2(A-2I), B_3 = A(A-I)(A-2I)^2 \\ B_4 &= A(A-I)^2(A-2I)^2, B_5 = A(A-I)(A-2I)^3. \end{aligned}$$

Choose the first B_i that is zero; if none is zero, set $i = 6$. We list below the options for J_1, J_2 corresponding to the different values of i:

$$i = 1, J_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad J_2 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$i = 2, J_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad J_2 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$i = 3, J_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad J_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$i = 4, J_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad J_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$i = 5, J_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad J_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

$$i = 6, J_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad J_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

(13) The matrix A satisfies the polynomial $X^n - 1$. Thus the minimal polynomial is a factor of $X^n - 1$. But $X^n - 1$ has no repeated roots. (To see this we can either list the n roots explicitly or observe that the derivative nX^{n-1} has no roots in common with $X^n - 1$). Thus the minimal polynomial also has no repeated roots. We can now quote the result, from lectures, that a matrix is diagonalisable if and only if its minimal polynomial has no repeated roots. Alternatively, the result can be derived, at somewhat greater length, directly from the Jordan canonical form theorem.

(14) To check that f is a homomorphism, we need to show that

$$f\left(\begin{bmatrix} a_1 & b_1 \\ 0 & c_1 \end{bmatrix}\begin{bmatrix} a_2 & b_2 \\ 0 & c_2 \end{bmatrix}\right) = f\left(\begin{bmatrix} a_1 & b_1 \\ 0 & c_1 \end{bmatrix}\right) f\left(\begin{bmatrix} a_2 & b_2 \\ 0 & c_2 \end{bmatrix}\right).$$

The left hand side of the above is

$$f\left(\begin{bmatrix} a_1a_2 & a_1b_2 + b_1c_2 \\ 0 & c_1c_2 \end{bmatrix}\right) = \begin{bmatrix} a_1a_2 & 0 \\ 0 & c_1c_2 \end{bmatrix}$$

whereas the right hand side is

$$\begin{bmatrix} a_1 & 0 \\ 0 & c_1 \end{bmatrix}\begin{bmatrix} a_2 & 0 \\ 0 & c_2 \end{bmatrix} = \begin{bmatrix} a_1a_2 & 0 \\ 0 & c_1c_2 \end{bmatrix}.$$

Thus f is a homomorphism.
The kernel $K$of f is the set of elements of G which satisfy

$$f\left(\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix};$$

that is, the set of matrices of the form $\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$. The image of f is H.
By the Isomorphism Theorem, $G/\ker(f) \cong \operatorname{im}(f)$; that is,

$$G/K \cong H.$$

(15) Each element of D_n is a symmetry of the n-gon which must therefore take one diagonal to another. Thus each element of D_n induces a permutation of the diagonals. That is, there is a homomorphism $\phi : D_n \to \mathrm{Sym}(X)$.
The kernel of the homomorphism is the set of all symmetries of the n-gon which leave each diagonal fixed. If n is even, this will be done by a rotation through π as well as the identity element. Thus the kernel will have order 2 and so will be a normal subgroup of order 2.

(16) Consider the one dimensional pattern below

$$\ldots \mathrm{EEEEEEEEEEEEEE}\ldots$$

which is assumed to be repeated indefinitely in both horizontal directions.

Choose co-ordinates so that the x-axis is horizontal and bisects each E. Choose the unit of distance so that the letters are distance 1 apart. The symmetry group of this pattern contains horizontal translations through any integral distance. If τ is a horizontal translation rightwards through distance 1, then any translation is a power of τ. There is also a reflection, call it σ, in the x-axis. Finally, there are glide reflections of the form $\sigma\tau^n$ which are horizontal translations through distance n followed by the reflection in the x-axis.
The translation subgroup is infinite cyclic generated by τ; the point group is cyclic of order 2 generated by reflection in a horizontal axis.
Note that $\sigma\tau = \tau\sigma$ (in coordinates, both have the effect $(x, y) \to (x + n, -y)$) and so $\sigma\tau^n = \tau^n\sigma$. Any element can be written in the form $\sigma^\epsilon\tau^n$ where $\epsilon = 0, 1$ and n is an integer. Thus, working with two general elements of the group,

