Lecture Notes for

MAST20022 Group Theory and Linear Algebra

Lawrence Reeves
School of Mathematics and Statistics
University of Melbourne

Version 2.0

Contents

Chapter 1

Modular Arithmetic and Fields

We begin with some number theory, looking at divisibility properties of the natural numbers and the integers.

$$\begin{aligned} \textbf{natural numbers}: &\quad \mathbb{N} = \{1, 2, 3, 4, \dots\} \\ \textbf{integers}: &\quad \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\} \end{aligned}$$

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

1 Well-ordering and induction

An important property of the natural numbers that we will need is the following:

Well-ordering property: (WOP)

Every non-empty subset of $\mathbb{N}$ has a smallest element.

It is equivalent to the following:

Principle of mathematical induction: (PMI)

Suppose that we have a set of statements $\{S(n) \mid n \in \mathbb{N}\}$ satisfying:

1) $S(1)$ is true
2) $\forall n \in \mathbb{N} \quad S(n) \implies S(n+1)$

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

Which is equivalent to:

Strong form of induction: (SMI)

Suppose that we have a set of statements $\{S(n) \mid n \in \mathbb{N}\}$ satisfying:

1) $S(1)$ is true
2) $\forall n \in \mathbb{N} \quad (S(1) \wedge S(2) \wedge \cdots \wedge S(n)) \implies S(n+1)$

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

We will show that the above three properties of $\mathbb{N}$ are equivalent by showing that

$$\text{WOP} \implies \text{PMI} \implies \text{SMI} \implies \text{WOP}$$

WOP $\implies$ PMI: Suppose first that the well-ordering property holds. Assume that $S(n)$ are as in the statement of the principle of mathematical induction. We need to show that $S(n)$ is true for all $n \in \mathbb{N}$. Let $E = \{n \in \mathbb{N} \mid S(n) \text{ is false}\}$. Suppose, for a contradiction, that E is non-empty. By the WOP, E has a minimum element, call it $m \in E$. Since $S(1)$ is true, we have that $m \neq 1$. Therefore $m - 1 \in \mathbb{N}$ and $S(m-1)$ is true by the minimality of m. But $S(m-1)$ is true implies that $S(m)$ is true. From this contradiction we conclude that $E = \emptyset$.

PMI $\implies$ SMI: Exercise!

SMI $\implies$ WOP: Assume that SMI holds. Let $E \subseteq \mathbb{N}$ be such that E does not have a smallest element. We want to show that $E = \emptyset$. Let $S(n)$ be the statement $n \notin E$. Then $S(1)$ is true since otherwise $1 \in E$ would be minimal. Suppose that $S(1)$, S(2), $\dots, S(n)$ are all true. Then $1 \notin E, \dots, n \notin E$. If $n \in E$, then n would be minimal. Since E has no minimal element, we must have that $S(n)$ is true. From SMI we conclude that $S(n)$ is true for all $n \in \mathbb{N}$, that is, $E = \emptyset$.

Exercise 1. Show that WOP holds for every subset of $\mathbb{Z}$ that is bounded below. That is, if $E \subseteq \mathbb{Z}$ is bounded below, then every non-empty subset of E has a minimal element. What about subsets of $\mathbb{Q}$ (or $\mathbb{R}$) that are bounded below?

2 Integer division

Theorem 1.1

Let $a \in \mathbb{Z}$ and $d \in \mathbb{N}$. Then there exist $q, r \in \mathbb{Z}$ such that

$$a = qd + r \quad \text{and} \quad 0 \leqslant r < d$$

Moreover, q and r are uniquely determined by a and d.

The integers q and d are known as the **quotient** and **remainder** respectively.

Proof. Given $a \in \mathbb{Z}$ and $d \in \mathbb{N}$ define $E = \{k \in \mathbb{Z} \mid k \geqslant 0 \text{ and } k = a - qd \text{ for some } q \in \mathbb{Z}\}$. Note that $E \neq \emptyset$ since $a - (-|a|)d = a + d|a| \geqslant 0$. By WOP, E has a minimal element $r \in E$. Since $r \in E$, we have $r \geqslant 0$ and there is a $q \in \mathbb{Z}$ such that $r = a - qd$. Also, $r < d$ since

$$r \text{ is minimal in } E \implies r - d \notin E \implies r - d = a - (q+1)d \notin E \implies r - d < 0 \implies r < d$$

The uniqueness of q and r is left as an exercise. □

Remark. There are versions of this result that hold when $\mathbb{Z}$ is replaced by $\mathbb{R}[X]$ or $\mathbb{Z}[i]$ (and others). The proof is essentially the same.

Definition 1.2. Let $a, d \in \mathbb{Z}$. We say that d **divides** a if $\exists\, q \in \mathbb{Z}$ such that $a = qd$. It is often denoted by $d \mid a$. We also say that d is a **divisor** of a.

Lemma 1.3

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

1) $(a \mid b) \wedge (b \mid c) \implies a \mid c$

2) $(a \mid b) \wedge (a \mid c) \implies \forall x, y \in \mathbb{Z}, \quad a \mid (xb + yc)$

3) $(a \mid b) \wedge (b \mid a) \implies a = \pm b$

4) $a \mid 1 \implies a = \pm 1$

Proof. Left as an exercise. □

Definition 1.4. Let $a, b \in \mathbb{Z}$. A **greatest common divisor** (gcd) of a and b is an element $d \in \mathbb{Z}$ such that

1) $(d \mid a) \wedge (d \mid b)$

2) $\forall c \in \mathbb{Z}, \quad (c \mid a) \wedge (c \mid b) \implies c \mid d$

Lemma 1.5

Let $a, b \in \mathbb{Z}$ be such that at least one of a and b is non-zero. Then there is a unique $d \in \mathbb{N}$ such that d is a gcd of a and b. It will be denoted $d = \gcd(a, b)$.

Proof. We first show that there exists a greatest common divisor $d \in \mathbb{N}$. Let $E = \{k > 0 \mid \exists\, x, y \in \mathbb{Z}, \quad k = xa+yb\}$. Then $E \subseteq \mathbb{N}$ and $E \neq \emptyset$ (since $a^2 + b^2 \in E$). By the WOP, E has a minimal element $d \in S$. Fix $x, y \in \mathbb{Z}$ such that $d = xa + yb$. By Theorem 1.1 there exist $q, r \in \mathbb{Z}$ with $a = qd + r$ and $0 \leqslant r < d$. Then

$$r = a - qd = a - q(xa + yb) = (1 - qx)a + (-qy)b$$

We must, therefore, have $r = 0$ since otherwise $r \in E$ and $r < d$. Therefore $a = qd$, that is, $d \mid a$.

Similarly, $d \mid b$.

Now suppose that $c \mid a$ and $c \mid b$. Then $c \mid d = xa + yb$ (see Lemma 1.3).

Now for uniqueness. Suppose that d and d' are both satisfy the conditions for being a greatest common divisor of a and b. Then $d \mid d'$ and $d' \mid d$ which implies that $d' \pm d$. Since $d, d' \in \mathbb{N}$, we conclude that $d' = d$. □

Given the uniqueness pointed out in the above result we use the notation $\gcd(a, b)$ for *the* greatest common divisor of two integers a and b. In fact, the above proof establishes the following result.

Theorem 1.6: Bézout's Theorem

Let $a, b \in \mathbb{Z}$. Then there exist $x, y \in \mathbb{Z}$ such that $\gcd(a, b) = xa + yb$. □

Definition 1.7. We say that two integers $a, b \in \mathbb{Z}$ are **relatively prime** if $\gcd(a, b) = 1$.

2.1 Exercises

Exercise 2. Show that two integers $a, b \in \mathbb{Z}$ are relatively prime if and only if $\exists\, x, y \in \mathbb{Z}, \quad xa + yb = 1$.

Exercise 3. Find the quotient and remainder when:

(a) 25 is divided by 3 (b) 68 is divided by 7 (c) -33 is divided by 7

Exercise 4. Prove Lemma 1.3.

Exercise 5. Prove the uniqueness of q and d in Theorem 1.1. That is, show that if $qd+r = q'd+r'$ with $0 \leqslant r, r' < d$, then $q = q'$ and $r = r'$.

Exercise 6. Let $a, b, c \in \mathbb{Z}$ be integers with $\gcd(a, b) = 1$. Show that if $a \mid c$ and $b \mid c$, then $ab \mid c$.

Exercise 7. Let F_n be the n-th Fibonacci number, defined by $F_0 = 0$, $F_1 = 1$ and $F_{k+2} = F_k + F_{k+1}$.

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

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

3 Euclidean algorithm

The greatest common divisor of two integers can be computed by first finding the prime factorisations of the given integers. However, a mush more efficient method is given by the Euclidean algorithm. It is based on the following observations.

Lemma 1.8

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

1) $\gcd(a, b) = \gcd(b, a) = \gcd(-a, b)$
2) $\gcd(a, 0) = |a|$
3) If $a = qb + r$, then $\gcd(a, b) = \gcd(b, r)$

Proof. We shall prove 3) and leave the rest as an exercise. Since $\gcd(a, b)$ divides both a and b, we have

$$\gcd(a, b) \mid b \quad \text{and} \quad \gcd(a, b) \mid r = (a - qb)$$

which implies that $\gcd(b, r) \mid \gcd(a, b)$. Similarly,

$$\gcd(b, r) \mid a = qb + r \quad \text{and} \quad \gcd(b, r) \mid b$$

which implies that $\gcd(a, b) \mid \gcd(b, r)$. Since both are positive, we have $\gcd(a, b) = \gcd(b, r)$. □

Given $a \geqslant b > 0$, define q_i and r_i as follows:

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

Since the r_i are strictly decreasing, we must eventually arrive at a remainder of zero. By Lemma 1.8 we have

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

That is, the greatest common divisor of a and b is given by the last non-zero remainder obtained.

Example 1.9. We calculate the greatest common divisor of 4163 and 8869.

$$8869 = 2 \times 4163 + 543 \tag{1.1}$$
$$4163 = 7 \times 543 + 362 \tag{1.2}$$
$$543 = 1 \times 362 + 181 \tag{1.3}$$
$$362 = 2 \times 181 + 0 \tag{1.4}$$

Therefore $\gcd(4163, 8869) = 181$
Note that we can also express the greatest common divisor as a linear combination of a and b.
Working back through the above calculation, we get

$$\begin{aligned} 181 &= 543 - 362 && \text{(from 1.3)} \\ &= 543 - (4163 - 7 \times 543) && \text{(from 1.2)} \\ &= -4163 + 8 \times 543 \\ &= -4163 + 8(8869 - 2 \times 4163) && \text{(from 1.1)} \\ &= -17 \times 4163 + 8 \times 8869 \end{aligned}$$

3.1 Exercises

Exercise 8. Using the Euclidean Algorithm (by hand) find:

(a) $\gcd(14, 35)$
(b) $\gcd(105, 165)$
(c) $\gcd(1287, 1144)$
(d) $\gcd(1288, 1144)$
(e) $\gcd(1287, 1145)$

Exercise 9. 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) $(312, 377)$

Exercise 10. Complete the proof of Lemma 1.8

4 Primes

Lemma 1.10

Let $a, b, c \in \mathbb{Z}$. If $\gcd(a, b) = 1$ and $a \mid bc$, then $a \mid c$.

Proof. We have

$$xa + yb = 1 \qquad \text{for some } x, y \in \mathbb{Z}$$

and

$$bc = az \qquad \text{for some } z \in \mathbb{Z}$$

therefore

$$xac + ybc = c \implies xac + yaz = c \implies (xc + yz)a = c \implies a \mid c$$

□

Definition 1.11. A natural number $p \in \mathbb{N}$ is called **prime** if $p \neq 1$ and $\forall\, a, b \in \mathbb{Z}, \quad p \mid ab \implies (p \mid a) \vee (p \mid b)$

Lemma 1.12

A natural number $p \in \mathbb{N}$ is prime if and only if it has exactly two divisors in $\mathbb{N}$.

Proof. Assume that p is prime. Then $p \neq 1$ and both 1 and p are divisors of p. We need to show that these are the only positive divisors of p. Suppose that $p = ab$ for some $a, b \in \mathbb{N}$. Since $p \mid p = ab$ and p is prime, we have that $p \mid a$ or $p \mid b$. Note that

$$\begin{aligned} p \mid a &\implies a = pq && \text{(some } q \in \mathbb{N}\text{)} \\ &\implies p = pqb && \\ &\implies qb = 1 && \text{(since } p \neq 0\text{)} \\ &\implies b = 1 && \text{(since } b, q \in \mathbb{N}\text{)} \\ &\implies a = p && \text{(since } p = ab\text{)} \end{aligned}$$

Similarly, if $p \mid b$, then $a = 1$ and $b = p$.

For the converse, assume that p has exactly two positive divisors. First note that $p \neq 1$ by Lemma 1.3. The two distinct positive divisors of p are therefore 1 and p. Let $a, b \in \mathbb{Z}$ be such that $p \mid ab$. We want to show that either $p \mid a$ or $p \mid b$, which is equivalent to showing that $p \nmid a \implies p \mid b$.

$$\begin{aligned} p \nmid a &\implies \gcd(a, p) = 1 && \text{(since } \gcd(a, p) \mid p\text{)} \\ &\implies p \mid b && \text{(by Lemma 1.10 since } p \mid ab\text{)} \end{aligned}$$

□

Theorem 1.13: The Fundamental Theorem of Arithmetic

Let $n \in \mathbb{N}$ with $n \geqslant 2$. Then there exist $k \in \mathbb{N}$, and $i_1, \ldots, i_k \in \mathbb{N}$, and primes $p_1 < p_2 < \cdots < p_k \in \mathbb{N}$ such that

$$m = p_1^{i_1} p_2^{i_2} \ldots p_k^{i_k}$$

Moreover, this expression is uniquely determined by m.

Proof. Left as an exercise. Hint: For existence use Strong Mathematical Induction. □

4.1 Exercises

Exercise 11. (a) Give an example of natural numbers a, b, c such that $a \mid c$ and $b \mid c$ but $ab \not| c$.

(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 $ab \mid c$.

Exercise 12. Prove Theorem 1.13.

5 Modular arithmetic

5.1 Congruence

Definition 1.14. Fix $m \in \mathbb{N}$. We say that $a, b \in \mathbb{Z}$ are **congruent modulo m** if $m \mid (a - b)$. It is denoted

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

Example 1.15.

$$\begin{aligned} 42 &\equiv 6 \pmod{4} \\ 77 &\equiv -4 \pmod{9} \\ 15 &\equiv 0 \pmod{5} \end{aligned}$$

Remark. It follows from Theorem 1.1 that given $a \in \mathbb{Z}$ and $m \in \mathbb{N}$, there exists a unique $r \in \mathbb{Z}$ with $0 \leqslant r < m$ such that $a \equiv r \pmod{m}$.

The following are fundamental properties of the congruence relation. They say that, for a fixed m, being congruent modulo m is an "equivalence relation".

Lemma 1.16: congruence is an equivalence relation

Let $m \in \mathbb{N}$ and $a, b, c \in \mathbb{Z}$. Then

1) $a \equiv a \pmod{m}$ (reflexive)
2) $a \equiv b \pmod{m} \implies b \equiv a \pmod{m}$ (symmetric)
3) $a \equiv b \pmod{m} \wedge b \equiv c \pmod{m} \implies a \equiv c \pmod{m}$ (transitive)

Proof. For the first two, note that $m \mid 0$ and that $(a - b) \mid (b - a)$. For the third note that

$$(m \mid (a - b)) \wedge (m \mid (b - c)) \implies m \mid (a - b) + (b - c) = a - c$$

□

The next result is that congruence works well with the arithmetic operations on $\mathbb{Z}$.

Lemma 1.17

Let $m, n \in \mathbb{N}$ and $a, b, c, d \in \mathbb{Z}$. Suppose that $a \equiv c \pmod{m}$ 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) $a^n \equiv c^n \pmod{m}$

Proof. For the first note that

$$m \mid (a-c) \wedge m \mid (b-d) \implies m \mid ((a-c)+(b-d)) \implies m \mid ((a+b)-(c+d)) \implies a+b \equiv c+d \pmod{m}$$

For the third note that

$$\begin{aligned} m \mid (a-c) \wedge m \mid (b-d) &\implies (a-c=mk) \wedge (b-d=ml) && \text{(for some } k,l \in \mathbb{Z}\text{)} \\ &\implies (a=mk+c) \wedge (b=ml+d) \\ &\implies ab = cd + cml + mkd + m^2kl \\ &\implies ab - cd = m(cl+kd+mkl) \\ &\implies ab \equiv cd \pmod{m} \end{aligned}$$

The remaining cases are left as an exercise. □

Example 1.18. We find an r with $0 \leqslant r < 12$ such that $29^4 \equiv r \pmod{12}$. Rather than calculating 29^4, we will use Lemma 1.17. Noting first that $29 \equiv 5 \pmod{12}$, we have

$$\begin{aligned} & 29^4 \equiv 5^4 \pmod{12} && \text{(Lemma 1.17.4)} \\ \implies & 29^4 \equiv 25^2 \pmod{12} \\ \implies & 29^4 \equiv 1^2 \pmod{12} && \text{(since } 25 \equiv 1 \pmod{12}\text{)} \\ \implies & 29^4 \equiv 1 \pmod{12} \end{aligned}$$

Definition 1.19. Let $a \in \mathbb{Z}$ and $m \in \mathbb{N}$. The **congruence class** of a modulo m is the following subset of $\mathbb{Z}$ which is denoted $[a]_m$.

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

Example 1.20.

$$\begin{aligned} [0]_3 &= \{\dots,-6,-3,0,3,6,\dots\} \\ [1]_3 &= \{\dots,-5,-2,1,4,7,\dots\} \\ [2]_3 &= \{\dots,-4,-1,2,5,8,\dots\} \\ [3]_3 &= \{\dots,-6,-3,0,3,6,\dots\} \end{aligned}$$

Exercise 13. Prove the following:

(a) $[a]_m = [b]_m$ if and only if $a \equiv b \pmod{m}$

(b) $[a]_m \cap [b]_m \neq \emptyset \implies [a]_m = [b]_m$

(c) $[0]_m \cup [1]_m \cup \cdots \cup [m-1]_m = \mathbb{Z}$

5.2 Integers modulo m

Definition 1.21. The set of congruence classes is denoted $\mathbb{Z}/m\mathbb{Z}$ and is called the **integers modulo** m. That is,

$$\mathbb{Z}/m\mathbb{Z} = \{[0]_m, \dots, [m-1]_m\}$$

Remark. Notice that $|\mathbb{Z}/m\mathbb{Z}| = m$

Example 1.22. $\mathbb{Z}/3\mathbb{Z} = \{[0]_3, [1]_3, [2]_3\}$

We define two binary operations on $\mathbb{Z}/m\mathbb{Z}$ in the following way:

$$\begin{aligned} [a]_m + [b]_m &= [a+b]_m \\ [a]_m \times [b]_m &= [a \times b]_m \end{aligned}$$

Remark.

1) It is important to realise that what is being defined here is what the symbols '+' and '×' mean when used as on the left. On the right of the above definitions the same symbols are used to represent the usual operations of addition and multiplication in $\mathbb{Z}$.

2) Each binary relation is a function $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$

3) These operations are 'well-defined', meaning that if a, b, α, β are such that $[a]_m = [\alpha]_m$ and $[b]_m = [\beta]_m$, then $[a+b]_m = [\alpha+\beta]_m$ and $[a \times b]_m = [\alpha \times \beta]_m$.

4) As with $\mathbb{Z}$, we often omit the symbol '$\times$' and write $[a]_m[b]_m$ in place of $[a]_m \times [b]_m$.

5) $[a_m]_m[1]_m = [a]_m$ and $[a_m]_m + [0]_m = [a]_m$for all $[a]_m \in \mathbb{Z}/m\mathbb{Z}$.

6) Equipped with these operations, $\mathbb{Z}/m\mathbb{Z}$ forms what is called a 'commutative ring'.

Definition 1.23. We say that $[a]_m$ is the **multiplicative inverse** of $[b]_m$ if $[a]_m[b]_m = [1]_m$.

Exercise 14. Show that $[a]_m$ has at most one multiplicative inverse in $\mathbb{Z}/m\mathbb{Z}$.

Example 1.24. The multiplication table for $\mathbb{Z}/6\mathbb{Z}$ is shown. Note that in this table we have written a in place of $[a]_6$. The element $[5]_6$ is the multiplicative inverse of itself. The element $[2]_6$ has no multiplicative inverse.

$(\mathbb{Z}/6\mathbb{Z}, \times)$

	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Theorem 1.25

Let $m \in \mathbb{N}$ and $a \in \mathbb{Z}$. Then $[a]_m$ has a multiplicative inverse if and only if $\gcd(a, m) = 1$.

Proof. Suppose that $[a]_m$ has a multiplicative inverse. That is, there exists $b \in \mathbb{Z}$ such that $[a]_m[b]_m = [1]_m$. Then note that

$$\begin{aligned}
[a]_m[b]_m = [1]_m &\implies [ab]_m = [1]_m \\
&\implies ab \equiv 1 \pmod{m} \\
&\implies m \mid (ab-1) \\
&\implies ab - 1 = mk && \text{(for some } k \in \mathbb{Z}) \\
&\implies ab - mk = 1 \\
&\implies \gcd(a, m) = 1 && \text{(see Exercise 2)}
\end{aligned}$$

For the converse, we have

$$\begin{aligned}
\gcd(a, m) = 1 &\implies xa + my = 1 && \text{(for some } x, y \in \mathbb{Z}\text{, by Theorem 1.6)} \\
&\implies xa - 1 = -my \\
&\implies xa \equiv 1 \pmod{m} \\
&\implies [x]_m[a]_m = [1]_m
\end{aligned}$$

□

Corollary 1.26

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

5.3 Exercises

Exercise 15. Write down the multiplication tables for $\mathbb{Z}/7\mathbb{Z}$ and $\mathbb{Z}/8\mathbb{Z}$.

Exercise 16. 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) $-2 \equiv 933 \pmod{55}$

Exercise 17. Simplify the following, writing your answers in the form $a \bmod m$ where $0 \le a < m$.

(a) $482 \pmod{14}$

(b) $511 \pmod{9}$

(c) $-374 \pmod{11}$

(d) $933 \pmod{55}$

(e) $102725 \pmod{10}$

(f) $57102725 \pmod{9}$

Exercise 18. Calculate the following, giving answers in the form $a \bmod m$ where $0 \le a < m$.
(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) $1 \times 2 \times 3 \times \cdots \times 20 \times 21 \pmod{22}$

Exercise 19. Use congruences modulo 9 to show that the following multiplication is incorrect:

$$326 \times 4471 = 1357546.$$

Exercise 20. 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.

Exercise 21. 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) $\mathbb{Z}/15\mathbb{Z}$

Exercise 22. Using the Euclidean 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 $\mathbb{Z}/911\mathbb{Z}$

Exercise 23. Find the smallest positive integer giving a remainder of 3 when divided by 7, and a remainder of 8 when divided by 11.

Exercise 24. (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 and it is unique modulo $m_1 m_2$.
Generalise to an arbitrary number of congruences.

6 Fields

Definition 1.27. A **commutative ring** is a set R together with two binary operations '+' and '×' satisfying

1) $\forall\, x, y, z \in R, \quad (x + y) + z = x + (y + z)$ (addition is associative)

2) $\forall\, x, y \in R, \quad x + y = y + x$ (addition is commutative)

3) $\exists\, 0 \in R\, \forall x \in R, \quad x + 0 = x$ (additive identity)

4) $\forall x \in R\, \exists y \in R, \quad x + y = 0$ (additive inverses)

5) $\forall\, x, y, z \in R, \quad (x \times y) \times z = x \times (y \times z)$ (multiplication is associative)

6) $\forall\, x, y \in R, \quad x \times y = y \times x$ (multiplication is commutative)

7) $\exists\, 1 \in R\, \forall x \in R, \quad 1 \times x = x$ (multiplicative identity)

8) $\forall\, x, y, z \in R, \quad x \times (y + z) = (x \times y) + (x \times z)$ (distriibutivity)

Example 1.28. 1. $\mathbb{Z}$ with the usual operations is a commutative ring

2. $\mathbb{N}$ with the usual operations is not a commutative ring (why not?)

3. $\mathbb{Z}/m\mathbb{Z}$ with the operations defined above is a commutative ring

4. $\mathbb{R}[X]$, the set of all polynomials equipped with the usual operations is a commutative ring

5. $\{*\}$ with operations given by $* + * = *$ and $* \times * = *$ is a commutative ring

Exercise 25. Show that the additive inverse of an element $x \in R$ (whose existence is given by axiom 4 above) is unique. It is denoted $-x$.

Exercise 26. Let R be a commutative ring, and $x \in R$ any two elements. Show that

a) $0x = 0$

b) $(-1)x = -x$

Exercise 27. Let R be a commutative ring. Show that if $1 = 0$ (i.e., the additive and multiplicative identities coincide), then R consists of a single element.

Definition 1.29. A **field** is a commutative ring having at least two elements and in which every non-zero element has a multiplicative inverse.

Example 1.30. 1. $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are fields

2. $\mathbb{Z}$ is not a field

3. $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\} \subset \mathbb{R}$ is a field

4. $\mathbb{Z}/p\mathbb{Z}$ for any prime p is a field. It will sometimes be denoted $\mathbb{F}_p$.

5. $\{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\} \subset M_2(\mathbb{Z}/2\mathbb{Z})$ with the usual matrix operations is a field.

Definition 1.31. A field K is called **algebraically closed** if every non-constant polynomial over K has a root in K. That is,

$$\forall p(X) \in K[X], \quad \deg(p(X)) \geqslant 1 \implies (\exists\, k \in K, \quad p(k) = 0)$$

Example 1.32. 1. $\mathbb{Q}$ and $\mathbb{R}$ are not algebraically closed since $X^2 + 1$ has no root in $\mathbb{R}$

2. $\mathbb{F}_2$ is not algebraically closed since $X^2 + X + 1$ has no root in $\mathbb{F}_2$

We will need the following facts, we we state without proof.

Theorem 1.33: Fundamental Theorem of Algebra

The field $\mathbb{C}$ is algebraically closed.

Theorem 1.34: Algebraic closure

Any field can be embedded in an algebraically closed field.

6.1 Exercises

Exercise 28. 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 $\mathbb{R}$.)

Exercise 29. 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 of $\mathbb{Q}[\sqrt{2}]$, but which contains $\sqrt[3]{2}$ as well as the rational numbers?

Exercise 30. Find an element a of $\mathbb{F}_7$ so that every non-zero element of $\mathbb{F}_7$ is a power of a.

Exercise 31. Show that the set of all polynomials with real coefficients does not form a field (using the usual operations).

Exercise 32. (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 series. Show that $\mathbb{C}((t))$ forms a field. (You should ignore the question of whether the series are convergent.)

Exercise 33. Show that the field $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$ is not algebraically closed.

Exercise 34. Let K be a field having only finitely many elements. Show that K is not algebraically closed.

7 RSA cryptography

Cryptography is the study of keeping messages secret by coding the messages so only the intended recipient can read them. In a **public key cryptosystem**, the method of encryption can be made public, but decryption is not possible (in a reasonable amount of time) except by the intended recipient. In this section we outline the use of modular arithmetic in the **RSA cryptosystem**. It was developed in 1977 by Rivest, Shamir and Adleman and 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 integers (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.) By contrast, large primes can be found efficiently using known primality tests.

7.1 Fermat's little theorem and Euler's theorem

We will need the following results.

Theorem 1.35: Fermat's Little Theorem

Let $p \in \mathbb{N}$ a prime number. If $a \in \mathbb{Z}$ is any integer which is not a multiple of p, then

$$a^{p-1} \equiv 1 \pmod{p}$$

Proof. Let $a \in \mathbb{Z}$ with $p \nmid a$. We will show that it must then be the case that p divides $a^{p-1} - 1$.
Since $p \nmid a$, we have that $[a]_p$ has a multiplicative inverse in $\mathbb{Z}/p\mathbb{Z}$ (Corollay 1.26). Let $[b]_p \in \mathbb{Z}/p\mathbb{Z}$ be such that $[b]_p[a]_p = [1]_p$. Consider the map $\mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ given by $[x]_p \mapsto [a]_p[x]_p$. This map is injective since

$$[a]_p[x]_p = [a]_p[y]_p \implies [b]_p[a]_p[x]_p = [b]_p[a]_p[y]_p \implies [x]_p = [y]_p$$

An injective map from a finite set to itself is necessarily also surjective. Therefore

$$\{[0]_p, [1]_p, \ldots, [p-1]_p\} = \{[0]_p, [a]_p, \ldots, [(p-1)a]_p\}$$

Multiplying together the non-zero elements, we obtain

$$\begin{aligned}
[a]_p \times [2a]_p \times \cdots \times [(p-1)a]_p &= [1]_p \times [2]_p \times \cdots \times [(p-1)]_p \\
[a^{p-1}(p-1)!]_p &= [(p-1)!]_p \\
\implies \quad & p \mid (a^{p-1}-1)(p-1)! \\
\implies \quad & p \mid (a^{p-1}-1) \text{ or } p \mid (p-1)! && (p \text{ is prime}) \\
\implies \quad & p \mid (a^{p-1}-1) && (\text{since } p \not| (p-1)!)
\end{aligned}$$

□

Theorem 1.36: Euler

Let $p, q \in \mathbb{N}$ be primes with $p \neq q$. Suppose that $N \in \mathbb{N}$ satisfies $N \equiv 1 \pmod{(p-1)(q-1)}$. Then

$$\forall a \in \mathbb{Z}, \qquad a^N \equiv a \pmod{pq}$$

Proof. Let $k \in \mathbb{Z}$ be such that $N = 1 + k(p-1)(q-1)$. By Fermat's Little Theorem, if $p \not| a$ 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 therefore

$$a^N \equiv a \pmod{p}$$

This equation also holds if $p \mid a$, since both sides are then 0. Therefore

$$\forall a \in \mathbb{Z} \qquad a^N \equiv a \pmod{p}$$

Similarly,

$$\forall a \in \mathbb{Z} \qquad a^N \equiv a \pmod{q}$$

Since $p \mid (a^N - a)$ and $q \mid (a^n - a)$ and $\gcd(p, q) = 1$, it follows that $pq \mid a^N - a$ (Exercise 11). Hence

$$\forall a \in \mathbb{Z} \qquad a^N \equiv a \pmod{pq}$$

□

7.2 RSA cryptosystem

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* of a message unit $X \in \mathbb{Z}/m\mathbb{Z}$ is given by

$$X \mapsto X^e \pmod{m}$$

- *decryption* is given by

$$Y \mapsto Y^d \pmod{m}$$

The original message is recovered since (by Euler's theorem)

$$(X^e)^d \equiv X^{ed} \equiv X \pmod{m}$$

Why is the RSA cryptosystem secure? To decrypt a message efficiently requires finding $n = (p-1)(q-1)$ 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).

Example 1.37. A very small example:

- Choose $p = 7, q = 13$
- Then $m = 7 \times 13 = 91$ and $n = 6 \times 12 = 72$
- Choose $e = 5$. (This is OK since $\gcd(5, 72) = 1$)
- Then $d = 29$ since $5 \times 29 = 145 \equiv 1 \pmod{72}$
- 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:

$$23 \quad 85$$

7.3 Exercises

Exercise 35. 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'.

Exercise 36. 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.

Ada 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 $n = (p-1)(q-1)$. 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) Xav wants to send Ada the message: 'hi there'. What is the encrypted sequence Xav should send?

(c) Ada 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.

Chapter 2

Linear Algebra I

Let V be a finite dimensional K-vector space and $f : V \to V$ a linear transformation. We would like to find a basis $\mathcal{B}$ such that $[f]_{\mathcal{B}}$ is an 'simple' as possible. We know that not all linear transformations are diagonalisable. We will show that the matrix can be chosen to be in a slightly generalised form known as 'Jordan normal form'. Along the way we will state and prove the Cayley-Hamilton Theorem.

1 Revision on vector spaces

We list here some important definitions and results that will be used. More are given (with some overlap with the present chapter) in an appendix.

1.1 Bases and dimension

Theorem 2.1

Let V be a vector space.

1. Every spanning set for V contains a basis.
2. Every linearly independent subset of V can be extended to a basis.
3. Any two bases of V have the same cardinality.

Lemma 2.2

Let V be a vector space and $U, W \leqslant V$ two subspaces of V. Suppose that $U + V$ is finite dimensional. Then

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

1.2 Matrix representation of a linear transformation

Let V and W be two finite dimensional K-vector spaces. Fix bases $\mathcal{B} = \{v_1, \dots, v_m\}$ and $\mathcal{C} = \{w_1, \dots, w_n\}$ for V and W respectively.

Lemma 2.3

Let $f : V \to W$ be a linear transformation. There exists a unique matrix $[f]_{\mathcal{C},\mathcal{B}} \in M_{n\times m}(K)$ with the property that

$$\forall v \in V \quad [f(v)]_{\mathcal{C}} = [f]_{\mathcal{C},\mathcal{B}} \times [v]_{\mathcal{B}}$$

Definition 2.4. The matrix given by the above lemma is called the **matrix representation** of f with respect to $\mathcal{B}$ and $\mathcal{C}$

Remark (on notation). In the case in which $V = W$ and $\mathcal{B} = \mathcal{C}$ we sometimes write $[f]_{\mathcal{B}}$ in place of $[f]_{\mathcal{B},\mathcal{B}}$.

Lemma 2.5

The entries a_{ij} of $[f]_{\mathcal{C},\mathcal{B}}$ are given by the equation $f(v_j) = \sum_{i=1}^{n} a_{ij} w_i$

Remark. The lemma says that the j-th column of $[f]_{\mathcal{C},\mathcal{B}}$ is exactly the coordinate matrix $[f(v_j)]_{\mathcal{C}}$.

Lemma 2.6

The matrix $[f]_{\mathcal{C},\mathcal{B}}$ is invertible if and only if f is a bijection.

$$\begin{tikzcd}
{[v]_{\mathcal{B}}} \arrow[r, "{[f]_{\mathcal{C},\mathcal{B}}\times}"] \arrow[d, "{[\mathrm{Id}_V]_{\mathcal{B}',\mathcal{B}}\times}"'] & {[f(v)]_{\mathcal{C}}} \arrow[d, "{[\mathrm{Id}_W]_{\mathcal{C}',\mathcal{C}}\times}"] \\
{[v]_{\mathcal{B}'}} \arrow[r, "{[f]_{\mathcal{C}',\mathcal{B}'}\times}"'] & {[f(v)]_{\mathcal{C}'}}
\end{tikzcd}$$

Notice that for all $v \in V$ we have

$$[\mathrm{Id}_W]_{\mathcal{C}',\mathcal{C}}[f]_{\mathcal{C},\mathcal{B}}[v]_{\mathcal{B}} = [f(v)]_{\mathcal{C}'} = [f]_{\mathcal{C}',\mathcal{B}'}[\mathrm{Id}_V]_{\mathcal{B}',\mathcal{B}}[v]_{\mathcal{B}}$$

$$\begin{tikzcd}
{[v]_{\mathcal{B}}} \arrow[r, "{[f]_{\mathcal{B}}\times}"] \arrow[d, "P\times"'] & {[f(v)]_{\mathcal{B}}} \arrow[d, "P\times"] \\
{[v]_{\mathcal{B}'}} \arrow[r, "{[f]_{\mathcal{B}'}\times}"'] & {[f(v)]_{\mathcal{B}'}}
\end{tikzcd}$$

Suppose we have a linear transformation $f : V \to V$ and two bases $\mathcal{B}$ and $\mathcal{B}'$ for V. Letting $P = [\mathrm{Id}_V]_{\mathcal{B}',\mathcal{B}}$ we have

$$P[f]_{\mathcal{B}}[v]_{\mathcal{B}} = [f]_{\mathcal{B}'}P[v]_{\mathcal{B}}$$

Since this holds for all $v \in V$ we have that

$$P[f]_{\mathcal{B}} = [f]_{\mathcal{B}'}P$$

Definition 2.7. Let $A, B \in M_n(K)$. We say that A and B are **similar** if there exists an invertible matrix $P \in \mathrm{GL}_n(K)$ such that $B = P^{-1}AP$. It is denoted $A \sim B$.

From the above observations we see that, if $\mathcal{B}$ and $\mathcal{B}'$ are bases of V, then $[f]_{\mathcal{B}}$ and $[f]_{\mathcal{B}'}$ are similar.

The next lemma says that, if we fix a basis for V, there is a correspondence between linear transformations and matrices.

Lemma 2.8

Let V be an n-dimensional K-vector space and $\mathcal{B}$ a basis for V. The map $\mathrm{End}_K(V) \to M_n(K)$, $f \mapsto [f]_{\mathcal{B}}$ is an isomorphism of K-vector spaces.

1.3 Exercises

Exercise 37. Show that the relation of similarity is an equivalence relation on $M_n(K)$. That is, show that the relation is reflexive, symmetric and transitive.

Exercise 38. Let V be an n-dimensional K-vector space. Show that every element of $\mathrm{GL}_n(K)$ is a change of basis matrix for V. That is, show that for all $P \in \mathrm{GL}_{(}\,K)$ there exist bases $\mathcal{B}$ and $\mathcal{B}'$ for V such that $P = [\mathrm{Id}_V]_{\mathcal{B}',\mathcal{B}}$.

Exercise 39. Let's note that if we allow different bases for domain and codomain, then f does have a diagonal matrix representation. Given a linear transformation $f : V \to V$, show that there exist bases $\mathcal{B}$ and $\mathcal{B}'$ for V such that $[f]_{\mathcal{B}',\mathcal{B}}$ is diagonal and all entries are either 0 or 1. (Hint: start with a basis for the kernel of f.)

2 Invariant decompositions

Our approach to finding a 'simple' matrix representation of a linear transformation $f : V \to V$ will be to decompose V into smaller pieces that are each preserved by f.

Definition 2.9. Let V be a K-vector space (not necessarily finite dimensional). Let $f : V \to V$ be a linear transformation. An **eigenvalue** of f is an element $\lambda \in K$ such that there exists $v \in V \setminus \{0\}$ with $f(v) = \lambda v$. Given an eigenvalue λ, the corresponding **eigenspace** is given by $V_\lambda = \{v \in V \mid f(v) = \lambda v\}$. The non-zero elements of V_λ are called **eigenvectors** of f.

Exercise 40. Show that V_λ is a subspace of V and has dimension at least 1.

Similarly, we define eigenvalues (etc) for a square matrix $A \in M_n(K)$.

Definition 2.10. Let $A \in M_n(K)$. Consider the linear transformation $f : M_{n,1}(K) \to M_{n,1}(K)$ defined by $f(v) = Av$. We say that $\lambda \in K$ is an **eigenvalue** of A is if is an eigenvalue of f. Smilarly, the eigenspaces and eigenvectors of A are defined to be those of f.

Exercise 41. Let $A, B \in M_n(K)$ be similar matrices. Show that $\lambda \in K$ is an eigenvalue of A if and only of λ is an eigenvalue of B.

Example 2.11. Some linear transformations with their eigenvalues:

1. $f : \mathbb{R}^3 \to \mathbb{R}^3$ given by orthogonal reflection across the plane Π given by $ax + by + cz = 0$. There are two eigenvalues: -1 and 1. The eigenspaces are $V_1 = \Pi$ and $V_{-1} = \mathrm{span}\{(a, b, c)\}$
2. $f : \mathbb{R}^3 \to \mathbb{R}^3$ given by rotation through an angle $\theta \in (0, \pi)$ about the line $\ell = \mathrm{span}\{(a, b, c)\}$. There is only one eigenvalue: 1. The eigenspace is $V_1 = \ell$
3. $A = \begin{pmatrix} 2 & 1 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 7 \end{pmatrix}$ has eigenvalues 2,3 and 7.
4. $V = \{\varphi \in \mathbb{R}^\mathbb{R} \mid \varphi \text{ is smooth }\}$, $f : V \to V$ given by $f(\varphi) = \frac{d\varphi}{dx}$. Every $\lambda \in \mathbb{R}$ is an eigenvalue of f! Given any $\lambda \in \mathbb{R}$, the function given by $\varphi(x) = e^{\lambda x}$ is an element of V that is an eigenvector with eigenvalue λ.

Exercise 42. Let $f : V \to V$ be a linear transformation and λ an eigenvalue of f. Show that if $v \in V_\lambda$, then $f(v) \in V_\lambda$.

Definition 2.12. Let $f : V \to V$ be a linear transformation. A subspace $W \leqslant V$ is called an **invariant subspace** if $\forall w \in W$, $f(w) \in W$. Given an invariant subspace W, the **restriction** of f to W is the linear transformation

$$f|_W : W \to W \quad \text{given by} \quad f|_W(w) = f(w)$$

Exercise 43. Let V be a finite dimensional K-vector space and $f : V \to V$ a linear transformation. Suppose that $W \leqslant V$ is an f-invariant subspace. Fix a basis $\{w_1, \ldots, w_m\}$ for W and extend to a basis $\mathcal{B} = \{w_1, \ldots, w_m, v_1, \ldots, v_n\}$ for V. Show that

$$[f]_\mathcal{B} = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$

for some $A \in M_m(K)$, $B \in M_{m,n}(K)$, $D \in M_n(K)$.

Example 2.13. Consider the linear transformation $f : \mathbb{Q}^3 \to \mathbb{Q}^3$ determined by

$$f(1, 0, 0) = (1, -1, 0), \quad f(0, 1, 0) = (2, 1, 0), \quad f(0, 0, 1) = (1, 0, 1)$$

Then $W = \mathrm{span}\{(1, 0, 0), (0, 1, 0)\}$ is invariant and $[f]_\mathcal{S} = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Definition 2.14. Let V be a vector space and $W \leqslant V$ a subspace. A **complement** of W is a subspace $U \leqslant V$ such that: $U \cap W = \{0\}$ and $U + W = V$. We write $V = U \oplus W$ and say that V is a **direct sum** of U and W.

Remark. Clearly, if U is a complement of W, then W is a complement of U.