$$\begin{aligned}(\sigma^{\epsilon_1}\tau^{n_1})(\sigma^{\epsilon_2}\tau^{n_2}) &= \sigma^{\epsilon_1}(\tau^{n_1}\sigma^{\epsilon_2})\tau^{n_2} = \sigma^{\epsilon_1}(\sigma^{\epsilon_2}\tau^{n_1})\tau^{n_2} = \sigma^{\epsilon_1+\epsilon_2}\tau^{n_1+n_2}\\ = \sigma^{\epsilon_2}(\sigma^{\epsilon_1}\tau^{n_2})\tau^{n_1} &= \sigma^{\epsilon_2}(\tau^{n_2}\sigma^{\epsilon_1})\tau^{n_1} = (\sigma^{\epsilon_2}\tau^{n_2})(\sigma^{\epsilon_1})\tau^{n_1})\end{aligned}$$

which shows that the group is abelian.

5.7 Solutions to the 1999 Examination

Section A

(1) We check if these matrices span $M_{2\times 2}(\mathbb{R})$.

$$a\begin{bmatrix}0 & 1\\ 1 & 1\end{bmatrix}+b\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}+c\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}+d\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}=\begin{bmatrix}x & y\\ z & w\end{bmatrix}$$

gives the system of linear equations

$$\begin{array}{ccccccc} & b & +c & +d & = & x\\ a & & +c & +d & = & y\\ a & +b & & +d & = & z\\ a & +b & +c & & = & w.\end{array}$$

Taking the sum of the equations gives $a+b+c+d=\frac{1}{3}(x+y+z+w)$. Then we can solve giving $a=\frac{1}{3}(-2x+y+z+w), b=\frac{1}{3}(x-2y+z+w)$, $c=\frac{1}{3}(x+y-2z+w), d=\frac{1}{3}(x+y+z-2w)$. Since there is a solution for all $x,y,z,w\in\mathbb{R}$, the matrices span $M_{2\times 2}(\mathbb{R})$. Also $x=y=z=w=0$ implies $a=b=c=d=0$, so the matrices are linearly independent. Hence they form a basis for $M_{2\times 2}(\mathbb{R})$.

(2) The matrix has eigenvalues $\lambda=1,2$. Since these are distinct, the matrix is diagonalisable.

(3) The matrix is 5×5 since the characteristic polynomial has degree 5. From the minimal polynomial we see that for $\lambda=2$ and $\lambda=3$ the Jordan blocks are 1×1; and for $\lambda=1$ there is a 2×2 Jordan block. Hence the Jordan normal form is

$$\begin{bmatrix}1 & 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 2 & 0 & 0\\ 0 & 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 0 & 3\end{bmatrix}.$$

(4) Possible inner products include:

$$\langle p,q\rangle=\int_0^1 p(x)q(x)\,dx,$$

or

$$\langle p,q\rangle=a_0b_0+a_1b_1+a_2b_2+a_3b_3+a_4b_4,$$

for $p=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4$, $q=b_0+b_1x+b_2x^2+b_3x^3+b_4x^4$. To verify that these give inner products, we need to check that for all $p,q,r\in\mathcal{P}_4(\mathbb{R})$ and all $\alpha,\beta\in\mathbb{R}$:

a) $\langle p,q\rangle=\langle q,p\rangle$

b) $\langle \alpha p+\beta q,r\rangle=\alpha\langle p,r\rangle+\beta\langle q,r\rangle$

c) $\langle p,p\rangle\geq 0$

d) $\langle p,p\rangle=0$ implies $p\equiv 0$.

(5) f^* is defined by the property:

$$\langle f^*(x),y\rangle=\langle x,f(y)\rangle,\ \text{for all } x,y.$$

Thus

$$\langle (f^*)^*(x),y\rangle=\langle x,f^*(y)\rangle=\langle f(x),y\rangle \text{ for all } x,y.$$

Hence

$$\langle (f^*)^*(x)-f(x),y\rangle=0 \text{ for all } x,y.$$

Taking $y=(f^*)^*(x)-f(x)$ in the last equation shows that $y=0$, hence $(f^*)^*(x)=f(x)$ for all x, i.e. $(f^*)^*=f$.