Lemma 2.15

Let V be a finite dimensional vector space and U, W two subspaces of V. Then the following are equivalent:

1. $V = U \oplus W$
2. Given any bases $\mathcal{B}$ and $\mathcal{C}$ for U and W (respectively), $\mathcal{B} \cup \mathcal{C}$ is a basis for V and $\mathcal{B} \cap \mathcal{C} = \emptyset$
3. There exist bases $\mathcal{B}$ and $\mathcal{C}$ for U and W (respectively) such that $\mathcal{B} \cup \mathcal{C}$ is a basis for V and $\mathcal{B} \cap \mathcal{C} = \emptyset$
4. $U \cap W = \{0\}$ and $\dim V = \dim U + \dim W$
5. $U + W = V$ and $\dim V = \dim U + \dim W$

Proof. $(1 \Rightarrow 2)$ Assume that the first holds. That is, assume that we have $U \cap W = \{0\}$ and $U + W = V$. Fix bases $\mathcal{B} = \{u_1, \dots, u_m\}$ and $\mathcal{C} = \{w_1, \dots, w_n\}$ for U and W. It is clear that $\mathcal{B} \cup \mathcal{C}$ is a spanning set for V since $V = U + W$. That the set $\mathcal{B} \cup \mathcal{C}$ is linearly independent follows from the following

$$\begin{aligned}
& \sum_{i=1}^{m} \alpha_i u_i + \sum_{i=1}^{n} \beta_i w_i = 0 && \text{(where } \alpha_i, \beta_i \in K\text{)} \\
\implies & \sum_{i=1}^{m} \alpha_i u_i = \sum_{i=1}^{n} (-\beta_i) w_i \in U \cap W \\
\implies & \sum_{i=1}^{m} \alpha_i u_i = 0 \quad \text{and} \quad \sum_{i=1}^{n} (-\beta_i) w_i = 0 \\
\implies & \alpha_i = 0 \quad \text{and} \quad \beta_i = 0 \quad \text{for all } i
\end{aligned}$$

Therefore $\mathcal{B} \cup \mathcal{C}$ is a basis for V. Also

$$v \in \mathcal{B} \cap \mathcal{C} \implies v \in U \cap W \implies v = 0$$

Since any set containing the zero vector is linearly dependent, we conclude that $\mathcal{B} \cap \mathcal{C} = \emptyset$.

$(2 \Rightarrow 3)$ Is immediate.

$(3 \Rightarrow 4)$ Assume now that $\mathcal{B}$ and $\mathcal{C}$ are as in 3. Then

$$\begin{aligned}
\dim V &= |\mathcal{B} \cup \mathcal{C}| && (\mathcal{B} \cup \mathcal{C} \text{ is a basis for } V) \\
&= |\mathcal{B}| + |\mathcal{C}| && (\mathcal{B} \cap \mathcal{C} = \emptyset) \\
&= \dim U + \dim W
\end{aligned}$$

Also

$$\begin{aligned}
& \dim(U + W) + \dim(U \cap W) = \dim U + \dim W && \text{(Lemma 2.2)} \\
\implies & \dim V + \dim(U \cap W) = \dim V \\
\implies & \dim(U \cap W) = 0 \\
\implies & U \cap W = \{0\}
\end{aligned}$$

$(4 \Rightarrow 5)$ Assume that $U \cap W = \{0\}$ and $\dim V = \dim U + \dim W$. We have

$$\begin{aligned}
\dim V = \dim U + \dim W & \implies \dim V = \dim(U + W) + \dim(U \cap W) && \text{(using Lemma 2.2)} \\
& \implies \dim V = \dim(U + W) && (U \cap W = \{0\}) \\
& \implies V = U + W && \text{(since } U + W \leqslant V\text{)}
\end{aligned}$$

$(5 \Rightarrow 1)$ Assume that $V = U + W$ and $\dim V = \dim U + \dim W$. We have

$$\begin{aligned}
\dim V = \dim U + \dim W & \implies \dim V = \dim(U + W) + \dim(U \cap W) && \text{(Lemma 2.2)} \\
& \implies \dim V = \dim V + \dim(U + W) && \text{(since } V = U + W\text{)} \\
& \implies \dim(U + V) = 0 \\
& \implies U + V = \{0\}
\end{aligned}$$

□

Exercise 44. Show that the first three conditions in the above lemma remain equivalent even without the hypothesis that V be finite dimensional.

The following observations will be an important tool in finding a simple matrix representation.

Lemma 2.16

Let V be a finite dimensional vector space and $f : V \to V$ a linear transformation. Suppose that U and W are f-invariant subspaces of V such that $V = U \oplus W$. Let $\mathcal{B}$ and $\mathcal{C}$ be bases for U and W respectively. Then

$$[f]_{\mathcal{B}\cup\mathcal{C}} = \begin{bmatrix} [f|_U]_{\mathcal{B}} & 0 \\ 0 & [f|_W]_{\mathcal{C}} \end{bmatrix}$$

Proof. Let $\mathcal{B} = \{u_1, \dots, u_m\}$ and $\mathcal{C} = \{w_1, \dots, w_n\}$. For $j \in \{1, \dots, m\}$ the j-th column of $[f]_{\mathcal{B}\cup\mathcal{C}}$ is equal to $[f(u_j)]_{\mathcal{B}\cup\mathcal{C}}$. Since U is f-invariant, we have $f(u_j) = \sum_{i=1}^{m} a_{ij} u_i$ for some $a_{ij} \in K$. Therefore

$$[f|_U(u_j)]_{\mathcal{B}} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \quad \text{and} \quad [f(u_j)]_{\mathcal{B}\cup\mathcal{C}} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Similar considerations apply to the last n columns of the matrix $[f]_{\mathcal{B}\cup\mathcal{C}}$ □

Given a linear transformation f we want to find complementary f-invariant subspaces. To help do this we consider the minimal polynomial of f.

3 Minimal polynomial

The minimal polynomial (to be defined below) is the monic polynomial of lowest degree that is satisfied by a linear transformation. We will see that it divides the characteristic polynomial.

Let V be a finite dimensional K-vector space and let $p(X) = a_0 + a_1 X + \cdots + a_n X^n \in K[X]$ be a polynomial. Given a a linear transformation $f : V \to V$, define $p(f) : V \to V$ by $p(f) = a_0 \operatorname{Id}_V + a_i f + \cdots + a_n f^n$. Similarly, given a square matrix $A \in M_m(K)$ define $p(A) \in M_m(K)$ by $p(A) = a_0 I_m + a_1 A + \cdots + a_n A^n$.

Remark. If $\mathcal{B}$ a basis for V and $p(X) \in K[X]$, we have

$$[p(f)]_{\mathcal{B}} = p([f]_{\mathcal{B}})$$

Now consider the set of polynomials

$$S = \{p(X) \in K[X] \mid p(f) = 0, p(X) \neq 0\} \subseteq K[X]$$

To see that $S \neq \emptyset$. Let $n = \dim V$. We know from Lemma 2.8 that $\dim(\operatorname{End}_K(V)) = n^2$. Therefore the set $\{\operatorname{Id}_V, f, f^2, \dots, f^{n^2}\} \subseteq \operatorname{End}_K(V)$ is linearly dependent, that is, there exist $a_i \in K$ (not all equal to zero) such that $\sum_{i=0}^{n^2} a_i f^i = 0$. It follows that $\sum_{i=0}^{n^2} a_i X^i \in S$.

Since the set $\{\deg(p(X)) \mid p(X) \in S\} \subseteq \mathbb{N}$ is non-empty, there is (by the Well Ordering property of $\mathbb{N}$) a polynomial $m(X) \in S$ such that $\deg(m(X)) \leqslant \deg(p(X))$ for all $p(X) \in S$. Multiplying by an element of K if necessary, we can assume that $m(X)$ is monic.

Definition 2.17. Let $f : V \to V$. The **minimal polynomial** of f is the element of S that is of lowest degree and is monic.

Remark. If is clear from the above discussion that $m(X)$ is uniquely determined by f and that $\deg(m(X)) \geqslant 1$.

Exercise 45. 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.

To show that $m(X)$ divides all elements of S, we use the polynomial versions of Theorems 1.1 and 1.6.

Theorem 2.18

Let K b a field and $a(X), d(X) \in K[X]$ with $d(X) \neq 0$. There there exist $q(X), r(X) \in K[X]$ such that

$$a(X) = q(x)d(X) + r(X) \quad \text{and either } r = 0 \text{ or } \deg(r(X)) < \deg(d(X))$$

Moreover, $q(X)$ and $r(X)$ are uniquely determined by $a(X)$ and $d(X)$.

Proof. The proof follows exactly the same reasoning as in the version for $\mathbb{Z}$ (Theorem 1.1). □

Exercise 46. Modify the proof of Theorem 1.1 to produce a proof of Theorem 2.18.

Exercise 47. Use Theorem 2.18 to show that $\forall p(X) \in K[X]\ \forall k \in K, \quad p(k) = 0 \implies (X - k) \mid p(X)$

Definition 2.19. Let $a(X), b(X) \in K[X]$. A **greatest common divisor** (gcd) of $a(X)$ and $b(X)$ is an element $d(X) \in K[X]$ such that

1) $(d(X) \mid a(X)) \wedge (d(X) \mid b(X))$

2) $\forall c(X) \in K[X], \quad (c(X) \mid a(X)) \wedge (c(X) \mid b(X)) \implies c(X) \mid d(X)$

We say that $a(X)$ and $b(X)$ are **relatively prime** if 1 is a gcd of $a(X)$ and $b(X)$.

Remark. The gcd of two polynomials is not unique, but any two gcd's differ only up to multiplication by an element of K.

Example 2.20. 1. $X^2 + X + 1$ is a gcd of $X^3 - X^2 - X - 2$ and $X^4 + 2X^3 + 2X^2 + X$ in $\mathbb{R}[X]$

2. $X^2 + X + 1$ is *not* a gcd of $X^3 - X^2 - X - 2$ and $X^4 + 2X^3 + 2X^2 + X$ in $\mathbb{F}_2[X]$. A gcd is $X^3 + X^2 + X$.

Exercise 48. Let $a, b \in K$ with $a \neq b$.

a) Show that $(X - a)$ and $(X - b)$ are relatively prime.

b) Show that if $d(X)$ is monic and divides $(X - a)^m$, then $d(X) = (X - a)^k$ for some $0 \leqslant k \leqslant n$.

c) Let $m, n \in \mathbb{N}$. Show that $(X - a)^m$ and $(X - b)^n$ are relatively prime.

Theorem 2.21

Let $a(X), b(X) \in K[X]$ be two polynomials at least one of which in non-zero. Then there exists a gcd $d(X)$ of $a(X)$ and $b(X)$. Moreover, for any gcd $d(X)$ there exist $\alpha(X), \beta(X) \in K[X]$ such that

$$d(X) = \alpha(X)a(X) + \beta(X)b(X)$$

Proof. The proof follows exactly the same reasoning as in the version for $\mathbb{Z}$ (Theorem 1.6). □

Exercise 49. Modify the proof of Theorem 1.6 to produce a proof of Theorem 2.21.

We can now show that the minimal polynomial divides any polynomial that has f as a root.

Proposition 2.22

Let $m(X) \in K[X]$ be the minimal polynomial of a linear transformation f. Then

$$\forall\, p(X) \in K[X], \quad p(f) = 0 \implies m(X) \mid p(X)$$

Proof. Let $p(X) \in K[X]$ be such that $p(f) = 0$. By Theorem 2.18, there exist $q(X), r(X) \in K[X]$ such that $p(X) =$

$q(X)m(X) + r(X)$ and either $r(X) = 0$ or $\deg(r(X)) < \deg(m(X))$.

$$\begin{aligned} & r(X) = p(X) - m(X)q(X) \\ \implies & r(f) = p(f) - m(f) \circ q(f) \\ & \quad = 0 && \text{(since } p(f) = m(f) = 0\text{)} \end{aligned}$$

We must therefore have that $r(X) = 0$, since otherwise $r(X) \in S$ and has lower degree than $m(X)$. Therefore $p(X) = q(X)m(X)$. □

Examples 2.23.

1. Let V be an n-dimensional K-vector space, let $\lambda \in K$, and consider the linear transformation $f = \lambda \operatorname{Id}_V$. The minimal polynomial is $m(X) = (X - \lambda)$. The characteristic polynomial is $(X - \lambda)^n$.
2. Consider a reflection (across a line through the origin) $f : \mathbb{R}^2 \to \mathbb{R}^2$. Since $f^2 = \operatorname{Id}_V$, we know that $m(X) \mid (X^2 - 1)$. Therefore $m(X)$ is one of $X + 1$, $X - 1$ or $X^2 - 1$ (as these are the only monic divisors of $X^2 - 1$). But if $m(X) = X + 1$, then $f = -\operatorname{Id}_V$. Similarly, if $m(X) = X - 1$, then $f = \operatorname{Id}_V$. Therefore $m(X) = X^2 - 1$. The characteristic polynomial is also $X^2 - 1$.
3. Fix a basis $\mathcal{B}$ of $\mathbb{R}^4$ and let $f : \mathbb{R}^4 \to \mathbb{R}^4$ be given by $[f]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$. Calculating the characteristic polynomial of the matrix gives $c(X) = (X - 1)(X - 2)^3$. Therefore

$$[c(f)]_{\mathcal{B}} = c([f]_{\mathcal{B}}) = 0 \qquad \text{(by the Cayley-Hamilton theorem)}$$

Therefore $c(f) = 0$ and hence $m(X)$ must be one of the monic divisors of $c(X)$:

$$X - 1, X - 2, (X - 1)(X - 2), (X - 1)(X - 2)^2, (X - 1)(X - 2)^3$$

Since

$$[f]_{\mathcal{B}} - I \neq 0 \qquad [f]_{\mathcal{B}} - 2I \neq 0 \qquad ([f]_{\mathcal{B}} - I)([f]_{\mathcal{B}} - 2I) \neq 0 \qquad ([f]_{\mathcal{B}} - I)([f]_{\mathcal{B}} - 2I)^2 = 0$$

we conclude that $m(X) = (X - 1)(X - 2)^2$.

Since the minimal polynomial divides the characteristic polynomial, any root of the minimal polynomial is also a root of the characteristic polynomial and therefore an eigenvalue. The next result is that, conversely, all eigenvalues are roots of the minimal polynomial.

Lemma 2.24

Let V be a finite dimensional K-vector space, $f : V \to V$ a linear transformation, and $m(X) \in K[X]$ the minimal polynomial of f. Then

$$\forall \lambda \in K, \quad m(\lambda) = 0 \iff \lambda \text{ is an eigenvalue of } f$$

Proof. Suppose that $\lambda \in K$ is an eigenvalue of f. Let $v \in V \setminus \{0\}$ be such that $f(v) = \lambda v$. For any $p(X) \in K[X]$ we have $p(f)(v) = p(\lambda)v$ and therefore

$$\begin{aligned} & m(f)(v) = m(\lambda)v \\ \implies & 0(v) = m(\lambda)v && (m(f) = 0 \in \operatorname{End}_K(V)) \\ \implies & 0_V = m(\lambda)v && (0_{\operatorname{End}_K(V)}(v) = 0_V) \\ \implies & m(\lambda) = 0_K && (v \neq 0_V) \end{aligned}$$

Now for the converse. Suppose that $m(\lambda) = 0$. We will show, without appealing to the Cayley-Hamilton theorem, that λ is an eigenvalue of f. From Exercise 47 we know that $m(\lambda) = 0$ implies that $(X - \lambda) \mid m(X)$. Let $t \in \mathbb{N}$ be given by

$$t = \max\{n \in \mathbb{N} \mid (X - \lambda)^n \mid m(X)\}$$

Then $m(X) = (X - \lambda)^t q(X)$ with $\deg(q(X)) < \deg(m(X))$ and $q(\lambda) \neq 0$. Since $\deg(q(X)) < \deg(m(X))$ we have $q(f) \neq 0_{\operatorname{End}_K(V)}$. Let $v \in V$ be such that $q(f)(v) \neq 0_V$. Letting $w = q(f)(v)$ we have

$$(f - \lambda \operatorname{Id}_V)^t(w) = (f - \lambda \operatorname{Id}_V)^t q(f)(v) = m(f)(v) = 0_{\operatorname{End}_K(V)}(v) = 0_V$$

Now define $s \in \mathbb{Z}$, with $0 \leqslant s < t$ to be maximal with the property that $(f - \lambda \operatorname{Id}_V)^s(w) \neq 0$. Then letting $u = (f - \lambda \operatorname{Id}_V)^s(w)$ we have $u \neq 0$ and

$$(f - \lambda \operatorname{Id}_V)u = (f - \lambda \operatorname{Id}_V)^{s+1}(w) = 0$$

Therefore u is an eigenvector with eigenvalue λ.

□

Exercise 50. Find the minimal polynomials of the matrices:

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

Exercise 51. 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?

Exercise 52. 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.

Exercise 53. 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?

Exercise 54. Suppose that A is an $n \times n$ upper triangular matrix with zeros on the diagonal. Prove that $A^n = 0$.

Lemma 2.25

Let V be a (not necessarily finite dimensional) K-vector space and let $p(X) \in K[X]$. Then $\ker(p(f))$ is an f-invariant subspace of V.

Proof. Let $p(X) = \sum_{i=0}^n a_i X^i$. Let $v \in \ker(p(f))$. Then

$$\begin{aligned} p(f)(f(v)) &= (\sum_{i=0}^n a_i f^i)(f(v)) = \sum_{i=0}^n a_i f^{i+1}(v) = f(\sum_{i=0}^n a_i f^i(v)) \\ &= f(p(f)(v)) = f(0) = 0 \end{aligned}$$

Therefore $f(v) \in \ker(p(f))$. □

Remark. Notice that the point in the above proof is the f and $p(f)$ commute.

The following will be a crucial tool in developing Jordan normal form.

Lemma 2.26

Let V be a finite dimensional K-vector space. Let $f : V \to V$ be a linear transformation and $m(X) \in K[X]$ its minimal polynomial. Suppose that $m(X) = p(X)q(X)$ where $p(X), q(X) \in K[X]$ are monic polynomials that are relatively prime. Then

1. $V = \ker(p(f)) \oplus \ker(q(f))$
2. the minimal polynomial of $f|_{\ker(p(f))}$ is $p(X)$

3. the minimal polynomial of $f|_{\ker(q(f))}$ is $q(X)$

Proof.

$$\begin{aligned}
&\exists\, a(X), b(X) \in K[X], \quad a(X)p(X) + b(X)q(X) = 1 && \text{Theorem 2.21}\\
&a(f)p(f) + b(f)q(f) = \mathrm{Id}_V\\
&\forall v \in V, \quad v = a(f)p(f)(v) + b(f)q(f)(v) && (*)
\end{aligned}$$

It follows that

$$V = \ker(q(f)) + \ker(p(f))$$

and

$$v \in \ker(q(f)) \cap \ker(p(f)) \implies v = a(f)(0) + b(f)(0) \qquad \text{(by } *)$$

Therefore $V = \ker(q(f)) \oplus \ker(p(f))$

To see that $p(X)$ is the minimal polynomial of $g = f|_{\ker(p(f))}$ note first that

$$p(g) = p(f|_{\ker(p(f))}) = p(f)|_{\ker(p(f))} = 0$$

Let $p'(X)$ be any non-zero polynomial such that $p'(g) = 0$. Then, for all $v \in V$ we have

$$\begin{aligned}
p'(f)q(f)(v) &= p'(f)q(f)(u + w) && \text{(for some } u \in \ker(q(f)) \text{ and } w \in \ker(p(f)))\\
&= p'(f)q(f)(u) + p'(f)q(f)(w)\\
&= p'(f)q(f)(w) && (u \in \ker(q(f)))\\
&= p'(f)(z) && \text{(letting } z = q(f)(w))\\
&= p'(g)(z) && \text{(since } z \in \ker(p(f)))\\
&= 0(z) = 0
\end{aligned}$$

Therefore $(p(X)q(X)) \mid (p'(X)q(X))$ since $p(X)q(X)$ is the minimal polynomial of f. This implies that $\deg(p(X)) \leqslant \deg(p'(X))$.

The same argument can be used to show that $q(X)$ is the minimal polynomial of $f|_{\ker(q(f))}$. □

Proposition 2.27

Let V be a finite dimensional K-vector space and $f : V \to V$ a linear transformation. Suppose that the minimal polynomial of f can be factorised as a product of pairwise relatively prime monic polynomials $q_i \in K[X]$:

$$m(X) = q_1(X)q_2(X)\dots q_N(X)$$

Let $\mathcal{B}_i$ be a basis for $\ker(q_i(f))$ and $A_i = [f|_{\ker(q_i(f))}]_{\mathcal{B}_i}$

Then $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_N$ is a basis for V and

$$[f]_{\mathcal{B}} = \begin{bmatrix} A_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & A_N \end{bmatrix}$$

Proof. We consider the case $N = 2$. Extension to the general case is then an easy induction argument.

Suppose $m(X) = q_1(X)q_2(X)$. Let $K_i = \ker(q_i(f))$. By Lemma 2.25 K_i is f-invariant and by Lemma 2.26 $V = K_1 \oplus K_2$. Applying Lemma 2.16 we have

$$[f]_{\mathcal{B}_1 \cup \mathcal{B}_2} = \begin{bmatrix} A_1 & 0\\ 0 & A_2 \end{bmatrix}$$

□

Example 2.28. As an illustration of the above proposition, consider $f : \mathbb{R}^3 \to \mathbb{R}^3$ given by $[f]_{\mathcal{S}} = A = \begin{bmatrix} -13 & 4 & 7 \\ -18 & 6 & 9 \\ -14 & 4 & 8 \end{bmatrix}$

A quick calculation gives the characteristic polynomial as $X^2(X-1)$ and so the eigenvalues as 0 and 1. Multiplication gives that $A(A-I) \neq 0$ and $A^2(A-I) = 0$. Therefore the minimal polynomial of f is $m(X) = X^2(X-1)$. Let $q_1(X) = X^2$ and $q_2(X) = X - 1$. Note that they are relatively prime. Using that

$$q_1(A) = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ -2 & 0 & 2 \end{bmatrix} \qquad q_2(A) = \begin{bmatrix} -14 & 4 & 7 \\ -18 & 5 & 9 \\ -14 & 4 & 7 \end{bmatrix}$$

we obtain bases $\mathcal{B}_1 = \{(1,0,1),(0,1,0)\}$ and $\mathcal{B}_2 = \{(1,0,2)\}$ for $\ker(q_1(f))$ and $\ker(q_2(f))$ respectively. Letting $\mathcal{B} = \{(1,0,1),(0,1,0),(1,0,2)\}$, we have

$$[f]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}^{-1} A \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -6 & 4 & 0 \\ -9 & 6 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

We want to analyze the blocks A_i that arise in the Proposition 2.27 from factors of the form $(X-\lambda)^m$.

Lemma 2.29

Let V be an n-dimensional K-vector space and let $f : V \to V$ be a linear transformation. Suppose that the minimal polynomial of f if $(X-\lambda)^m$ for some $\lambda \in K$ and $m \in \mathbb{N}$. Then $m \leqslant n$ and there exists a basis $\mathcal{B}$ of V such that $[f]_{\mathcal{B}}$ is in upper-triangular form with all entries on the diagonal equal to λ. That is,

$$[f]_{\mathcal{B}} = \begin{bmatrix} \lambda & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda \end{bmatrix}$$

Proof. Since λ is a root of the minimal polynomial, it is an eigenvalue (Lemma 2.24). For $0 \leqslant i \leqslant m$ define $W_i = \ker(f - \lambda \operatorname{Id}_V)^i$. Note that

$$\{0\} = W_0 \subseteq W_1 \subseteq W_2 \subseteq \cdots \subseteq W_m = V$$

We now show that $W_{i-1} \neq W_i$. Suppose, for a contradiction, that $W_{i-1} = W_i$ for some i. Then for all $v \in V$, we have

$$\begin{aligned} & (f - \lambda \operatorname{Id}_V)^m(v) = 0 \\ \Longrightarrow\ & (f - \lambda \operatorname{Id}_V)^i \circ (f - \lambda \operatorname{Id}_V)^{m-i}(v) = 0 \\ \Longrightarrow\ & (f - \lambda \operatorname{Id}_V)^{i-1} \circ (f - \lambda \operatorname{Id}_V)^{m-i}(v) = 0 && \text{(since } W_{i-1} = W_i\text{)} \\ \Longrightarrow\ & (f - \lambda \operatorname{Id}_V)^{m-1}(v) = 0 \end{aligned}$$

Since this holds for all $v \in V$ we have that $(f - \lambda \operatorname{Id}_V)^{m-1} = 0$, contradicting the minimality of $m(X)$.

Now choose $\mathcal{B}_i \subset W_i$ such that $\mathcal{B}_i$ is a basis for W_i and $\mathcal{B}_1 \subsetneq \mathcal{B}_2 \subsetneq \cdots \subsetneq \mathcal{B}_m$. Since $|B_{i+1}| \geqslant |\mathcal{B}_i| + 1$ we have that $n = |\mathcal{B}_m| \geqslant m$.

We also have

$$\begin{aligned} v \in W_i &\Longrightarrow (f - \lambda \operatorname{Id}_V)^i(v) = 0 \\ &\Longrightarrow (f - \lambda \operatorname{Id}_V)^{i-1}(f - \lambda \operatorname{Id}_V)(v) = 0 \\ &\Longrightarrow (f - \lambda \operatorname{Id}_V)(v) \in W_{i-1} \\ &\Longrightarrow f(v) - \lambda v = w \quad \text{for some } w \in W_{i-1} \\ &\Longrightarrow f(v) = \lambda v + w \end{aligned}$$

That $[f]_{\mathcal{B}_m}$ has the desired form then follows. □

Exercise 55. Let $f : (\mathbb{F}_5)^3 \to (\mathbb{F}_5)^3$ be given by $[f]_{\mathcal{S}} = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \end{bmatrix}$

(a) Calculate the eigenvalues of f.

(b) Find the minimal polynomial of f.

(c) Find a basis $\mathcal{B}$ of $(\mathbb{F}_5)^3$ such that $[f]_{\mathcal{B}}$ is in upper triangular form.

Proposition 2.30

Let V be an finite dimensional K-vector space and let $f : V \to V$ be a linear transformation. Suppose that the minimal polynomial of f is of the form

$$m(X) = (X-\lambda_1)^{m_1}(X-\lambda_2)^{m_2}\dots(X-\lambda_N)^{m_N}$$

for some $N \in \mathbb{N}$, $m_i \in \mathbb{N}$ and $\lambda_i \in K$ with $\lambda_i \neq \lambda_j$ if $i \neq j$.
Then there exists a basis $\mathcal{B}$ of V such that $[f]_{\mathcal{B}}$ is in upper triangular form.

Proof. Note that for $i \neq j$ the polynomials $(X-\lambda_i)^{m_i}$ and $(X-\lambda_j)^{m_j}$ are relatively prime.
Let $q_i(X) = (X-\lambda_i)^{m_i} \in K[X]$, $V_i = \ker(q_i(f)) \leqslant V$, and $f_i = f|_{V_i}$. Choose a basis $\mathcal{B}_i$ for V_i and let $A_i = [f_i]_{\mathcal{B}_i}$.
By Lemma 2.26, $V = V_1 \oplus \dots \oplus V_N$ and the minimal polynomial of f_i is $q_i(X)$.
By Lemma 2.29

$$A_i = \begin{bmatrix} \lambda_i & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_i \end{bmatrix}$$

Let $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_N$. By Proposition 2.27, $\mathcal{B}$ is a basis for V and

$$[f]_{\mathcal{B}} = \begin{bmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_N \end{bmatrix} = \begin{bmatrix} \begin{smallmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_1 \end{smallmatrix} & & 0 \\ & \ddots & \\ 0 & & \begin{smallmatrix} \lambda_N & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_N \end{smallmatrix} \end{bmatrix}$$

□

Corollary 2.31

Let V be an finite dimensional K-vector space and let $f : V \to V$ be a linear transformation. If K is algebraically closed, then there exists a basis $\mathcal{B}$ of V such that $[f]_{\mathcal{B}}$ is in upper triangular form.

Proof. Since K is algebraically closed, any polynomial in $K[X]$ can be written as a product of linear terms. In particular, for the minimal polynomial of f we have

$$m(X) = (X-\lambda_1)^{m_1}(X-\lambda_2)^{m_2}\dots(X-\lambda_N)^{m_N}$$

for some $N \in \mathbb{N}$, $m_i \in \mathbb{N}$ and $\lambda_i \in K$ with $\lambda_i \neq \lambda_j$ if $i \neq j$. Apply the preceding result. □

4 The Cayley-Hamilton theorem

We first recall here the definition of the characteristic polynomial.

Definition 2.32. Let K be a field and $A \in M_n(K)$. The **characteristic polynomial** of A is the polynomial $c_A(X) \in K[X]$ given by

$$c_A(X) = \det(XI_n - A)$$

Remark. The polynomial $c_A(X)$ is always monic and of degree n. The constant term of $c(X)$ is equal to $(-1)^n \det(A)$.

Exercise 56. Show that if $A, B \in M_n(K)$ are similar, then $c_A(X) = c_B(X)$.

Definition 2.33. Let V be an finite dimensional K-vector space and let $f : V \to V$ be a linear transformation. The **characteristic polynomial** of f is denoted $c_f(X)$ and given by $c_f(X) = c_A(X)$ for some matrix representation $A = [f]_{\mathcal{B}}$ of f.

Remark. The above exercise shows that $c_f(X)$ does not depend on the choice of A.

Theorem 2.34: Cayley-Hamilton Theorem

Let V be an finite dimensional K-vector space and let $f : V \to V$ be a linear transformation. Let $c_f(X) \in K[X]$ be the characteristic polynomial of f. Then $c_f(f) = 0$.
(That is, 'f satisfies its own characteristic equation'.)

Proof. Assume first that K is algebraically closed. Then the minimal polynomial has the form

$$m(X) = (X - \lambda_1)^{m_1}(X - \lambda_2)^{m_2} \dots (X - \lambda_N)^{m_N}$$

for some $N \in \mathbb{N}$, $m_i \in \mathbb{N}$ and $\lambda_i \in K$ with $\lambda_i \neq \lambda_j$ if $i \neq j$. As in the proof of Proposition 2.30 define $V_i = \ker(f - \lambda_i)^{m_i}$. By Lemma 2.25, V_i is f-invariant. Let $f_i = f|_{V_i}$. By Lemma 2.26, $V = V_1 \oplus \cdots \oplus V_N$ andt he minimal polynomial of f_i is $(X - \lambda_i)^{m_i}$. Let $n_i = \dim(V_i)$. By Lemma 2.29, $m_i \leqslant n_i$ and there exists a basis $\mathcal{B}_i$ for V_i such that

$$[f_i]_{\mathcal{B}_i} = \begin{bmatrix} \lambda_i & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_i \end{bmatrix} \in M_{n_i}(K)$$

By Lemma 2.16 we have

$$[f]_{\mathcal{B}} = \begin{bmatrix} [f_i]_{\mathcal{B}_i} & & 0 \\ & \ddots & \\ 0 & & [f_i]_{\mathcal{B}_i} \end{bmatrix} = \begin{bmatrix} \begin{smallmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_1 \end{smallmatrix} & & 0 \\ & \ddots & \\ 0 & & \begin{smallmatrix} \lambda_N & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_N \end{smallmatrix} \end{bmatrix}$$

Letting $A = [f]_{\mathcal{B}}$, we have We have

$$\begin{aligned} c_f(X) = c_A(X) &= \det(XI_n - A) \\ &= \Pi_{i=1}^{N} \det(XI_{n_i} - A_i) \\ &= \Pi_{i=1}^{N} (X - \lambda_i)^{n_i} \end{aligned}$$

Since $m_i \leqslant n_i$, we have

$$m(X) = (X - \lambda_1)^{m_1} \dots (X - \lambda_N)^{m_N} \mid (X - \lambda_1)^{n_1} \dots (X - \lambda_N)^{n_N} = c(X)$$

Therefore $c(f) = 0$ and we're done in the case in which K is algebraically closed.
Now consider the case in which K is not algebraically closed. Let L be an algebraically closed field with $K \subseteq L$ (see Theorem 1.34). Let $\mathcal{B}$ be any basis of V and let $A = [f]_{\mathcal{B}}$. Then $A \in M_n(K) \subseteq M_n(L)$. Applying the result already obtained for algebraically closed fields, we have that A satisfies its characteristic equation and therefore

$$[c(f)]_{\mathcal{B}} = c(A) = 0$$

□

Recall that a linear transformation $f : V \to V$ is called **diagonalisable** if there exists a basis $\mathcal{B}$ of V such that $[f]_{\mathcal{B}}$ is a diagonal matrix.

Proposition 2.35

Let K be an algebraically closed field and let V be a finite dimensional K-vector space. A linear transformation $f : V \to V$ is diagonalisable if and only if its minimal polynomial can be written as a product of distinct linear factors (in $K[X]$).

Proof. Denote the minimal polynomial by $m(X) \in K[X]$. Since K is algebraically closed, $m(X)$ can be written as a product of linear factors

$$m(X) = (X - \lambda_1)^{m_1} \dots (X - \lambda_N)^{m_N}$$

with $\lambda_i \neq \lambda_j$ if $i \neq j$. Define V_i, f_i and $\mathcal{B}_i$ as above. Then

$$\begin{aligned} f \text{ is diagonalisable} &\iff \forall i, \quad f_i \text{ is diagonalisable} \\ &\iff \forall i, \quad f_i = \lambda_i \operatorname{Id}_{V_i} && (\lambda_i \text{ is the only eigenvaluie of } f_i) \\ &\iff \forall i, \quad \text{the minimal polynomial of } f_i \text{ is } (X - \lambda_i) \\ &\iff \forall i, \quad m_i = 1 && (\text{the minimal polynomial of } f_i \text{ is } (X - \lambda_i)^{m_i}) \end{aligned}$$

□

5 Jordan normal form

Linear transformations that are not diagonalisable do nonetheless have a nice matrix representation which is block diagonal with each block of a simple form.

Definition 2.36. Let $\lambda \in K$ and $n \in \mathbb{N}$. Define a matrix $J(\lambda, n) \in M_n(K)$ to be the matrix with λ at all entries on the main diagonal, 1 at all entries directly above the main diagonal, and 0 elsewhere.

$$J(\lambda, n) = \begin{bmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{bmatrix}$$

A marix of this form is called a **Jordan block**.

Example 2.37. $J(4,3) = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}$

Exercise 57.

(a) Show that the characteristic polynomial of $J(\lambda, n)$ is $(X - \lambda)^n$

(b) Show that the minimal polynomial of $J(\lambda, n)$ is $(X - \lambda)^n$

(c) Show that the eigenspace has dimension 1.

Definition 2.38. A square matrix is said to be in **Jordan normal form** (JNF) if it is block diagonal and each of the blocks is a Jordan block.

Example 2.39. The matrix shown is in JNF. Note that the characteristic and minimal polynomials are

$$c(X) = (X-2)^3(X-i)^2 \qquad m(X) = (X-2)^2(X-i)^2$$

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

Definition 2.40. A linear transformation $T : V \to V$ is called **nilpotent** if there exists $n \in \mathbb{N}$ such that $T^n = 0$. Similarly, a square matrix $A \in M_m(K)$ is called nilpotent if $A^n = 0$ for some $n \in \mathbb{N}$.

Example 2.41.

1. The linear transformation $D : \mathcal{P}_d(K) \to \mathcal{P}_d(K)$ given by differentiation is nilpotent.
2. The matrix $\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$ is nilpotent.

Exercise 58. Show that if a linear transformation $T : V \to V$ is nilpotent, then it is not injective.

The following technical looking lemma will be used in the proof of the existence of a Jordan normal form matrix representative.

Lemma 2.42

Let K be a field and V a finite dimensional K-vector space. Suppose that $N : V \to V$ is a nilpotent linear transformation. Then there exist $v_1, \ldots, v_k \in V$ and $m_1, \ldots, m_k \in \mathbb{N}$ such that

$$\forall i \in \{1, \ldots, k\}, \quad N^{m_i}(v_i) = 0$$

and

$$\begin{aligned} \{\, & N^{m_1-1}(v_1), \ldots, N^2(v_1), N(v_1), v_1, \\ & N^{m_2-1}(v_2), \ldots, N^2(v_2), N(u_2), v_2, \\ & \qquad\qquad \vdots \\ & N^{m_k-1}(v_k), \ldots, N^2(v_k), N(v_k), v_k \,\} \end{aligned}$$

is a basis for V.

Remark. If we let $\mathcal{B}$ denote the (ordered) basis given in the above lemma, then $[N]_{\mathcal{B}}$ is in Jordan normal form:

$$[N]_{\mathcal{B}} = J(0, m_1) \oplus J(0, m_2) \oplus \cdots \oplus J(0, m_k)$$

Proof of Lemma 2.42. We use (strong) induction on the dimension of V. For the base case ($\dim(V) = 1$), let $u \in \ker(N) \setminus \{0\}$. Then $\{u\}$ is a basis for V and $Nu = 0$.

Now suppose that $\dim(V) \geqslant 2$ and that the lemma holds in all cases with lower dimension. Note that, by the rank-nullity theorem, we have

$$\dim(\mathrm{im}(N)) = \dim(V) - \dim(\ker(N)) < \dim(V)$$

By the induction hypothesis, the result holds for the transformation $N|_{\mathrm{im}(N)} : \mathrm{im}(N) \to \mathrm{im}(N)$. Let $u_1, \ldots, u_k \in \mathrm{im}(N)$ and $m_1, \ldots, m_k \in \mathbb{N}$ be such that

$$\mathcal{A} = \{N^{m_1-1}(u_1), \ldots, N(u_1), u_1, N^{m_2-1}(u_2), \ldots, N(u_2), u_2, \ldots, N^{m_k-1}(u_k), \ldots, N(u_k), u_k\}$$

is a basis for $\mathrm{im}(N)$ and $N^{m_i}(u_i) = 0$ for all i. Choose $v_i \in V$ such that $Nv_i = u_i$ and define

$$\begin{aligned} \mathcal{B} &= \mathcal{A} \cup \{v_1, \ldots, v_k\} \\ &= \{N^{m_1}(v_1), \ldots, N(v_1), v_1, N^{m_2}(v_2), \ldots, N(v_2), v_2, \ldots, N^{m_k}(v_k), \ldots, N(v_k), v_k\} \end{aligned}$$

The set $\mathcal{B}$ is linearly independent since

$$\begin{aligned} \sum_{i=1}^{k} \sum_{j=0}^{m_i} \alpha_{ij} N^j v_i = 0 &\implies \sum_{i=1}^{k} \sum_{j=0}^{m_i} \alpha_{ij} N^{j+1} v_i = 0 && \text{(applying } N \text{ to both sides)} \\ &\implies \sum_{i=1}^{k} \sum_{j=0}^{m_i} \alpha_{ij} N^j u_i = 0 && (u_i = Nv_i) \\ &\implies \sum_{i=1}^{k} \sum_{j=0}^{m_i-1} \alpha_{ij} N^j u_i = 0 && (N^{m_i} u_i = 0) \\ &\implies \alpha_{ij} = 0 \text{ for all } i \in \{1, \ldots, k\} \text{ and } j \in \{0, \ldots, m_i - 1\} && (\mathcal{A} \text{ is linear independent}) \end{aligned}$$

which then also gives

$$\begin{aligned} & \sum_{i=1}^{k} \alpha_{i,m_i} N^{m_i} v_i = 0 \\ \implies & \sum_{i=1}^{k} \alpha_{i,m_i} N^{m_i-1} u_i = 0 \\ \implies & \alpha_{i,m_i} = 0 \text{ for all } i \in \{1, \ldots, k\} && (\mathcal{A} \text{ is linear independent}) \end{aligned}$$

Having shown that $\mathcal{B}$ is linearly independent, we know that it can be extended to a basis of V. Let $\tilde{w}_i \in V$ be such that $\mathcal{B} \cup \{\tilde{w}_1, \dots, \tilde{w}_\ell\}$ is a basis of V. We have (for all i)

$$\begin{aligned}
&N\tilde{w}_i \in \operatorname{im}(V) \\
\implies &N\tilde{w}_i \in \operatorname{span}(\mathcal{A}) \\
\implies &N\tilde{w}_i \in \operatorname{span}(N(\mathcal{B})) && (N(\mathcal{B}) = \mathcal{A} \cup \{0\}) \\
\implies &N\tilde{w}_i = N\hat{w}_i && (\text{for some } \hat{w}_i \in \operatorname{span}(\mathcal{B})) \\
\implies &\tilde{w}_i - \hat{w}_i \in \ker(N)
\end{aligned}$$

Letting $w_i = \tilde{w}_i - \hat{w}_i$ we have that

$$\mathcal{C} = \mathcal{B} \cup \{w_1, \dots, w_\ell\}$$

is a basis for V and is of the desired form. □

Theorem 2.43: Jordan normal form

Let K be an algebraically closed field, V a finite dimensional K-vector space, and $f : V \to V$ a linear transformation. There exists a basis $\mathcal{B}$ of V such that

$$[f]_{\mathcal{B}} = \begin{bmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_N \end{bmatrix}$$

where each J_i is a Jordan block.

Proof. By Lemmas 2.25, 2.26, 2.27 it is enough to consider the case in which f has only one eigenvalue. (In the notation of the previous proof, it is enough to consider the transformation f_i.) So assume that we have a linear transformation $f : V \to V$ with minimal polynomial $(X - \lambda)^m$. Define $N : V \to V$ by $N = f - \lambda \operatorname{Id}$. Then N is nilpotent since $N^m = (f - \lambda \operatorname{Id})^m = 0$. Applying Lemma 2.42, there is a basis $\mathcal{B}$ of V such that $[N]_{\mathcal{B}}$ is in JNF and each Jordan block is of the form $J(0, n_i)$ for some $n_i \in \mathbb{N}$. Since $f = N + \lambda \operatorname{Id}$ we have that $[f]_{\mathcal{B}} = [N]_{\mathcal{B}} + \lambda I$ where I is the identity matrix of size $\dim(V)$. Therefore $[f]_{\mathcal{B}}$ is in JNF and has blocks of the form $J(\lambda, n_i)$. □

Remark. The JNF is unique up to rearrangement of the Jordan blocks. See Exercise 69. Two matrices in JNF are similar if and only if one can be obtained from the other by permuting the Jordan blocks.

The minimal and characteristic equation can be read from the Jordan normal form.

Exercise 59. Let K be an algebraically closed field, V a finite dimensional K-vector space, and $f : V \to V$ a linear transformation. Suppose that $\lambda \in K$ is an eigenvalue of f and let $m, n \in \mathbb{N}$ be maximal with the property that $(X - \lambda)^m$ divides that minimal polynomial and $(X - \lambda)^n$ divides the characteristic polynomial. Show that

(a) $m =$ the size of the largest Jordan block having λ on the diagonal

(b) $n =$ is the sum of the sizes of all Jordan blocks having λ on the diagonal

(c) the dimension of the λ-eigenspace is equal to the number of Jordan blocks having λ on the diagonal.

Example 2.44. Suppose that $A \in M_7(\mathbb{C})$ is similar to the matrix shown (which is in JNF). Then the characteristic and minimal polynomials of A are

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

$$c(X) = (X-2)^5(X-i)^2 \qquad m(X) = (X-2)^2(X-i)^2$$

The eigenspaces have dimensions

$$\dim(V_2) = 3 \qquad \dim(V_i) = 1$$

See Example 2.39 for another example.

Example 2.45. 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

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

Remark. For square matrices of size 2 or 3, the JNF can be determined from the minimal and characteristic polynomials. However, this is not true for larger matrices

Example 2.46. The following two matrices (both in JNF) are not similar.

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

However both have $m(X) = (X-2)^3$, $c(X) = (X-2)^7$ and $\dim(V_2) = 3$. One way to see that they are not similar is to note that $\dim(\ker(A-2I)^2) = 5$ but $\dim(\ker(B-2I)^2) = 6$

5.1 Exercises

Exercise 60. 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.