(6) The product is $(167253)(48) * (1645)(28) = (17243)(568)$. The order of the product is the least common multiple of 5 and 3, i.e. 15.

(7) $D_2 = C_2 \times C_2$ is abelian, but not cyclic since there is no element of order 4. (Each element satisfies $g^2 = e$.)

(8) The symmetry group has order 120. By Lagrange's theorem, each element of the group has order dividing $120 = 8 \times 3 \times 5$. Hence the group contains no element of order 9.

(9) G/N consists of the cosets $Ng, g \in G$. Multiplication in G/N satisfies

$$(Ng_1)(Ng_2) = Ng_1g_2 = Ng_2g_1 = (Ng_2)(Ng_1)$$

since G is abelian. Hence G/N is abelian.

(10) Let

$$g = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, h = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix};$$

these have non-zero determinants so belong to $GL(2, \mathbb{R})$. Then

$$hgh^{-1} = hgh = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

is conjugate to, but not equal to g.

Section B

(11) For any $u \in V$ we can write

$$u = a_1v_1 + \ldots + a_nv_n, \text{ for some } a_i \in \mathbb{C},$$

since $\{v_1, \ldots, v_n\}$ is a basis for V. Now

$$\begin{aligned} (u, v_i) &= (a_1v_1 + \ldots + a_nv_n, v_i) \\ &= a_1(v_1, v_i) + \ldots + a_i(v_i, v_i) + \ldots + a_n(v_n, v_i) = a_i, \end{aligned}$$

since $(v_i, v_i) = 1$ and $(v_i, v_j) = 0$ if $i \neq j$. Hence

$$u = (u, v_1)v_1 + (u, v_2)v_2 + \cdots + (u, v_n)v_n.$$

(12) Since $f^2 = 0$, the minimal polynomial $m(X)$ for f divides X^2. But since $f \neq 0$, $m(X) \neq X$. Hence, the minimal polynomial for f is $m(X) = X^2$. So the Jordan form of f contains a 2×2 Jordan block. Further, each eigenvalue of f is a root of $m(X) = X^2$, so the characteristic polynomial for f is $c(X) = X^3$. Hence the Jordan form of f is

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

Similarly g has minimal polynomial $m(X) = X^2$ and characteristic polynomial $c(X) = X^4$. Then the largest Jordan form for g is 2×2, so the possibilities for the Jordan normal form are:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

(13) For all $v \in V$,

$$(f(v), f(v)) = (f^*f(v), v) = (ff^*(v), v) = (f^*(v), f^*(v)).$$

Then $v \in \ker f$ iff $f(v) = 0$ iff $(f(v), f(v)) = 0$ iff $(f^*(v), f^*(v)) = 0$ iff $f^*(v) = 0$ iff $v \in \ker f^*$. Hence $\ker f = \ker f^*$.

(14) $|U_n| = n/2 = 2^{k-1}$ (since the number of "even" elements is equal to the number of "odd" elements in $\mathbb{Z}_n$). If $x = [a]_n \in U_n$ then $|x|$ divides $|U_n|$ by Lagrange's theorem, hence $x^{|U_n|}$ is the identity $[1]_n$ in U_n. This means that

$$a^{2^k-1} \equiv 1 \bmod 2^k,$$

i.e. $a^{2^{k-1}} - 1$ is a multiple of 2^k if a is an odd integer.

(15) First note that $abs : \mathbb{C}^* \to \mathbb{R}^*$ is a homomorphism since

$$abs(z_1 z_2) = |z_1 z_2| = |z_1| \cdot |z_2| = abs(z_1) \cdot abs(z_2).$$

$\ker(abs) = \{z : |z| = 1\}$ is the unit circle S^1 in $\mathbb{C}$; $\mathrm{im}(abs) = \{x \in \mathbb{R} : x > 0\}$ is the group of positive reals Pos under multiplication.
Hence $\mathbb{C}^*/S^1 = \mathbb{C}^*/\ker(abs)$ is isomorphic to the subgroup $\mathrm{im}(abs) = \mathrm{Pos}$ of $\mathbb{R}^*$.