Exercise 61. 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.

Exercise 62. 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.

Exercise 63. Find the Jordan normal form of the following matrices:

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

Exercise 64. 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
$(X-1)^2(X+1)^2$	$(X-1)^4(X+1)^4$.

Exercise 65. 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) $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

Exercise 66. 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.)

Exercise 67. 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.

Exercise 68. In the language of the previous question, show that $JN = NJ$ and $JD = DJ$. Deduce that $fd = df$ and $fn = nf$.

Exercise 69. (Harder) Show that the Jordan normal form of a complex matrix A is completely determined by the dimensions of the nullspaces of $(A - \lambda I)^i$, $i = 1, 2, 3, \ldots$ for all the eigenvalues λ of A.

Chapter 3

Groups I

1 Definition of a group and some examples

Definition 3.1. A **group** is a non-empty set G together with a binary operation $* : G \times G \to G$ (the image of (g, h) being denoted $g * h$ or simply gh) that satisfies the following properties:

1) $\forall g, h, k \in G, \quad (g * h) * k = g * (h * k)$ (associativity)

2) $\exists e \in G \, \forall g \in G, \quad g * e = g \wedge e * g = g$ (identity element)

3) $\forall g \in G \, \exists h \in G, \quad g * h = e \wedge h * g = e$ (inverses)

Remark. Since a group consists of a set G *and* an operation, a good notation would be $(G, *)$. However, it is common to suppress explicit mention of the operation and refer to the group simply as G.

Exercise 70. (a) Prove that the identity element is unique. That is, show that

$$(\forall g \in G, \quad g * e = g \wedge e * g = g) \wedge (\forall g \in G, \quad g * e' = g \wedge e' * g = g) \implies e = e'$$

(b) Show that the element h in the third axiom is uniquely determined by g. That is, for a given $g \in G$,

$$(g * h = e \wedge h * g = e) \wedge (g * h' = e \wedge h' * g = e) \implies h = h'$$

In light of this uniqueness, the element is denoted g^{-1} and called *the* inverse of g.

(c) Let $g, h \in G$. Show that $(g^{-1})^{-1} = g$ and $(g * h)^{-1} = h^{-1} * g^{-1}$.

Definition 3.2. A group G is called **abelian** if $\forall g, h \in G, \ gh = hg$. A group G is called **finite** if the underlying set G is finite.

Example 3.3. 1. $(\mathbb{Z}, +)$ is an infinite abelian group.

2. $(\mathbb{Z}, \times)$ is not a group.
3. $(\mathbb{Z}/2\mathbb{Z}, +)$ is a finite group. It has two elements.
4. If $(K, +, \times)$ is a field, then $(K, +)$ and $(K \setminus \{0\}, \times)$ are (different) abelian groups.
5. $(M_n(K), \times)$ is not a group.
6. $GL(n, K)$ is a (non-abelian) group.
7. Other matrix groups include:

$O(n)$	the group of all $n \times n$ orthogonal matrices (real matrices A such that $A^T A = I$)
$U(n)$	the group of all $n \times n$ unitary matrices (complex matrices U such that $\overline{U}^t U = I$)
$SL(n, K)$	the group of all $n \times n$ matrices of determinant 1 with entries from the field K
$SO(n)$	the group of all $n \times n$ orthogonal matrices having determinant 1
$SU(n)$	the group of all $n \times n$ unitary matrices having determinant 1

Lemma 3.4

Let G be a group and $g, h, k \in G$. Then

1) $gh = gk \implies h = k$
2) $\exists!\, l \in G, \quad gl = h$
3) The map $L_g : G \to G$, $L_g(x) = gx$ is a bijection.
   The map $R_g : G \to G$, $R_g(x) = xg$ is also a bijection.

Exercise 71. Write out a proof of Lemma 3.4.

Example 3.5. Here are two groups of size 4. Let

$$V = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \right\} \subset GL(2, \mathbb{R})$$

$$C_4 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\} \subset GL(2, \mathbb{R})$$

with the operation in both cases defined to be matrix multiplication.
Notice that in V every element has square equal to the identity. That's not the case in C_4 where, for example, $\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]^2 = \left[\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right] \neq \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$.

1.1 Exercises

Exercise 72. Write down the multiplication tables for V and C_4.

Exercise 73. Show that the set of all rotations of the plane about a fixed centre P, together with the operation of composition, forms a group. What about all of the reflections for which the axis (or mirror) passes through P?

Exercise 74. 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?

Exercise 75. Set $X = \mathbb{R} \setminus \{0, 1\}$. Show the following set of functions $X \to X$, together with the operation of composition, forms a group.

$$\begin{array}{lll} f(x) = \frac{1}{1-x} & g(x) = \frac{x-1}{x} & h(x) = \frac{1}{x} \\ i(x) = x & j(x) = 1 - x & k(x) = \frac{x}{x-1} \end{array}$$

Exercise 76. If G is a group and $(gh)^2 = g^2h^2$ for all $g, h \in G$, prove that G is abelian.

2 The symmetric groups S_n

We investigate the permutations of a fixed set.

Definition 3.6. Let $n \in \mathbb{N}$. A **permutation** of the set $\{1, \ldots, n\}$ is a bijecion $\{1, \ldots, n\} \to \{1, \ldots, n\}$. The group of all permutations of the set $\{1, \ldots, n\}$ is denoted by S_n and called the **symmetric group** (on n letters). The operation is the usual composition of functions.

Remark. It is clear that $|S_n| = n!$.

2.1 Notations for permutations

Let $\sigma \in S_n$. One way of specifying σ is as two rows, with the image $\sigma(i)$ written directly below i. That is, as

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

Example 3.7. There are six permutations of the set $\{1, 2, 3\}$. We can list the six elements of S_3 as follows:

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \tau_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \tau_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \end{pmatrix}, \tau_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

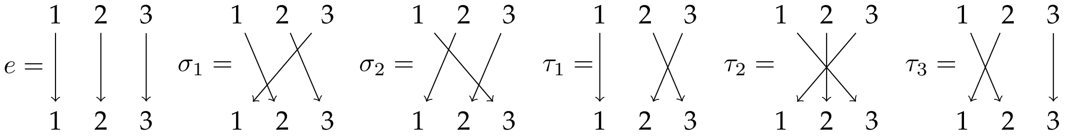
Since the operation is composition of functions we have, for example:

$$\tau_3\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \tau_1$$
$$\sigma_1\tau_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \tau_2$$

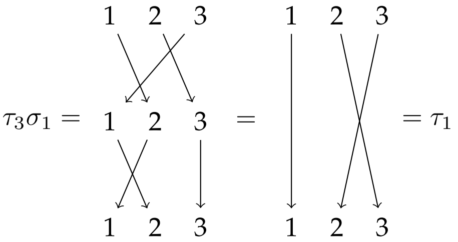
Notice that the group S_3 is not abelian since $\tau_3\sigma_1 \neq \sigma_1\tau_3$. The full multiplication table for S_3 is given on the right.

S_3	e	σ_1	σ_2	τ_1	τ_2	τ_3
e	e	σ_1	σ_2	τ_1	τ_2	τ_3
σ_1	σ_1	σ_2	e	τ_3	τ_1	τ_2
σ_2	σ_2	e	σ_1	τ_2	τ_3	τ_1
τ_1	τ_1	τ_2	τ_3	e	σ_1	σ_2
τ_2	τ_2	τ_3	τ_1	σ_2	e	σ_1
τ_3	τ_3	τ_1	τ_2	σ_1	σ_2	e

Another notation used for elements of S_n is to write the set $\{1, \ldots, n\}$ twice and then join i to $\sigma(i)$ by a directed edge. To illustrate, we list the elements of S_3 in this notation:



To multiply elements in this notation, we simply place one diagram on top of the other and amalgamate the directed edges. For example:



Cycle notation

A third more compact notation is known as **cycle notation**. In this notation each element $\sigma \in S_n$ is represented by a collection tuples ('cycles') in which each element $i \in \{1, \ldots, n\}$ appears exactly once as in followed immediately by $\sigma(i)$ (with the last element of a tuple being 'followed' by the first). Some examples will make this clear. We list the elements of S_3 in cycle notation:

$$e = (1)(2)(3) \quad \sigma_1 = (1, 2, 3) \quad \sigma_2 = (1, 3, 2) \quad \tau_1 = (1)(2, 3) \quad \tau_2 = (1, 3)(2) \quad \tau_3 = (1, 2)(3)$$

It is common to adopt the further conventions that singletons are omitted and commas are dropped (unless the notation would be made ambiguous). With these conventions we have:

$$\sigma_1 = (123) \quad \sigma_2 = (132) \quad \tau_1 = (23) \quad \tau_2 = (13) \quad \tau_3 = (12)$$

The identity element will be denoted as (1) or simply as e.

We will generally use cyclic notation and give here an example of multiplication written in cycle notation.

Example 3.8. Consider $\sigma, \tau \in S_7$ given by $\sigma = (1234)(567)$, $\tau = (143)(267)$. Then

$$\sigma\tau = (1234)(567)(143)(267) = (1)(273)(4)(56) = (273)(56)$$
$$\tau\sigma = (143)(267)(1234)(567) = (162)(3)(4)(57) = (162)(57)$$

Remark. Cycle notation for a permutation is *not* unique, for example $(123) = (231) = (312)$ as they all represent the permutation mapping $1 \mapsto 2$, $2 \mapsto 3$ and $3 \mapsto 1$. Also, $(123)(45) = (45)(123)$.

Exercise 77. Find the product of the following permutations:

(a) $(123)(456) * (134)(25)(6)$ (b) $(12345) * (1234567)$ (c) $(123456) * (123) * (123) * (1)$

3 Subgroups

Definition 3.9. Let G be a group. A **subgroup** of G is a subset $H \subset G$ which, when equipped with the operation from G (restricted to H), itself forms a group. We will use the notation $H \leqslant G$.

Remark. It is clear from the definition that $\{e\} \leqslant G$. It is called the **trivial subgroup**.

Example 3.10. Some examples of groups G and a subgroup $H \leqslant G$.

1. $G = (\mathbb{Z}/4\mathbb{Z}, +)$, $H = \{[0], [2]\}$
2. $G = S_3$, $H = \{e, (123), (132)\}$
3. $G = S_3$, $H = \{e, (13)\}$
4. $G = (\mathbb{Z}, +)$, $H = 2\mathbb{Z}$
5. $G = GL(n, K)$, $H = SL(n, K)$
6. $G = GL(n, K)$, $H = \{\left[\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right], \left[\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right], \left[\begin{smallmatrix}-1 & 0\\ 0 & 1\end{smallmatrix}\right], \left[\begin{smallmatrix}-1 & 0\\ 0 & -1\end{smallmatrix}\right]\}$

Example 3.11. $\{e, (12), (23)\} \subset S_3$ is *not* a subgroup of S_3.

Lemma 3.12

Let G be a group and $H \subseteq G$ a non-empty subset. Then the following are equivalent:

1) H is a subgroup of G

2) $\forall x, y \in H,\ (xy \in H) \wedge (x^{-1} \in H)$

3) $\forall x, y \in H,\ xy^{-1} \in H$

Proof. That the first implies the second is immediate from the definition of a subgroup.
Assume the the second holds. Let $x, y \in H$. Then $y^{-1} \in H$ and therefore $xy^{-1} \in H$. Therefore the second implies the third.
Assume that the third condition holds. We will show that H is a subgroup. Note first that H is non-empty by hypothesis. Let $h \in H$. Then $e = hh^{-1} \in H$ by (3).

$$\begin{aligned} k \in H &\implies ek^{-1} \in H && \text{(by (3))} \\ &\implies k^{-1} \in H \end{aligned}$$

and therefore

$$\begin{aligned} h, k \in H &\implies h, k^{-1} \in H \\ &\implies h(k^{-1})^{-1} \in H && \text{(by (3))} \\ &\implies hk \in H && ((k^{-1})^{-1} = k) \end{aligned}$$

Therefore the group operation $G \times G \to G$ restricts to an operation $H \times H \to H$. We need to show that the axioms of a group are satisfied by H equipped with this operation. Let $h, k, l \in H$. Then we have

$$\begin{aligned} h(kl) &= (hk)l && \text{(since this holds for the original, unrestricted, operation)} \\ eh &= he = h && \text{(and } e \in H \text{ as noted above)} \\ hh^{-1} &= h^{-1}h = e && \text{(and } h^{-1} \in H \text{ as noted above)} \end{aligned}$$

□

Exercise 78. Let G be a group and $\{H_i \leqslant G \mid i \in I\}$ a set of subgroups of G. Show that $\cap_{i \in I} H_i$ is a subgroup of G.

Definition 3.13. Let G be a group and let $S \subseteq G$ be a subset of G. The **subgroup generated** by S is denoted by $\langle S \rangle$ and defined to be the subgroup given by theintersection of all subgroups that contain S. That is,

$$\langle S \rangle = \bigcap_{\substack{H \leqslant G \\ S \subseteq H}} H$$

Remark. It follows from the definition that $\langle\emptyset\rangle = \langle\{e\}\rangle = \{e\}$.

Example 3.14. We give some examples of subsets S of a group G and the generated generated.

G	$S \subset G$	$\langle S\rangle \leqslant G$
S_3	$\{(123)\}$	$\{e, (123), (132)\}$
S_3	$\{(12), (23)\}$	S_3
$(\mathbb{C} \setminus \{0\}, \times)$	{i}	{ 1,i,-1,i}
$\mathrm{GL}(2, \mathbb{R})$	$\{\left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right]\}$	$SL(2, \mathbb{Z})$

G	$S \subset G$	$\langle S\rangle \leqslant G$
$(\mathbb{Z}, +)$	$\{0\}$	$\{0\}$
$(\mathbb{Z}, +)$	$\{1\}$	$\mathbb{Z}$
$(\mathbb{Z}, +)$	$\{-1\}$	$\mathbb{Z}$
$(\mathbb{Z}, +)$	$\{2, 9\}$	$\mathbb{Z}$
$(\mathbb{Z}, +)$	$\{6, 9\}$	$3\mathbb{Z}$

The following result reflects the fact that the subgroup generated by S is the smallest subgroup of G that contains S.

Lemma 3.15

Let G be a group, $H \leqslant G$ a subgroup of G and $S \subseteq G$ a subset. Then

1) $S \subseteq \langle S\rangle$

2) $S \subseteq H \implies \langle S\rangle \leqslant H$

Proof. Both are almost immediate from the definition. □

3.1 Exercises

Exercise 79. List all of the subgroups of $\mathbb{Z}/12\mathbb{Z}$.

Exercise 80. Decide whether or not 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 in S_n which fix 1.

Exercise 81. 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. That is, show that the set $\{z \in \mathbb{C} \mid \exists n \in \mathbb{N},\ z^n = 1\}$ forms a subgroup of $\mathbb{C}^\times$.

Exercise 82. If H is a subgroup of a group G and if $g \in G$, show that $gHg^{-1} = \{ghg^{-1} \mid h \in H\}$ is a subgroup of G.

4 Cyclic groups

Lemma 3.16

Let $g \in G$. Then $\langle g\rangle = \{g^n \mid n \in \mathbb{Z}\}$.

Proof. Let $H = \{g^n \mid n \in \mathbb{Z}\}$. Note first that H is a subgroup of G, since $H \neq \emptyset$ and

$$\begin{aligned} h, k \in H &\implies h = g^m, k = g^n \quad \text{for some } m, n \in \mathbb{Z} \\ &\implies hk^{-1} = g^{m-n} \\ &\implies hk^{-1} \in H \end{aligned}$$

Therefore H is a subgroup of G and $g \in H$. Now suppose that K is a subgroup of G such that $g \in K$. For all $n \in \mathbb{Z}$ we have $g^n \in H$, because H is a subgroup. It follows that $K \leqslant H$ and hence $\langle g\rangle = H$.

□

Example 3.17. Let $g = (123) \in S_3$. Then $\langle g\rangle = \{e, (123), (132)\}$.

Definition 3.18. A group G is called **cyclic** if there exists $g \in G$ such that $\langle g \rangle = G$. Such an element g is called a **generator** for the cyclic group G.

Remark. It is clear from the definition that cyclic groups are abelian. The converse is false. The group V of Example 3.5 is abelian, but not cyclic.

Example 3.19.

1. $\mathbb{Z}$ is cyclic
2. $3\mathbb{Z}$ is a cyclic subgroup of $\mathbb{Z}$
3. $\langle 6, 9 \rangle \leqslant \mathbb{Z}$ is a cyclic subgroup
4. $\mathbb{Z}/6\mathbb{Z}$ is cyclic
5. S_3 is not cyclic
6. $\{\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right], \left[\begin{smallmatrix}-1&0\\0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}-1&0\\0&-1\end{smallmatrix}\right]\} \leqslant GL(2, \mathbb{R})$ is not a cyclic subgroup. (But it is abelian.)

Lemma 3.20

Every subgroup of a cyclic group is itself cyclic.

Proof. Let G be a cyclic group and $g \in G$ such that $G = \langle g \rangle$. Let H be a subgroup of G. If $H = \{e\}$, then H is cyclic. So assume that H is non-trivial. Let $d = \min\{m \in \mathbb{N} \mid g^n \in H\}$. We will show that $\langle g^d \rangle = H$. Let $h \in H$. Then, since $h \in G$, we have that $h = g^a$ for some $a \in \mathbb{Z}$. We need to show that $d \mid a$. Let $q, r \in \mathbb{Z}$ be such that $a = qd + r$ and $0 \leqslant r < d$. Then $h = (g^d)^q g^r$, which implies that $g^r \in H$. From the minimality of d we conclude that $r = 0$. □

5 Order of an element

Definition 3.21. Let G be a group and $g \in G$. Let $S\{n \in \mathbb{N} \mid g^n = e\}$. If $S = \emptyset$, we say that g has **infinite order**. If $S \neq \emptyset$ we say that g has **finite order** and define the **order** of g to be the minimal element of S. The order of g is denoted $o(g)$ or $|g|$.

Remark. The order of an element is equal to the size of the subgroup generated by g, i.e., $|g| = |\langle g \rangle|$.

Example 3.22.

1. The orders of the elements of S_3 are: $|e| = 1$, $|(123)| = 3$, $|(132)| = 3$, $|(12)| = 2$, $|(13) = 2$, $|(23)| = 2$.
2. $(12)(34) \in S_4$ has order 2
3. $(123)(45) \in S_5$ has order 6

Lemma 3.23

Let $g \in G$ and $n \in \mathbb{N}$. If $g^n = e$ then g has finite order and $|g|$ divides n.

Proof. That g has finite order is clear from the definition of order. Let $d = |g|$ and write $n = qd + r$ with $q, r \in \mathbb{Z}$ and $0 \leqslant r < d$. Then note that $g^n = g^{(qd+r)} = (g^d)^q g^r = e^q g^r = eg^r = g^r$. Therefore $g^r = e$ and $r < |g|$. Therefore $r = 0$ and hence $d \mid n$. □

Exercise 83. Let G be a group and $g \in G$.

(a) Suppose that g has infinite order. Show that $\forall m, n \in \mathbb{Z},\ g^m = g^n \implies m = n$.

(b) Suppose that g has finite order. Show that $\forall m, n \in \mathbb{Z},\ g^m = g^n \implies m \equiv n \pmod{|g|}$.

Lemma 3.24

Let G be a group and $g \in G$. Let $h \in \langle g \rangle \setminus \{e\}$.

1) If g has infinite order, then h has infinite order.

2) If g has finite order, then h has finite order and $|h| \mid |g|$.

Proof. Let $n \in \mathbb{Z} \setminus \{0\}$ be such that $h = g^n$. For that first part, we have the following.

$$\begin{aligned} h \text{ has finite order} &\implies \exists m \in \mathbb{N},\ h^m = e \\ &\implies (g^n)^m = e \\ &\implies g^{|mn|} = e && \text{(note that } mn \neq 0\text{)} \\ &\implies g \text{ has finite order} \end{aligned}$$

Now suppose that g has finite order. Note that $h^{|g|} = (g^n)^{|g|} = (g^{|g|})^n = e^n = e$, and therefore, by Lemma 3.23, we have that $|h| \mid |g|$. □

Example 3.25. We list the elements of $\langle g \rangle$ together with their orders for $g = (1243) \in S_4$.

$g^0 = e$	$g^1 = (1243)$	$g^2 = (14)(23)$	$g^3 = (1432)$
$\lvert g^0 \rvert = 1$	$\lvert g \rvert = 4$	$\lvert g^2 \rvert = 2$	$\lvert g^3 \rvert = 4$

5.1 Exercises

Exercise 84. Find the orders of the following elements:

(a) $(123)(4567)(89)$ in S_{10}

(b) $(14)(23567)$ in S_7

(c) a reflection in the plane

(d) a translation in the group of all 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 $\mathbb{Z}/13\mathbb{Z}$

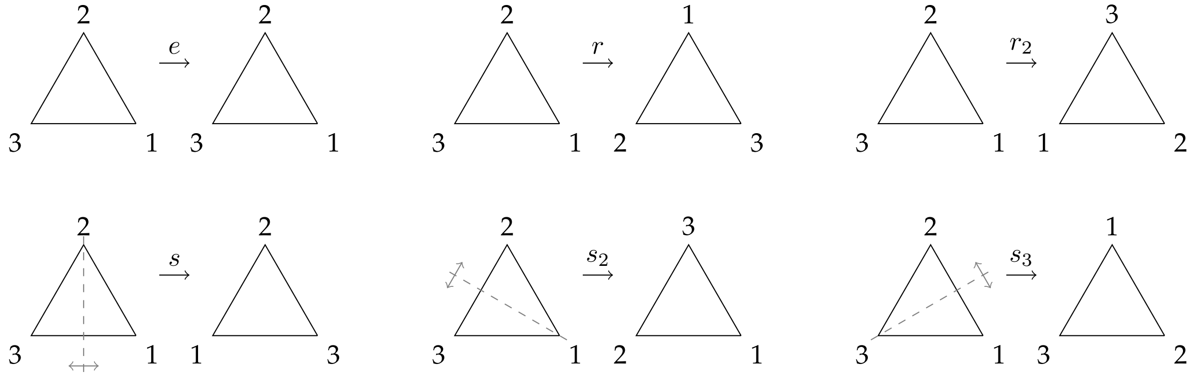
Exercise 85. If g is an element of a group G, prove that the orders of g and g^{-1} are equal.