(16) Each orbit is contained in X so has at most $|X| = 18$ elements. Further, by the orbit-stabiliser relation, each orbit has size dividing $|G| = 35$; so each orbit must contain 1,5 or 7 elements. Since the orbits are disjoint and fill up X, we also know that the sum of sizes of all orbits is equal to $|X| = 18$. Now it is easy to check that 18 can not be written as a sum of 5's and 7's. Hence there must be an orbit containing a single element $x \in X$. But then x is fixed by each element of g, as desired.
Let G be the group generated by the permutation

$$(1, 2, 3, 4, 5)(6, 7, 8, 9, 10, 11, 12)$$

acting on $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ in the usual way. Then $|G| = 35$, $|X| = 12$ and the orbits are $\{1, 2, 3, 4, 5\}$ and $\{6, 7, 8, 9, 10, 11, 12\}$. So this gives an example where no element of X is fixed by all of G.

5.8 Solutions to the 2000 Examination

Section A

(1) We first check if these polynomials span $\mathcal{P}_2(\mathbb{R})$. Given $a, b, c \in \mathbb{R}$ we try to find α, β, γ satisfying

$$\alpha(1-x+x^2)+\beta(1+x+x^2)+\gamma(1+x-x^2)=a+bx+cx^2.$$

This is equivalent to

$$\alpha+\beta+\gamma=a, -\alpha+\beta+\gamma=b, \alpha+\beta-\gamma=c.$$

Reducing the coefficient matrix to row echelon form gives

$$\begin{bmatrix} 1 & 1 & 1 & \mid & a \\ -1 & 1 & 1 & \mid & b \\ 1 & 1 & -1 & \mid & c \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 1 & \mid & a \\ 0 & 2 & 2 & \mid & a+b \\ 0 & 0 & -2 & \mid & c-a \end{bmatrix}.$$

Thus the equations have a solution for all a, b, c, and the polynomials span $\mathcal{P}_2(\mathbb{R})$. Further, taking $a=b=c=0$ we see that the only solution is $\alpha=\beta=\gamma=0$, so the polynomials are linearly independent. Hence they form a basis for $\mathcal{P}_2(\mathbb{R})$.

(2) The matrix is triangular, so its eigenvalues are the diagonal entries and the characteristic polynomial is $c(X)=(X-1)^2(X-2)$. The minimal polynomial $m(X)$ divides $c(X)$ and has the eigenvalues $1, 2$ as roots, so $m(X)=(X-1)(X-2)$ or $(X-1)^2(X-2)$. But

$$(A-I)(A-2)=\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

hence the minimal polynomial is $m(X)=(X-1)(X-2)$.

(3) Since $c(X)=X^2(X-1)^4$, the matrix is 6×6 with diagonal entries $0, 0, 1, 1, 1, 1$. Since $m(X)=X^2(X-1)^2$ the largest Jordan block with 0 on the diagonal is 2×2 and the largest Jordan bloack with 1 on the diagonal is 2×2. This gives two possibilities for the Jordan normal form (up to reordering Jordan blocks):

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(4) We have $(x, y, z)\in V^\perp$ iff $(x, y, z)\cdot(1, 1, 0)=0$ and $(x, y, z)\cdot(0, 1, 2)=0$. This gives the system of linear equations

$$x+y=0, y+2z=0$$

with general solution $z=t, y=-2t, x=2t$ with $t\in\mathbb{R}$. Hence

$$V^\perp=\{(2t, -2t, t): t\ \in\mathbb{R}\}=\{t(2, -2, 1): t\ \in\mathbb{R}\}.$$

(5) One example is $A=\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$. This satisfies $A^*=A$ so is Hermitian. The matrix is also normal since $A^*A=A^2=AA^*$.

(6) (i) P is not a group under addition. For example, there is no identity since $0\notin P$.