Exercise 86. Show that, in an abelian group, the product of two elements of finite order again has finite order.

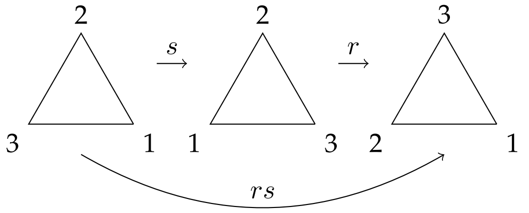
Exercise 87. Let $A, B \in GL(2, \mathbb{R})$ be given by $A = \left[\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right]$ and $B = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$. Show that A has order 3, that B has order 4, and that AB has infinite order.

6 The dihedral groups D_n

The dihedral groups are another important family of non-abelian finite groups. We start by describing the group D_3. Consider the ways in which two copies of an equilateral triangle can by placed one on top of the other. There are a total of six possibilities: three rotations (including the identity) and three reflections.



Two such maps can by combined. Denote by r the map given by rotating the triangle through $2\pi/3$ and by s the map given by reflection across the line indicated above. The product rs is the map given by first applying s and then applying r. The product rs is equal to s_2.



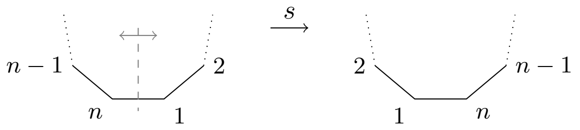
Similarly, we can show that $r^2s = sr = s_3$ and $r^2 = r_2$. Notice that $rs \neq sr$. Equipped with this operation the given set of six symmetries forms a (non-abelian, finite) group, which is denoted D_3. Given our calculations so far, we have $D_3 = \{e, r, r^2, s, rs, r^2s\}$. The multiplication table for this group is given on the right.

D_3	e	r	r^2	s	rs	r^2s
e	e	r	r^2	s	rs	r^2s
r	r	r^2	e	rs	r^2s	s
r^2	r^2	e	r	r^2s	s	rs
s	s	r^2s	rs	e	r^2	r
rs	rs	s	r^2s	r	e	r^2
r^2s	r^2s	rs	s	r^2	r	e

We can generalise from an equilateral triangle to a regular n-gon.

Definition 3.26. Let $n \in \mathbb{N}$ with $n \geqslant 3$. The **dihedral group D_n** is the group of symmetries of the regular n-gon. The group operation is composition.

For a fixed $n \geqslant 3$, we denote by $r \in D_n$ the element given by rotation through $2\pi/n$ and by $s \in D_n$ the element given by reflection across the perpendicular bisector of a fixed edge.



Proposition 3.27

The group D_n is a non-abelian group and has $2n$ elements. The elements r and s satisfy $r^n = e$, $s^2 = e$, $sr = r^{n-1}s$. The elements of D_n can be listed as

$$D_n = \{e, r, r^2, \dots, r^{n-1}, s, rs, r^2s, \dots r^{n-1}s\}$$

Proof. An element of D_n is uniquely determined by the image an edge. There are n choices for the image of the vertex labelled 1. Given a choice for the image of the vertex labelled 1, there are then two choices for the image of vertex labelled n. Hence $|D_n| = 2n$. It is obvious from the way in which they are defined that $r^n = e$ and $s^2 = e$. We now show that $sr = r^{n-1}s$. Given that an element of D_n is determined by the images of the vertices labelled 1 and n, it is enough to show that $sr(1) = r^{n-1}s(1)$ and $sr(n) = r^{n-1}s(n)$. We calculate

$$\begin{aligned} sr(1) &= s(2) = n-1 & \qquad r^{n-1}s(1) &= r^{n-1}(n) = r^{-1}(n) = n-1 \\ sr(n) &= s(1) = n & r^{n-1}s(n) &= r^{n-1}(1) = r^{-1}(1) = n \end{aligned}$$

Exercise 88. Finish the proof by showing that no two of the listed elements are equal.

□

Exercise 89. Determine the possible orders of elements in the dihedral group D_n.

7 Group homomorphisms

Definition 3.28. Let G and H be groups. A **homomorphism** from G to H is a function $\varphi : G \to H$ with the property that: $\forall x, y \in G,\ \varphi(xy) = \varphi(x)\varphi(y)$.

Example 3.29.

1. $\varphi : \mathbb{Z} \to \mathbb{Z}$, $\varphi(n) = 4n$
2. $\varphi : \mathbb{Z} \to \mathbb{Z}/6\mathbb{Z}$, $\varphi(n) = [n]_6$
3. $\varphi : GL(n, \mathbb{R}) \to \mathbb{R}^\times$, $\varphi(A) = \det(A)$
4. $\varphi : S_3 \to GL(3, K)$, $(\varphi(\sigma))_{ij} = \delta_{i,\sigma(j)}$

Lemma 3.30

Let $\varphi : G \to H$ be a homomorphism. Then

a) $\varphi(e_G) = e_H$

b) $\forall g \in G, \varphi(g^{-1}) = \varphi(g)^{-1}$

c) If $g \in G$ has finite order, then so does $\varphi(g)$ and $|\varphi(g)| \mid |g|$

d) If φ is a bijection, then the inverse function $\varphi^{-1} : H \to G$ is a homomorphism.

Proof. For part (a) we have:

$$\begin{aligned} \varphi(e_G) &= \varphi(e_G e_G) = \varphi(e_G)\varphi(e_G) \\ \implies \varphi(e_G)^{-1}\varphi(e_G) &= \varphi(e_G)^{-1}\varphi(e_G)\varphi(e_G) \\ \implies e_H &= e_H\varphi(e_G) \\ \implies e_H &= \varphi(e_G) \end{aligned}$$

Part (b) is left as an exercise.
For part (c), let $n = |g|$. We have $\varphi(g)^n = \varphi(g^n) = e_H$, which implies that $|\varphi(g)| \mid n$ by Lemma 3.23.
For part (d) we need to show that $\forall\, h_1, h_2 \in H$ we have $\varphi^{-1}(h_1h_2) = \varphi^{-1}(h_1)\varphi^{-1}(h_2)$. Note that

$$\begin{aligned} h_1h_2 &= \varphi(\varphi^{-1}(h_1))\varphi(\varphi^{-1}(h_2)) \\ &= \varphi(\varphi^{-1}(h_1)\varphi^{-1}(h_2)) \\ \implies \varphi^{-1}(h_1h_2) &= \varphi^{-1}(h_1)\varphi^{-1}(h_2) \end{aligned}$$

□

Example 3.31. The map $\varphi : \mathbb{Z}/4\mathbb{Z} \to \mathbb{Z}/6\mathbb{Z}$, $\varphi([n]_4) = [3m]_6$ is a homomorphism. Note that $|\varphi([1]_4)| = |[3]_6| = 2$ and $|[1]_4| = 4$.

Example 3.32. Let $m \in \mathbb{N}$. There is only one homomorphism $\varphi : \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}$. To see this, note that every element $g \in \mathbb{Z}/m\mathbb{Z}$ has finite order. Therefore, $\varphi(g) = 0$ as this is the only element of $\mathbb{Z}$ that has finite order. The only homomorpism is therefore $\varphi : \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}$, $\varphi(g) = 0$.

Definition 3.33. A bijective homomorphism is called an **isomorphism**. Two groups G and H are said to be **isomorphic** (denoted $G \cong H$) if there exists an isomorphism $G \to H$.

Remark. If two groups are isomorphic, then they are essentially the 'same' group. More precisely, any algebraic property satisfied by one will also be satisfied by the other. For example, if $G \cong H$ and G is abelian, then H is abelian.

Example 3.34.

1. $(\mathbb{Z}/4\mathbb{Z}, +) \cong (\{1, i, -1, -i\}, \times)$
2. $D_3 \cong S_3$
3. $(\mathbb{R}, +) \cong (\mathbb{R}^+, \times)$
4. $(\mathbb{Z}/4\mathbb{Z}, +) \ncong (\mathbb{Z}/3\mathbb{Z}, +)$
5. $(\mathbb{Z}/4\mathbb{Z}, +) \ncong V$

Exercise 90. Suppose that $\varphi : G \to H$ is an isomorphism. Show that

(a) $\varphi^{-1} : H \to G$ is an isomorphism

(b) $\forall\, g \in G,\ |\varphi(g)| = |g|$

Proposition 3.35

Let G be a cyclic group. If G is infinite, then $G \cong \mathbb{Z}$. If G is finite, then $G \cong \mathbb{Z}/m\mathbb{Z}$ where $m = |G|$.

Proof. Let $g \in G$ be such that $\langle g \rangle = G$.
Suppose first that g has infinite order. Define $\varphi : \mathbb{Z} \to G$ by $\varphi(m) = g^m$. Note that φ is a homomorphism since:

$$\varphi(m+n) = g^{m+n} = g^m g^n = \varphi(m)\varphi(n)$$

That φ is surjective follows from Lemma 3.16. It is also injective since

$$\varphi(m) = \varphi(n) \implies g^m = g^n \implies g^{m-n} = e \implies m - n = 0$$

Now suppose that g has finite order and let $m = |g|$. Define $\psi : \mathbb{Z}/m\mathbb{Z} \to G$ by $\psi([a]_m) = g^a$. Note that this map is well-defined because

$$[a]_m = [b]_m \implies m \mid (a-b) \implies a - b = mk \quad (\text{for some } k \in \mathbb{Z}) \implies g^{a-b} = g^{mk} = e^k = e \implies g^a = g^b$$

It is clear that ψ is surjective (Lemma 3.16). For injectivity we have

$$\psi([a]_m) = \psi([b]_m) \implies g^a = g^b \implies g^{a-b} = e \implies m \mid (a-b) \implies [a]_m = [b]_m$$

□

7.1 Exercises

Exercise 91. Show that the matrix group $SO(2)$ is isomorphic to the group $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ of complex numbers having modulus 1 (and operation given by multiplication of complex numbers).

Exercise 92. Show that if m divides n, then D_m is isomorphic to a subgroup of D_n.

Exercise 93. Show that:

(a) $(\mathbb{R}, +)$ and $(\mathbb{R} \setminus \{0\}, \times)$ 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 $(\mathbb{Q}^+, \times)$.

Exercise 94. Let $\mathcal{H} = \left\{ \left[\begin{smallmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{smallmatrix}\right] \mid a, b, c \in \mathbb{F}_3 \right\} \leqslant GL(2, \mathbb{F}_3)$. Show that the group $|\mathcal{H}| = 27$ and that all non-identity elements have order 3. The group $G = \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ also has order 27 and has all non-identity elements of order 3. Are $\mathcal{H}$ and G isomorphic?

8 Direct product

Definition 3.36. Let G and H be groups. The **direct product** of G and H is the group with underlying set the cartesian product

$$G \times H = \{(g, h) \mid g \in G, h \in H\}$$

and operation given by

$$(g_1, h_1) * (g_2, h_2) = (g_1 g_2, h_1 h_2)$$

Exercise 95. (a) Show that $(G \times H, *)$ forms a group and that $e_{G \times H} = (e_G, e_H)$.

(b) Show that if G and H are both abelian, then $G \times H$ is abelian

Example 3.37. 1. $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ is a (non-cyclic) group of size 4

2. $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ and $\mathbb{Z}/8\mathbb{Z}$ are all abelian groups of size 8. No two are isomorphic.

9 Cosets and Lagrange's theorem

Definition 3.38. Let G be a group and $H \leqslant G$ a subgroup. The set $gH = \{gh \mid h \in H\}$ is called a **left coset** of H in G. The set $Hg = \{hg \mid h \in H\}$ is called a **right coset** of H in G.

Remark. 1. H itself is both a left and right coset: $eH = He = H$.

2. If $g \notin H$, then gH is not a subgroup of G. Similarly, Hg is not a subgroup.

Example 3.39. 1. If $G = \mathbb{Z}$ and $H = 3\mathbb{Z}$, there are three (left) cosets: $0 + H = [0]_3, 1 + H = [1]_3, 2 + H = [2]_3$.

2. Let $G = S_3$ and $H = \{e, (123), (132)\}$. There are two left cosets:

$$eH = (123)H = (132)H = H \quad \text{and} \quad (12)H = (13)H = (23)H = \{(12), (13), (23)\}$$

There are two right cosets:

$$He = H(123) = H(132) = H \quad \text{and} \quad (12)H = H(13) = H(23) = \{(12), (13), (23)\}$$

3. Let $G = S_3$ and $H = \{e, (12)\}$. There are three left cosets:

$$eH = (12)H = H \quad \text{and} \quad (123)H = (13)H = \{(123), (13)\} \quad \text{and} \quad (132)H = (23)H = \{(132), (23)\}$$

There are three right cosets:

$$He = H(12) = H \quad \text{and} \quad H(123) = H(23) = \{(123), (23)\} \quad \text{and} \quad H(132) = H(13) = \{(132), (13)\}$$

Note that, in this example, the left and right cosets are not the same.

Lemma 3.40

Let G be a group and $H \leqslant G$ a subgroup. Let $a, b \in G$.

a) (i) $aH = bH \iff a^{-1}b \in H$ (ii) $Ha = Hb \iff ab^{-1} \in H$

b) (i) The left cosets partition G. (ii) The right cosets partition G.

c) (i) The map $aH \to bH$, $ah \mapsto bh$ is a bijection. (ii) The map $Ha \to Hb$, $ha \mapsto hb$ is a bijection.

Proof. We prove the statements for left cosets, and leave the right coset versions as an exercise.

$$\begin{aligned} aH = bH &\implies b \in aH \\ &\implies b = ah && \text{(for some } b \in H) \\ &\implies a^{-1}b = h \in H \end{aligned}$$

Conversely, suppose that $a^{-1}b \in H$. Then

$$x \in aH \implies x = ah \quad (\text{for some } h \in H) \implies x = b(a^{-1}b)^{-1}h \implies x \in bH \quad (\text{since } (a^{-1}b)^{-1} \in H)$$
$$x \in bH \implies x = bh \quad (\text{for some } h \in H) \implies x = a(a^{-1}b)h \implies x \in aH \quad (\text{since } (a^{-1}b) \in H)$$

Therefore (a) holds.
For (b) we need to show that every element of G is contained in exactly one coset. Let $g \in G$. There is at least one coset that contains g since $g \in gH$. Suppose now that $g \in kH$. Our aim is to show that $kH = gH$. Using part (a) we have

$$g \in kH \implies g = kh \quad (\text{for some } h \in H) \implies k^{-1}g \in H \implies kH = gH$$

For part (c), let $f : aH \to bH$ be the map $f(ah) = bh$. We have

$$f(ah_1) = f(ah_2) \implies bh_1 = bh_2 \implies h_1 = h_2 \implies ah_1 = ah_2$$
$$x \in bH \implies x = bh \quad (\text{for some } h \in H) \implies x = f(ah)$$

□

Remark. 1. It follows from part (c) that $\forall\, g \in G,\ |gH| = |Hg| = |H|$. That is, all cosets (left and right) have the same size as H.

2. It follows from the lemma that the number of left cosets is equal to the number of right cosets.

Definition 3.41. Let G be a group and $H \leqslant G$ a subgroup. The number of cosets of H in G is called the **index** of H in G and is denoted by $[G : H]$. That is,

$$[G : H] = |\{gH \mid g \in G\}|$$

Example 3.42. (cf. Example 3.39)

1. $[\mathbb{Z} : 3\mathbb{Z}] = 3$
2. $[S_3 : \langle(123)\rangle] = 2$
3. $[S_3 : \langle(12)\rangle] = 3$

That the cosets partition G and all have the same size leads directly to the following fundamental and useful result.

Theorem 3.43: Lagrange's Theorem

Let G be a finite group and $H \leqslant G$ a subgroup. Then $|G| = [G : H]|H|$.

Proof. We saw in Lemma 3.40 that the left cosets partition G and all have size equal to $|H|$. Let $k = [G : H]$ and $g_1, \dots, g_k \in G$ be such that the (distinct) cosets are $g_1H, \dots, g_kH$. Then

$$\begin{aligned} |G| &= |g_1H| + \cdots + |g_kH| && \text{(cosets are disjoint)} \\ &= k|H| && (|g_iH| = |H|) \\ &= [G : H]|H| \end{aligned}$$

□

Example 3.44. 1. $|S_3| = 6 = 2 \times 3 = [S_3 : \langle(123)\rangle]\,|\langle(123)\rangle|$

2. $|S_3| = 6 = 3 \times 2 = [S_3 : \langle(12)\rangle]\,|\langle(12)\rangle|$

Example 3.45. Since $|S_4| = 24$ and $|\langle(12), (34)\rangle| = 4$, we deduce that $[S_4 : \langle(12), (34)\rangle] = 6$.

Corollary 3.46

Let G be a finite group and $g \in G$. Then $g^{|G|} = e$ and $|g| \mid |G|$.

Corollary 3.47

Let G be a finite group. If $|G|$ is prime, then $G \cong \mathbb{Z}/p\mathbb{Z}$, where $p = |G|$.

9.1 Exercises

Exercise 96. If H and K are subgroups of a group G and if $|H| = 7$ and $|K| = 29$, show that $H \cap K = \{e_G\}$.

Exercise 97. Let G be the subgroup of $GL(2, \mathbb{R})$ of the form

$$G = \left\{ \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} \mid x, y \in \mathbb{R}, x > 0 \right\}$$

Let H be the subgroup of G defined by

$$H = \left\{ \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} \mid z \in \mathbb{R}, z > 0 \right\}$$

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.

Exercise 98. Consider the set of linear equations of the form $AX = B$, 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 $(\mathbb{R}^n, +)$.

Exercise 99. (a) Let H be a subgroup of index 2 in a group G. Show that if $a, b \in G \setminus H$, then $ab \in H$.

(b) Let H be a subgroup of a group G with the property that if $a, b \in G \setminus H$, then $ab \in H$. Show that H has index 2 in G.

Exercise 100. Determine all subgroups of the dihedral group D_5.

Exercise 101. Determine all subgroups of the dihedral group D_4 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.

Exercise 102. Let G denote the group of rotational symmetries of a regular tetrahedron. Note that $|G| = 12$.

(a) Show that G has subgroups of order 1,2,3,4 and 12.

(b) Show that G has no subgroup of order 6.

Exercise 103. Let G be a group of order 841 (which is $(29)^2$). Show that if G is not cyclic, then every element $g \in G$ satisfies $g^{29} = 1$.

10 Normal subgroups and quotient groups

Given a group G and a subgroup $H \leqslant G$, we would like to define a group G/H in a way that mimics the construction of $(\mathbb{Z}/m\mathbb{Z}, +)$. The set will be the set of all (left) cosets, but what should the operation be? The natural choice to make is to define $aH * bH = (ab)H$. However, this is not always well-defined.

For example, consider $G = S_3$ and $H = \{e, (12)\}$. The left cosets are $C_1 = \{e, (12)\}$, $C_2 = \{(23), (132)\}$, $C_3 = \{(13), (123)\}$. What should $C_1 * C_2$ be? The coset $(ab)H$ depends on the choice of a and b:

$$C_1 * C_2 = eH * (23)H = (e(23))H = (23)H = C_2$$

but, also

$$C_1 * C_2 = (12)H * (23)H = ((12)(23))H = (123)H = C_3$$

The solution is to put a condition on the subgroup H.

Definition 3.48. A subgroup $H \leqslant G$ is called a **normal subgroup** if $\forall g \in G$, $gH = Hg$. This will be denoted $H \lhd G$.

Remark. It is immediate from the definition that $\{e\} \lhd G$ and $G \lhd G$.

Exercise 104. Let H be a subgroup of a group G. Show that H is normal if and only if $\forall g \in G\, \forall h \in H$, $ghg^{-1} \in H$.

Remark. If G is abelian, then all subgroups of G are normal.

Example 3.49. 1. $3\mathbb{Z} \lhd \mathbb{Z}$

2. $\langle (123) \rangle \lhd S_3$

3. $SL(n, K) \lhd GL(n, K)$

4. $\langle (12) \rangle \ntriangleleft S_3$

5. $\langle (123) \rangle \ntriangleleft S_4$

Exercise 105. Let $G = S_4$ and $H = \{e, (12)(34), (13)(24), (14)(23)\}$. Show that $H \lhd G$.

Exercise 106. Let G be a group and $H \leqslant G$ a subgroup. Show that if $[G : H] = 2$, then $H \lhd G$.

Definition 3.50. Let G be a group and $H \lhd G$ a normal subgroup. The **quotient group** G/H is the group whose elements are the (left) cosets $G/H = \{gH \mid g \in G\}$ and whose operation is given by $(g_1H) * (g_2H) = (g_1g_2)H$.

Exercise 107. Check that the above operation is well-defined and that G/H is a group and $e_{G/H} = e_G H$.

Remark. 1. If G is finite, from Lagrange's Theorem we have $|G/H| = |G|/|H|$.

2. If $G = \mathbb{Z}$ and $H = m\mathbb{Z}$, the notation G/H agrees with for our existing notation for $\mathbb{Z}/m\mathbb{Z}$.

Example 3.51. Let $G = D_4$ and $r, s \in D_4$ as in Definition 3.26. Then $H = \{e, r^2\}$ is a normal subgroup of G. The multiplication table for $D_4/\langle r^2 \rangle$ is

	eH	rH	sH	rsH
eH	eH	rH	sH	rsH
rH	rH	eH	rsH	sH
sH	sH	rsH	eH	rH
rsH	rsH	sH	rH	eH

In fact, $D_4/\langle r^2\rangle \cong (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$ (see Example 3.57).

10.1 Exercises

Exercise 108. 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.

Exercise 109. Show that if K and L are normal subgroups of a group G, then $K \cap L$ is a normal subgroup of G.

Exercise 110. Let G be a group and $n \in \mathbb{N}$. Show that if there is exactly one subgroup of order n, then it is normal.

Exercise 111. Find all of the normal subgroups of D_4. (See Exercise 101.)

Exercise 112. 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 hence that these 8 elements do give a subgroup of $GL(2,\mathbb{C})$.

(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) Show that all subgroups of Q_8 are normal. (Even though Q_8 is not abelian.)

(e) Are Q_8 and D_4 isomorphic?

Exercise 113. (a) Show that if G is an abelian group, then every quotient G/N is abelian.

(b) Show that if G is a cyclic group, then every quotient G/N is cyclic.

Exercise 114. 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 $\mathbb{R}/\mathbb{Z}$.

Exercise 115. Let H denote the subgroup of D_8 generated by r^4 (where, as in Definition 3.26, r is rotation by $\pi/4$).

(a) Show that H is normal.

(b) Write out the multiplication table of D_8/H.

11 The first isomorphism theorem

Definition 3.52. Let $\varphi : G \to H$ be a homomorphism. The **kernel** of φ is defined to be

$$\ker(\varphi) = \{g \in G \mid \varphi(g) = e_H\}$$

The **image** of φ is defined to be

$$\operatorname{im}(\varphi) = \{\varphi(g) \mid g \in G\}$$

Example 3.53. 1. $\varphi : \mathbb{Z} \to \mathbb{Z}$, $\varphi(m) = 4m$. Then $\ker(\varphi) = \{0\}$ and $\operatorname{im}(\varphi) = 4\mathbb{Z}$.

2. $\varphi : \mathbb{Z} \to \mathbb{Z}/6\mathbb{Z}$, $\varphi(m) = [4m]_6$. Then $\ker(\varphi) = 3\mathbb{Z}$ and $\operatorname{im}(\varphi) = \{[0]_6, [2]_6, [4]_6\}$.

3. $\varphi : GL(n, \mathbb{R}) \to \mathbb{R}^\times$, $\varphi(A) = \det(A)$. Then $\ker(\varphi) = SL(n, \mathbb{R})$ and $\operatorname{im}(\varphi) = \mathbb{R}^\times$.

Exercise 116. Show that $\ker(\varphi)$ is a subgroup of G and that $\operatorname{im}(\varphi)$ is a subgroup of H.

Lemma 3.54

Let $\varphi : G \to H$ be a homomorphism.

1. $\ker(\varphi) \lhd G$
2. φ is injective if and only if $\ker(\varphi) = \{e\}$

Proof. The $\ker(\varphi)$ is a subgroup of G is shown in Exercise 116. To see that it is normal, let $g \in G$ and $k \in \ker(\varphi)$. Then $\varphi(gkg^{-1}) = \varphi(g)\varphi(k)\varphi(g)^{-1} = \varphi(g)\varphi(h)\varphi(g)^{-1} = e_H$. Therefore $gkg^{-1} \in \ker(\varphi)$ and hence $\ker(\varphi)$ is normal by 104.

If φ is injective, then $k \in \ker(\varphi) \implies \varphi(k) = \varphi(e_G) \implies k = e_G$, and therefore $\ker(\varphi) = \{e_G\}$.

Now suppose that $\ker(\varphi) = \{e_G\}$. For $g_1, g_2 \in G$ we have

$$\varphi(g_1) = \varphi(g_2) \implies \varphi(g_1)\varphi(g_2)^{-1} = e_H \implies \varphi(g_1g_2^{-1}) = e_H \implies g_1g_2^{-1} \in \ker(\varphi) \implies g_1g_2^{-1} = e_G \implies g_1 = g_2$$

Therefore, if $\ker(\varphi) = \{e_G\}$ then φ is injective. □

Not only is the kernel of a homomorphism normal, every normal subgroup is the kernel of some homomorphism.

Lemma 3.55

Let G be a group and $H \lhd G$ a normal subgroup. Then the map $\varphi : G \to G/H$, $\varphi(g) = gH$ is a surjective homomorphism and $\ker(\varphi) = H$.

Remark. The above map $G \to G/H$ is often called the **projection map**.

Proof. That the map is a surjective homomorphism is clear from the definition of the quotient group G/H. Further, $k \in \ker(\varphi) \iff \varphi(k) = e_{G/H} \iff kH = H \iff k \in H$. □

Theorem 3.56: First isomorpism theorem

Let $\varphi : G \to H$ be a homomorphism and let $K = \ker(\varphi)$. Then the map $\bar{\varphi} : G/K \to H$ given by $\bar{\varphi}(gK) = \varphi(g)$ is an injective homomorphism. It follows that $G/\ker(\varphi) \cong \operatorname{im}(\varphi)$.

Proof. First we verify that the given map is well-defined:

$$g_1K = g_2K \implies g_1^{-1}(g_2) \in K \implies \varphi(g_1^{-1}g_2) = e_H \implies \varphi(g_1)^{-1}\varphi(g_2) = e_H \implies \varphi(g_1) = \varphi(g_2)$$

Now that $\bar{\varphi}$ is a homomorphism:

$$\bar{\varphi}((g_1K)(g_2K)) = \bar{\varphi}((g_1g_2)K) = \varphi(g_1g_2) = \varphi(g_1)\varphi(g_2) = \bar{\varphi}(g_1K)\bar{\varphi}(g_2K)$$

It is injective:

$$\bar{\varphi}(gK) = e_H \implies \varphi(g) = e_H \implies g \in K \implies gK = K \implies gK = e_{G/K}$$

□

Example 3.57. (cf. Example 3.51) Let $\varphi : D_4 \to (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$ be given by

$$\varphi(e) = e, \varphi(r) = (1,0), \varphi(r^2) = (0,0), \varphi(r^3) = (1,0), \varphi(s) = (0,1), \varphi(rs) = (1,1), \varphi(r^2s) = (0,1), \varphi(r^3s) = (1,1)$$

Then φ is a surjective homomorphism and $\ker(\varphi) = \{e, r^2\}$. Therefore $D_4/\langle r^2\rangle \cong (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$.

Exercise 117. Let $\varphi : \mathbb{Z}/8\mathbb{Z} \to H$ be a homomorphism. Show that $\operatorname{im}(\varphi)$ is isomorphic to one of: $\{e\}$, $\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/8\mathbb{Z}$.

Chapter 4

Linear Algebra II

1 Inner product spaces

In this chapter, unless explicitly stated otherwise, K will denote either $\mathbb{R}$ or $\mathbb{C}$. We start by recalling the definition of an inner product on a vector space. Having an inner product will enable us to define geometric notions such as length.

Definition 4.1. Let V be a K-vector space. An **inner product** on V is a function $V \times V \to K$ (with the image of (u, v) being denoted $\langle u, v\rangle$) that satisfies the following conditions.

1) $\forall u, v \in V \quad \langle v, u\rangle = \overline{\langle u, v\rangle}$

2) $\forall u, v, w \in V\ \forall k, l \in K \quad \langle ku + lv, w\rangle = k\langle u, w\rangle + l\langle v, w\rangle$

3) (a) $\forall u \in V \quad \langle u, u\rangle \geqslant 0$

   (b) $\forall u \in V \quad \langle u, u\rangle = 0 \implies u = 0$

An **inner product space** is a vector space equipped with an inner product.

Remark. 1. The first condition implies that $\forall u \in V$, $\langle u, u\rangle \in \mathbb{R}$.

2. The first and second conditions imply that $\forall u, v \in V\ \forall k \in K$, $\langle u, kv\rangle = \overline{k}\langle u, v\rangle$.

Exercise 118. Show, using the above axioms, that $\forall u \in V$, $\langle 0, u\rangle = 0$.

Exercise 119. Show that the following defines an inner product on $\mathbb{C}^2$.

$$\langle (x_1, x_2), (y_1, y_2)\rangle = x_1\overline{y_1} + ix_1\overline{y_2} - ix_2\overline{y_1} + 2x_2\overline{y_2}$$

Example 4.2. 1. $V = \mathbb{R}^n$ equipped with the usual dot product.

2. The **standard inner product** on $\mathbb{C}^n$ is $\langle (u_1, \ldots, u_n), (v_1, \ldots, v_n)\rangle = u_1\overline{v_1} + \cdots + u_n\overline{v_n}$

3. $V = M_n(K)$, $\langle A, B\rangle = \operatorname{tr}(A(\overline{B})^t)$

4. $V = \mathcal{C}([a, b], \mathbb{C})$, $\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}\, dt$

Definition 4.3. Let V be an inner product space.

1) The **length** (or norm) of a vector $u \in V$ is defined to be $\|u\| = \sqrt{\langle u, u\rangle}$

2) The **distance function** (or metric) on V is defined to be $d : V \times V \to \mathbb{R}_{\geqslant 0}$ given by $d(u, v) = \|u - v\|$

3) Two vectors $u, v \in V$ are said to be **orthogonal** if $\langle u, v\rangle = 0$

4) The **orthogonal complement** of a subspace $W \leqslant V$ is defined to be $W^\perp = \{u \in V \mid \forall w \in W\ \langle u, w\rangle = 0\}$

5) A subset $S \subseteq V$ is said to be **orthonormal** if

$$\forall u, v \in S\ \langle u, v\rangle = \begin{cases} 1 & \text{if } u = w \\ 0 & \text{if } u \neq w \end{cases}$$

Example 4.4. 1. With $V = \mathbb{C}^2$, $\langle (x_1, x_2), (y_1, y_2)\rangle = x_1\overline{y_1} + ix_1\overline{y_2} - ix_2\overline{y_1} + 2x_2\overline{y_2}$ we have $\|(1, i)\| = \sqrt{5}$

2. $V = \mathcal{C}([0, 1], \mathbb{C})$, $\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\,dt$. Let $f, g \in V$ be given by $f(t) = e^{2\pi it}$ and $g(t) = e^{4\pi it}$. Then $\|f\| = 1$ and $\langle f, g\rangle = 0$.

3. $V = \mathcal{C}(\mathbb{R}, \mathbb{R})$, $\langle f, g\rangle = \int_{-1}^1 f(t)g(t)\,dt$, $S = \{\frac{1}{\sqrt{2}}, \sin(\pi t), \cos(\pi t), \sin(2\pi t), \cos(2\pi t), \dots\}$ is an (infinite) orthonormal set.

Lemma 4.5

Let V be an inner product space and $S \subseteq V$ a subset. If S is orthonormal, then S is linearly independent.

Proof. Let $u_1, \dots, u_n \in S$ and $k_i, \dots, k_n \in K$ be such that $\sum_{i=1}^n \alpha_i u_i = 0$. Then for all j we have

$$0 = \langle 0, u_j\rangle = \langle \sum_{i=1}^n k_i u_i, u_j\rangle = \sum_{i=1}^n k_i \langle u_i, u_j\rangle = k_j$$

□

Exercise 120. Find the length of

(a) $(2 + i, 3 - 2i, -1)$ in the standard inner product on $\mathbb{C}^3$.

(b) $x^2 - 3x + 1 \in \mathcal{P}_2(R)$ using inner product $\langle p(x), q(x)\rangle = \int_0^1 p(x)q(x)\,dx$.

(c) $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \in M_2(\mathbb{C})$ using inner product $\langle A, B\rangle = \operatorname{tr}(A(\overline{B})^t)$.

Exercise 121. 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?

2 Gram-Schmidt

Bases that are orthonormal are convenient to work with (see Proposition 4.8 below, for example). Although not all vector spaces admit an orthonormal basis, all finite dimensional vector spaces do.

Theorem 4.6: Gram-Schmidt

Let V be a finite dimensional inner product space. Any orthonormal set $S \subset V$ can be extended to an orthonormal basis.

Remark. It follows that every finite dimensional inner product space has an orthonormal basis.

Proof. Let $S \subset V$ be an orthonormal set. Then S is linearly independent and therefore $|S| \leqslant \dim(V)$. Say $S = \{u_i, \dots, u_k\}$. We want to show that there is a basis $\mathcal{B}$ with $\mathcal{B} \supseteq S$. If S is a spanning set, we take $\mathcal{B} = S$. Otherwise, let $w \in V \setminus \operatorname{span}(S)$ and let $v = w - \sum_{i=1}^k \langle w, u_i\rangle u_i$. Note that $v \neq 0$ since $w \notin \operatorname{span}(S)$. Also, $\forall j \in \{1, \dots, k\}$ we have

$$\langle v, u_j\rangle = \langle w, u_j\rangle - \sum_{i=1}^k \langle w, u_i\rangle\langle u_i, u_j\rangle = \langle w, u_j\rangle - \langle w, u_j\rangle = 0$$

Defining $u_{k+1} = v/\|v\|$, the set $\{u_1, \dots, u_k, u_{k+1}\}$ is orthonormal. If S' is a spanning set for V, then it is a basis and we are done. Otherwise we repeat the above with S' in place of S. □

Example 4.7. Consider $\mathcal{P}_2(\mathbb{R})$ quipped with the inner product $\langle p(x), q(x)\rangle = \int_0^1 p(x)q(x)\,dx$. The set $S = \{1\}$ is an orthonormal set. We extend to an orthonormal basis in the way described in the above proof. Note that

$x \notin \operatorname{span}\{1\}$ and $x^2 \notin \operatorname{span}\{1, x\}$. We have

$$\begin{aligned}
v_1 &= x - \langle x, 1\rangle 1 = x - \frac{1}{2} \\
\|v_1\|^2 = \langle v_1, v_1\rangle &= \int_0^1 (x - \frac{1}{2})^2\, dx = \frac{1}{12} \\
u_1 &= v_1/\|v_1\| = \sqrt{3}(2x-1) \\
v_2 &= x^2 - \langle x^2, 1\rangle 1 - \langle x^2, u_1\rangle u_1 = x^2 - x + \frac{1}{6} \\
\|v_2\|^2 &= \int_0^1 (x^2 - x + \frac{1}{6})^2\, dx = \frac{1}{180} \\
u_2 &= v_2/\|v_2\| = \sqrt{5}(6x^2 - 6x + 1)
\end{aligned}$$

The set $\{1, \sqrt{3}(2x-1), \sqrt{5}(6x^2-6x+1)\}$ is an orthonormal basis for $\mathcal{P}_2(x)$.

Proposition 4.8

Let V be an inner product space and $S = \{u_1, \dots, u_n\}$ an orthonormal set. Let $v \in V$.

1) $\sum_{i=1}^n |\langle v, u_i\rangle|^2 \leqslant \|v\|^2$
2) If S is a basis, then $v = \sum_{i=1}^n \langle v, u_i\rangle u_i$

Proof. We have

$$\begin{aligned}
\|v - \sum_{i=1}^n \langle v, u_i\rangle u_i\|^2 &= \langle v - \sum_{i=1}^n \langle v, u_i\rangle u_i, v - \sum_{i=1}^n \langle v, u_i\rangle u_i\rangle \\
&= \langle v, v\rangle - \sum_{i=1}^n \overline{\langle v, u_i\rangle}\langle v, u_i\rangle - \sum_{i=1}^n \langle v, u_i\rangle\langle u_i, v\rangle + \sum_{i=1}^n\sum_{j=1}^n \langle v, u_i\rangle\overline{\langle v, u_j\rangle}\langle u_i, u_j\rangle \\
&= \langle v, v\rangle - \sum_{i=1}^n |\langle v, u_i\rangle|^2 - \sum_{i=1}^n |\langle v, u_i\rangle|^2 + \sum_{i=1}^n |\langle v, u_i\rangle|^2 \\
&= \langle v, v\rangle - \sum_{i=1}^n |\langle v, u_i\rangle|^2
\end{aligned}$$

Therefore $\langle v, v\rangle - \sum_{i=1}^n |\langle v, u_i\rangle|^2 \geqslant 0$. The final statement is left as an exercise. □

Corollary 4.9: Cauchy-Schwartz

Let V be an inner product space. Then $\forall\, u, v \in V, \quad |\langle u, v\rangle| \leqslant \|u\|\|v\|$

Proof. If $u = 0$, then the inequality holds since both sides are zero. So we can assume that $u \neq 0$. Apply Proposition 4.8 with $S = \{u/\|u\|\}$. □

Example 4.10. 1. If we take $V = \mathbb{R}^n$ and the dot product, this becomes

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

for any real numbers a_i, b_i.

2. If we take the inner product space of Example 4.4.2 above, then we have

$$\left|\int_0^1 f(t)\overline{g(t)}\, dt\right| \leqslant \left(\int_0^1 f(t)^2\, dt\right)^{\frac{1}{2}} \left(\int_0^1 g(t)^2\, dt\right)^{\frac{1}{2}}$$

for any $f, g \in \mathcal{C}([0,1], \mathbb{C})$.

Exercise 122. Let V be an inner product space. Show that the distance function $d : V \times V \to \mathbb{R}$ (defined by $d(u, v) = \|u - v\|$) satisfies the following properties:

(a) $d(u,v) = 0 \iff u = v$ (b) $d(u,v) = d(v,u)$ (c) $d(u,v) \leqslant d(u,w) + d(w,v)$

3 Orthogonal complements

Definition 4.11. Let V be an inner product space and let $W \leqslant V$ be a subspace. The **orthogonal complement** of W in V is denoted $W^\perp$ and defined to be

$$W^\perp = \{u \in V \mid \forall w \in W,\ \langle u, w\rangle = 0\}$$

Exercise 123. Show that

(a) $W^\perp$ is a subspace of V (b) $W \cap W^\perp = \{0\}$ (c) $W \subseteq (W^\perp)^\perp$

Proposition 4.12

Let V be a finite dimensional inner product space and let $W \leqslant V$ be a subspace. Then $V = W \oplus W^\perp$.

Proof. We know that $W \cap W^\perp = \{0\}$ from Exercise 123. It remains to show that $V = W + W^\perp$. From Theorem 4.6 we know that W has an orthonormal basis, say $\{w_1, \dots, w_k\}$. Given $u \in V$ define $w = \sum_{i=1}^k \langle u, w_i\rangle w_i$. Then $w \in W$ and $\langle u - w, w_i\rangle = 0$ for all i. Therefore $u - w \in W^\perp$ and we have $u = w + (u - w) \in W + W^\perp$. □

Remark. It follows from the proposition that $\dim(V) = \dim(W) + \dim(W^\perp)$.

Exercise 124. Show that if V is a finite dimensional inner product space and $W \leqslant V$ is a subspace of V, then $(W^\perp)^\perp = W$.

Example 4.13. This is an example in which $W \neq (W^\perp)^\perp$. Denote by ℓ^2 the vector space of all square-summable real-valued sequences, that is

$$\ell^2 = \{(x_1, x_2, \dots) \mid x_i \in \mathbb{R} \quad \text{and} \quad \sum_{i=1}^{\infty} |x_i|^2 \text{ converges}\}$$

The following is an inner product on ℓ^2

$$\langle (x_1, x_2, \dots), (y_1, y_2, \dots)\rangle = \sum_{i=1}^{\infty} x_i y_i$$

Let W be the subspace of ℓ^2 consisting of all sequences that are eventually zero, that is,

$$W = \{(x_1, x_2, \dots) \mid x_i \in \mathbb{R}, \quad \exists N \in \mathbb{N} \text{ such that } i \geqslant N \implies x_i = 0\}$$

Now define $v \in \ell^2$ to be the sequence $v = (1/i)_{i \in \mathbb{N}}$. Clearly, $v \notin W$, however $v \in (W^\perp)^\perp$ because for any $(\xi_i) \in W^\perp$ we have

$$\langle v, \xi\rangle = \sum_{i=1}^{\infty} \xi_i v_i = \lim_{N\to\infty} \sum_{i=1}^{N} \xi_i v_i = \lim_{N\to\infty} \langle \xi, u_i\rangle = \lim_{N\to\infty} 0 = 0$$

where $u_i \in W$ is the sequence given by $(u_i)_j = \begin{cases} v_j & j \leqslant i \\ 0 & j > i \end{cases}$

Therefore $W \subsetneqq (W^\perp)^\perp$.

4 Adjoint transformations

Definition 4.14. Let V be an inner product space and $f : V \to V$ a linear transformation. An **adjoint** of f is a linear transformation $f^* : V \to V$ satisfying

$$\forall u, v \in V \quad \langle f(u), v\rangle = \langle u, f^*(v)\rangle$$

For a matrix $A \in M_n(K)$ the notation A^* is used to denote the matrix $A^* = (\overline{A})^t$.

Lemma 4.15

Let V be an inner product space and $f : V \to V$ a linear transformation.

1. If an adjoint of f exists, it is unique. (This justifies the notation f^*.)
2. If V is finite dimensional, then an adjoint of f exists.