(ii) P is a group under multiplication:
(0) if $x>0, y>0$ then $xy>0$, so P is closed under multiplication
(1) $e=1$ is an identity element
(2) if $x>0$ then $x^{-1}=\frac{1}{x}\in P$ so inverses exist.
(3) multiplication is associative in $\mathbb{R}$, hence also in P.

(7) $(135)(2678) * (14)(23578) = (143)(25867)$.
This permutation has order $\operatorname{lcm}(3,5) = 15$.

(8) The group $\langle [2]_9 \rangle$ generated by $[2]_9$ is

$$\{[2]_9, [2]_9^2 = [4]_9, [2]_9^3 = [8]_9, [2]_9^4 = [7]_9, [2]_9^5 = [5]_9, [2]_9^6 = [1]_9\} = H.$$

Hence H is a cyclic group, and $[2]_9$ is a generator. (Note: $[5]_9$ is also a generator.)

(9) By Lagrange's theorem, $|G|$ is divisible by 10 and 25. Hence $|G|$ is a multiple of $\operatorname{lcm}(10,25) = 50$. Since $|G| < 100$ we must have $|G| = 50$.

(10) The orbit of a face F is the set of all six faces of the cube. The stabiliser of F is the cyclic group of order 4 generated by a 90 degree rotation about the axis passing through the centre of face F and the centre of the opposite face. Hence $|G| = |\text{orbit}(F)| \cdot |\text{stabilizer}(F)| = 6 \times 4 = 24$.

Section B

(11) a) Let $\alpha = a + b\sqrt{2}$. Then for $v, w \in V$ and $c \in \mathbb{Q}$ we have

$$f(v+w) = \alpha(v+w) = \alpha v + \alpha w = f(v) + f(w)$$

and

$$f(cv) = \alpha(cv) = c(\alpha v) = cf(v)$$

using the properties of addition and multiplication in the field $\mathbb{R}$. So f is a linear transformation.

b) To find the matrix of f, we apply f to the basis vectors $\{1, \sqrt{2}\}$ and write the images as linear combinations of these basis vectors with $\mathbb{Q}$ coefficients. We have

$$\begin{aligned} f(1) &= a + b\sqrt{2} &&= a \cdot 1 + b \cdot \sqrt{2} \\ f(\sqrt{2}) &= (a + b\sqrt{2})\sqrt{2} = 2b \cdot 1 + a \cdot \sqrt{2}. \end{aligned}$$

These coefficients give the columns of the matrix for f:

$$[f] = \begin{bmatrix} a & 2b \\ b & a \end{bmatrix}.$$

c) The nullspace of f is $\{v \in V : \alpha v = 0\} = \{0\}$ since $\alpha \neq 0$; hence $\text{nullity}(f) = 0$. Since $\text{rank}(f) + \text{nullity}(f) = \dim V = 2$, we have $\text{rank}(f) = 2$.

d) Since $\operatorname{im} f$ is a subspace of V and $\dim(\operatorname{im} f) = \text{rank}(f) = 2 = \dim V$, we have $\operatorname{im} f = V$ and f is surjective.

(12) The Spectral Theorem states that if $f : V \to V$ is a normal linear transformation on a finite dimensional complex inner product space V, then there exists an orthonormal basis for V consisting of eigenvectors for f. In matrix form, this shows that if A is a normal matrix with complex entries, then there is a unitary matrix U and a diagonal matrix D such that $U^*AU = D$ or $A = UDU^*$.

Let $\lambda_1, \ldots, \lambda_n$ be the diagonal entries of D, and let $\mu_i = \sqrt[3]{\lambda_i}$ be a complex cube root of λ_i for each i. Then the matrix E with diagonal entries $\mu_1, \ldots, \mu_n$ satisfies $E^3 = D$. Hence the matrix $B = UEU^*$ satisfies

$$B^3 = UEU^*UEU^*UEU^* = UE^3U^* = UDU^* = A,$$

since $U^*U = I$

(13) a) Let v be an eigenvector of f with eigenvalue λ. Then

$$(v,v) = (f^*f(v), v) = (f(v), f(v)) = (\lambda v, \lambda v) = \lambda\overline{\lambda}(v,v) = |\lambda|^2(v,v),$$

since $f^*f = \text{identity}$. Since v is non-zero, $(v,v) \neq 0$ hence $|\lambda|^2 = 1$ or $|\lambda| = 1$.

b) Let v_1, v_2 be eigenvectors of f corresponding to eigenvalues $\lambda_1 \neq \lambda_2$. Then $f(v_1) = \lambda_1 v_1$ and $f(v_2) = \lambda_2 v_2$, so

$$(v_1, v_2) = (f^* f(v_1), v_2) = (f(v_1), f(v_2)) = (\lambda_1 v_1, \lambda_2 v_2) = \lambda_1 \overline{\lambda_2}(v_1, v_2).$$

Since $\lambda_2 \overline{\lambda_2} = |\lambda_2|^2 = 1$ by part (a), $\overline{\lambda_2} = \frac{1}{\lambda_2}$. Hence we have

$$(v_1, v_2) = \frac{\lambda_1}{\lambda_2}(v_1, v_2) \text{ or } \lambda_2(v_1, v_2) = \lambda_1(v_1, v_2).$$

Since $\lambda_1 \neq \lambda_2$, this implies that $(v_1, v_2) = 0$, so v_1, v_2 are orthogonal.

(14) a) For any $a, b \in G$ we have $(ab)^2 = abab = e$. Multiplying on the left by a and on the right by b gives $aababb = aeb$ or $ebae = ab$, since $a^2 = e$ and $b^2 = e$. Thus $ba = ab$ and the group is abelian.

b) Assume G is a *finite group*. By Cauchy's theorem, if p is a prime dividing $|G|$, then G contains an element of order p. But each element of G has order 1 or 2, so the only prime dividing $|G|$ is 2, and $|G| = 2^k$ for some integer $k \geq 0$. [Note: in fact $G \cong \mathbb{Z}_2^k = \mathbb{Z}_2 \times \ldots \mathbb{Z}_2$.]

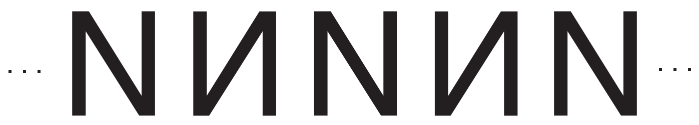
(15) a) For all $z_1, z_2 \in \mathbb{C}^*$, we have $f(z_1 z_2) = (z_1 z_2)^2 = z_1^2 z_2^2 = f(z_1) f(z_2)$, so f is a homomorphism.

b) $\ker f = \{z \in \mathbb{C}^* : z^2 = 1\} = \{\pm 1\}$
$\operatorname{im} f = \{z^2 : z \in \mathbb{C}^*\} = \mathbb{C}^*$

c) By the isomorphism theorem, $\mathbb{C}^* / \ker f \cong \operatorname{im} f = \mathbb{C}^*$.

d) f is not injective (i.e. not $1-1$), since $f(-1) = f(1) = 1$. So f is not an isomorphism.

(16) Let G be the group of all symmetries of the frieze pattern shown below. (The pattern repeats to fill out an infinite strip in the plane.)



a) (i) the centres of rotations are the centres of the N's.

(ii) mirror lines are vertical lines between each N and the adjacent reflected N.

(iii) axes of glide reflections are horizontal, through the middle of the N's.

b) The translation subgroup is the infinite cyclid group generated by a horizontal translation taking one N to the next (non-reflected) N. The point group is the dihedral group D_2 of order 4, consisting of the identity, two reflections about horizontal and vertical axes through a point, and the 180 degree rotation about that point.

c) The group G is not abelian. For example, if r_1, r_2 are two relections in different vertical axes, then $r_1 r_2 \neq r_2 r_1$ since $r_1 r_2$ and $r_2 r_1$ are translations in opposite directions.

Bibliography

[1] B. Fine and G. Rosenberger. *The Fundamental Theorem of Algebra.* Springer Undergraduate Texts in Mathematics and Technology. Springer New York, 1997.

[2] T. W. Hungerford. *Abstract algebra: an introduction.* Saunders College Publishing, second edition, 1997.

[3] S. Lang. *Algebra.* Graduate Texts in Mathematics. Springer New York, 2002.