Proof. For the first part, suppose that $g, h : V \to V$ are such that

$$\forall u, v \in V \quad \langle f(u), v\rangle = \langle u, g(v)\rangle = \langle u, h(v)\rangle$$

Let $v \in V$ and define $u = g(v) - h(v)$. We have

$$\begin{aligned}\langle u, u\rangle &= \langle u, g(v)\rangle - \langle u, h(v)\rangle \\ &= \langle f(u), v\rangle - \langle f(u), v\rangle \\ &= 0\end{aligned}$$

From which it follows that $g(v) = h(v)$. Since this holds for all $v \in V$, we have that $g = h$.
We now establish the second part. Since V is finite dimensional, there is an orthonormal basis. Let $\mathcal{B} = \{b_1, \dots, b_n\}$ be an orthonormal basis for V and let $A = [f]_{\mathcal{B}}$. Let $g : V \to V$ be the linear transformation determined by the condition that $[g]_{\mathcal{B}} = A^*$. We will now show that g is an adjoint for f. Denote the entries in the matrix A by A_{ij}.

$$\begin{aligned}\langle f(b_i), b_j\rangle &= \langle \sum_{k=1}^{n} A_{ki} b_k, b_j\rangle = \sum_{k=1}^{n} A_{ki}\langle b_k, b_j\rangle = A_{ji} \\ \langle b_i, g(b_j)\rangle &= \langle b_i, \sum_{i=1}^{k} (A^*)_{kj} b_k\rangle = \sum_{i=1}^{k} \overline{(A^*)_{kj}}\langle b_i, b_k\rangle = \overline{(A^*)_{ij}} = A_{ji}\end{aligned}$$

Therefore, for all $i, j \in \{1, \dots, n\}$ we have $\langle f(b_i), b_j\rangle = \langle b_i, g(b_j)\rangle$. It follows that for all $u, v \in V$ we have $\langle f(u), v\rangle = \langle u, g(v)\rangle$. □

Remark. As part of the above proof we showed that $[f^*]_{\mathcal{B}} = ([f]_{\mathcal{B}})^*$ for any orthonormal basis $\mathcal{B}$ of a finite dimensional V.

Example 4.16.

1. $f : \mathbb{C}^2 \to \mathbb{C}^2$, $f(x, y) = (x, 0)$ has adjoint $f^* = f$.
2. $f : \mathbb{R}^2 \to \mathbb{R}^2$, given by a rotation has adjoint $f^* = f^{-1}$
3. Let W be as in Example 4.13. The linear transformation $f : W \to W$ given by $f(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$ has adjoint given by $f^*(x_1, x_2, \dots) = (x_2, x_3, \dots)$.
4. Let $V = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ is infinitely differentiable and } \forall n \in \mathbb{Z}\ f(x + n) = f(x)\}$ with inner product $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$. Let $\Delta : V \to V$ be given by $\Delta(f) = \frac{d^2 f}{dt^2}$. Then $\Delta^* = \Delta$.

Lemma 4.17: Properties of the adjoint

Let V be an inner product space and let $f, g : V \to V$ be two linear transformations and $k \in K$. Then

1. $(f + g)^* = f^* + g^*$
2. $(kf)^* = \overline{k} f^*$
3. $(f \circ g)^* = g^* \circ f^*$
4. $(f^*)^* = f$

Exercise 125. Write out a proof of the above lemma. Note that there is no assumption that V be finite dimensional, merely that f^* and g^* exist.

Definition 4.18. Let $f : V \to V$ be a linear transformation on an inner product space. We say that f is:

1. **self-adjoint** if $f^* = f$ (also called **symmetric** if $K = \mathbb{R}$ or **hermitian** if $K = \mathbb{C}$)
2. **isometric** if $f^* \circ f = \mathrm{Id}_V$ (also called **orthogonal** if $K = \mathbb{R}$ or **unitary** if $K = \mathbb{C}$)
3. **normal** if $f^* \circ f = f \circ f^*$

Remark. It follows from the definitions that

1. If f is self-adjoint, then it is normal.
2. If V is finite dimensional and f is an isometry, then $f^* = f^{-1}$ and f is normal.

Example 4.19. Considering the linear transformations in Example 4.16 we see that:

1. f is self-adjoint and therefore normal,
2. f is an isometry and therefore normal (since $\mathbb{R}^2$ is finite dimensional),
3. f is an isometry since $f^* \circ f = \mathrm{Id}$, but f is not invertible and not normal,
4. Δ is self-adjoint and therefore normal.

Lemma 4.20

Let $f : V \to V$ be a linear transformation on an inner product space. The following are equivalent:

1. $f^* \circ f = \mathrm{Id}_V$ (i.e., f is an isometry as defined above)
2. $\forall u, v \in V$, $\langle f(u), f(v)\rangle = \langle u, v\rangle$
3. $\forall v \in V$, $\|f(v)\| = \|v\|$

Proof. If the first holds, then we have $\langle f(u), f(v)\rangle = \langle u, f^* \circ f(v)\rangle = \langle u, \mathrm{Id}_V(v)\rangle = \langle u, v\rangle$, so the second holds. If the second holds, then we have $\|f(v)\|^2 = \langle f(v), f(v)\rangle = \langle v, v\rangle = \|v\|^2$, so the third holds.
Now suppose that the third condition holds and define $g = f^* \circ f - \mathrm{Id}_V$. We will show that $g = 0$. From Lemma 4.17 we have that g is self-adjoint: $g^* = (f^* \circ f) - \mathrm{Id}_V^* = f^* \circ (f^*)^* - \mathrm{Id}_V = f^* \circ f - \mathrm{Id}_V = g$. For any $u, v \in V$ we have

$$\langle g(v), v\rangle = \langle f^* \circ f(v) - v, v\rangle = \langle f^* \circ f(v), v\rangle - \langle v, v\rangle = \langle f(v), f(v)\rangle - \langle v, v\rangle = \|f(v)\|^2 - \|v\|^2 = 0$$

and therefore

$$0 = \langle g(u+v), u+v\rangle = \langle g(u), v\rangle + \langle g(v), u\rangle = \langle g(u), v\rangle + \langle v, g(u)\rangle = \langle g(u), v\rangle + \overline{\langle g(u), v\rangle}$$

Letting $v = g(u)$ we obtain

$$0 = \langle g(u), g(u)\rangle + \overline{\langle g(u), g(u)\rangle} = 2\langle g(u), g(u)\rangle$$

Therefore $g(u) = 0$ for all $u \in V$ and hence $g = 0$. □

Lemma 4.21

Let $f : V \to V$ be a linear transformation on a inner product space and $W \leqslant V$ a subspace. If W is f-invariant, then $W^\perp$ is f^*-invariant.

Proof. Let $u \in W$ and $v \in W^\perp$. Then $\langle u, f^*(v)\rangle = \langle f(u), v\rangle = 0$ since $f(u) \in W$ and $v \in W^\perp$. □

4.1 Exercises

Exercise 126. If A is a transition matrix between orthonormal bases, show that A is isometric (i.e., $A^*A = I$).

Exercise 127. Suppose that f is a linear transformation on an inner product space V. Prove the following.

(a) If f is self-adjoint, then all eigenvalues of f are real.

(b) If f is isometric, then all eigenvalues of f have absolute value 1.

Exercise 128. 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 kernel 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.

Exercise 129. Consider the inner product space $\mathcal{P}(\mathbb{R})$ having inner product $\langle p(x), q(x)\rangle = \int_0^1 p(x)q(x)\,dx$. Show that the linear transformation $\delta : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ given by differentiation has no adjoint. (Hint: Try to find what $\delta^*(1)$ should be.)

Exercise 130. Show that a triangular matrix which is self-adjoint or unitary is diagonal.

Exercise 131. 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 $\langle f(v), w\rangle = \langle v, w_1\rangle$ for all $v \in V$. (Hint: First show that it will be enough to consider only those v that lie in some fixed orthonormal basis of V.)

Exercise 132. Let g be a self-adjoint linear transformation on a finite dimensional inner product space V. Suppose that $\langle g(v), v\rangle = 0$ for all $v \in V$.

(a) Show that $\langle g(u), w\rangle + \langle g(w), u\rangle = 0$ for all $u, w \in V$.

(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 $\langle g(u), w\rangle$ is imaginary for all $u, w \in V$.

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

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

Exercise 133. 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 $(f|_W)^* = (f^*)|_W$.

Exercise 134. 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.

Exercise 135. 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 $\left[\begin{smallmatrix} \cos\theta & -\sin\theta \\ \epsilon\sin\theta & \epsilon\cos\theta \end{smallmatrix}\right]$ where $\epsilon = \pm 1$.

5 Spectral theorem

We now come to the question of when a linear transformation can be diagonalised. We have seen necessary and sufficient conditions in terms of the minimal polynomial of the transformation. The spectral theorem gives a sufficient condition for diagonalisability (without reference to the minimal polynomial).

Theorem 4.22: Spectral theorem for normal linear transformations

Let V be a finite dimensional, complex inner product space vector space and let $f : V \to V$ be a linear transformation. If f is normal, then there exists an orthonormal basis $\mathcal{B}$ for V such that $[f]_{\mathcal{B}}$ is diagonal.

Proof. We use (strong) induction on $n = \dim(V)$. If $n = 1$, then the statement is trivially true. Assume now that $n > 1$ and that the statement holds for all cases in which the dimension is less than n. Let $\lambda \in \mathbb{C}$ be an eigenvalue of f and V_λ the corresponding eigenspace. By Proposition 4.12 we have $V = V_\lambda \oplus V_\lambda^\perp$. Note that $\dim(V_\lambda) < \dim(V)$ and $\dim(V_\lambda^\perp) < \dim(V)$. We will show that both V_λ and $V_\lambda^\perp$ are f-invariant, and then apply Lemma 2.16. That V_λ is f-invariant is clear (see Exercise 42). To show that $V_\lambda^\perp$ is f-invariant, we note first that V_λ is f^*-invariant since (using that f is normal):

$$u \in V_\lambda \implies f(f^*(u)) = f^*(f(u)) = f^*(\lambda u) = \lambda f^*(u) \implies f^*(u) \in V_\lambda$$

That $V_\lambda^\perp$ is f-invariant then follows from Lemma 4.21 since $(f^*)^* = f$. Let $f_1 : V_\lambda \to V_\lambda$ and $f_2 : V_\lambda^\perp \to V_\lambda^\perp$ be the restrictions of f to V_λ and $V_\lambda^\perp$ respectively. By the induction hypothesis, there exist orthonormal bases $\mathcal{B}_1$ and $\mathcal{B}_2$ for V_λ and $V_\lambda^\perp$ respectively, such that $[f_i]_{\mathcal{B}_i}$ is diagonal. By Lemma 2.16 $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ is a basis for V and $[f]_{\mathcal{B}} = [f_1]_{\mathcal{B}_1} \oplus [f_2]_{\mathcal{B}_2}$. In particular, $[f]_{\mathcal{B}}$ is diagonal. □

Theorem 4.23: Spectral theorem for normal matrices

Let $A \in M_n(\mathbb{C})$ be such that $AA^* = A^*A$. There exists a matrix $U \in M_n(\mathbb{C})$ such that $U^*U = I$ (i.e., U is unitary) and U^*AU is diagonal.

Proof. Define $f : M_{n\times 1} \to M_{n\times 1}$ by $f(X) = AX$ and apply Theorem 4.22. Letting U be the matrix whose columns are the elements of $\mathcal{B}$, we have $[f]_{\mathcal{B}} = U^{-1}AU$. That $U^{-1} = U^*$ follows from the fact that $\mathcal{B}$ is an orthonormal basis. □

Remark. The columns of U form an orthonormal basis and the diagonal entries or U^*AU are exactly the eigenvalues of A.

Example 4.24. Let $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Then $A^*A = AA^*$. By the spectral theorem there is unitary matrix $U \in M_2(\mathbb{C})$ such that U^*AU is diagonal. To find such a U we calculate an orthonormal basis of eigenvectors. The eigenvalues of the matrix A are $1-i, 1+i$. An orthonormal basis for the $(1-i)$-eigenspace is $\{\left[\begin{smallmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{smallmatrix}\right]\}$. An orthonormal basis for the $(1+i)$-eigenspace is $\{\left[\begin{smallmatrix} i/\sqrt{2} \\ 1/\sqrt{2} \end{smallmatrix}\right]\}$. So we can take

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \quad D = \begin{bmatrix} 1-i & 0 \\ 0 & 1+i \end{bmatrix}$$

Note however that the matrix A is *not* diagonalisable over $\mathbb{R}$. That is, there does not exist an invertible matrix $P \in M_2(\mathbb{R})$ such that $P^{-1}AP$ is diagonal.

For real inner product spaces we have the following.

Theorem 4.25: Spectral theorem for symmetric linear transformations

Let V be a finite dimensional real inner product space and let $f : V \to V$ be a self-adjoint linear transformation. Then there exists an orthonormal basis $\mathcal{B}$ of V such that $[f]_{\mathcal{B}}$ is diagonal.

Outline of proof. We use induction on $n = \dim(V)$. If $n = 1$, the result holds trivially.

Since f is self-adjoint, all eigenvalues are real (Exercise 127). Let $\lambda \in \mathbb{R}$ be an eigenvalue of f and let $u \in V$ be such that $f(u) = \lambda u$. Let $W = \operatorname{span}(u)$. Then $V = W \oplus W^\perp$ (Proposition 4.12) and W and $W^\perp$ are both f-invariant (Lemma 4.21). By the induction hypothesis, there exists an orthonormal basis $\mathcal{C} = \{c_1, \dots, c_{n-1}\}$ for $W^\perp$ such that $D = [f|_{W^\perp}]_{\mathcal{C}}$ is diagonal. Letting $\mathcal{B} = \{c_1, \dots, c_{n-1}, u/\|u\|\}$ we have that $[f]_{\mathcal{B}} = D \oplus [\lambda]$ (Lemma 2.16). It remains to show that $\mathcal{B}$ is orthonormal. This follows from the fact that both $\mathcal{C}$ and $\{u/\|u\|\}$ are orthonormal and that $\langle c_i, u\rangle = 0$ for all i. □

5.1 Exercises

Exercise 136. 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).

Exercise 137. Show that a linear transformation $f : V \to V$ on a complex inner product space V is normal if and only if $\langle f(u), f(v)\rangle = \langle f^*(u), f^*(v)\rangle$ for all $u, v \in V$.

Exercise 138. (a) Show that every normal matrix A has a square root; that is, a matrix B so that $B^2 = A$.

(b) Must every complex square matrix have a square root?

Exercise 139. 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 or not the following matrices are unitarily equivalent.

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

(b) $\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ and $\begin{bmatrix} -1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix}$

Exercise 140. Are f and f^* always unitarily equivalent?

Exercise 141. 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.

Exercise 142. If f is a normal linear transformation on a finite dimensional complex inner product space show that $f^* = p(f)$ for some polynomial p.

Exercise 143. 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.

Exercise 144. 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 $\operatorname{diag}(A_1, \dots, A_m)$ 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 $\operatorname{diag}(B_1, \dots, B_m)$ 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.

Chapter 5

Groups II

1 Group actions

Definition 5.1. Let G be a group and X a set. A left **action** of of G on X is a function $G \times X \to X$ (with the image of (g, x) being denoted $g \cdot x$) satisfying

1) $\forall x \in X, \quad e_G \cdot x = x$
2) $\forall x \in X \; \forall g, h \in G, \quad (gh) \cdot x = g \cdot (h \cdot x)$

We also say that G acts on X and denote this by $G \curvearrowright X$.

Example 5.2. 1. $S_n \curvearrowright \{1, 2, \ldots, n\}$, for example $(132) \cdot 3 = 2$

2. D_n acts on the vertices of a regular n-gon
3. $GL(n, K)$ acts on K^n (having fixed a basis for K^n)
4. $GL(n, K)$ acts on $\{W \mid W \leqslant K^n\}$ (having fixed a basis for K^n)
5. $\mathbb{Z}/2\mathbb{Z} \curvearrowright \mathbb{C}$, $[0] \cdot z = z$, $[1] \cdot z = \overline{z}$

Example 5.3. Here are two important examples in which a group acts on itself.

1. $G \curvearrowright G$ by left multiplication: $g \cdot x = gx$
2. $G \curvearrowright G$ by conjugation: $g \cdot x = gxg^{-1}$

Remark. Let S_X denote the group of all bijections from X to X (with operation given by function composition). An action $G \curvearrowright X$ corresponds to a homomorphism $G \to S_X$ in the following sense.

Exercise 145. (a) Suppose that a group G acts on a set X.

(i) Let $g \in G$. Show that the map $\varphi_g : X \to X$, $\varphi_g(x) = g \cdot x$ is a bijection.
(ii) Show that the map $\Phi : G \to S_X$ given by $\Phi(g) = \varphi_g$ is a homomorphism.

(b) Suppose that G is a group, X a set and that $\Psi : G \to S_X$ is a homomorphism. Show that there is an action of G on X defined by $g \cdot x = \Psi(g)(x)$.

Theorem 5.4: Cayley's Theorem

Let G be a finite group and $n = |G|$. Then G is isomorphic to a subgroup of S_n.

Proof. Consider the action of G on itself by left multiplication. From Exercise 145 there is a corresponding homomorphism $\Phi : G \to S_n$. The homomorphism Φ is injective since

$$\Phi(g) = e \implies g \cdot x = x \quad \text{(for all } x \in G) \implies ge_G = e_G \implies g = e_G$$

Because Φ is injective, $G \cong \mathrm{im}(\Pi)$. □

2 Orbits and stabilisers

Definition 5.5. Suppose that $G \curvearrowright X$ and let $x \in X$.

1) The **orbit** of x is the set $O(x) = \{g \cdot x \mid g \in G\} \subseteq X$ (sometimes denoted $G \cdot x$)
2) The **stabiliser** of x is $\mathrm{Stab}(x) = \{g \in G \mid g \cdot x = x\}$
3) $x \in X$ is a **fixed point** if $\mathrm{Stab}(x) = G$
4) The action is **transitive** if $\forall x, y \in X\ \exists g \in G,\ g \cdot x = y$
(i.e., there is only one orbit)

Exercise 146. Show that $\mathrm{Stab}(x)$ is a subgroup of G.

Example 5.6. 1. $S_3 \curvearrowright \{1,2,3\}$, $\mathrm{Stab}(2) = \{e, (13)\}$, $O(2) = \{1,2,3\}$, the action is transitive

2. $G = \langle(123)\rangle \leqslant S_5$, $X = \{1,2,3,4,5\}$, $\mathrm{Stab}(2) = \{e\}$, $O(2) = \{1,2,3\}$, $\mathrm{Stab}(5) = G$, $O(5) = \{5\}$
3. $X = \{1,2,3,4\}$ (identified with the vertices of a square), $G = D_4$, $\mathrm{Stab}(1) = \{e, rs\}$, $O(1) = \{1,2,3,4\}$ (using our standing notational conventions for the dihedral groups as in section 3.6.)
4. $G \curvearrowright G$ by left multiplication, $\mathrm{Stab}(g) = \{e\}$, $O(g) = G$
5. $G \curvearrowright G$ by conjugation, $\mathrm{Stab}(g)$ is called the **centraliser** of g

$$C_G(g) = \{h \in G \mid hg = gh\}$$

$O(g) = \{hgh^{-1} \mid h \in G\}$ is called the **conjugacy class** of g.

Lemma 5.7

Let G be a group acting on a set X. The orbits partition X.

Proof. We need to show that every element of X is contained in exactly one orbit. Clearly $x = e \cdot x \in O(x)$. We need to show that if $O(x) \cap O(y) \neq \emptyset$, then $O(x) = O(y)$. Let $z \in O(x) \cap O(y)$. Then there are $g, h \in G$ such that $z = g \cdot x$ and $z = h \cdot y$. Then $x = g^{-1} \cdot z$, $y = h^{-1} \cdot z$, and

$$\begin{aligned} w \in O(x) &\implies w = k \cdot x \quad \text{for some } k \in G \\ &\implies w = k \cdot (g^{-1} \cdot z) = (kg^{-1}) \cdot z = (kg^{-1}) \cdot (h \cdot y) = (kg^{-1}h) \cdot y \\ &\implies w \in O(y) \end{aligned}$$

So $O(x) \subseteq O(y)$. Similarly $O(y) \subseteq O(x)$. □

Exercise 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) $G = \langle(1234),(13)\rangle$ (which is isomorphic to D_4)

Exercise 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.

Exercise 149. Find the conjugacy classes in the quaternion group described in Exercise 112.

Definition 5.8. Let G be a group. Two elements $a, b \in G$ are called **conjugates** of one another if $\exists g \in G$, $gag^{-1} = b$. This is equivlalent to the condition that a and b lie in the same orbit of the conjugacy action of G on itself.

Exercise 150. Find the conjugates of the following:

(a) (123) in S_3

(b) (123) in S_4

(c) (1234) in S_4

(d) (1234) in S_n where $n \geqslant 4$

(e) $(12\dots m)$ in S_n where $n \geqslant m$

Exercise 151. Let $\tau \in S_n$. 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.

Exercise 152. 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.

Exercise 153. 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) $\begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}$

3 The orbit-stabiliser relation and applications

Theorem 5.9: The orbit-stabiliser relation

Let G be a group and $G \curvearrowright X$ an action on a set X. Denote by $G/\operatorname{Stab}(x)$ the set of left cosets of $\operatorname{Stab}(x)$. Then, for all $x \in X$ the map $G/\operatorname{Stab}(x) \to O(x)$ given by $g\operatorname{Stab}(x) \mapsto g \cdot x$ is a bijection. If G is finite, then

$$|G| = |O(x)|\,|\operatorname{Stab}(x)|$$

Proof. Denote the map by Φ. We first show that the map is well-defined.

$$g\operatorname{Stab}(x) = h\operatorname{Stab}(x) \implies g^{-1}h \in \operatorname{Stab}(x) \implies (g^{-1}h)\cdot x = x \implies h\cdot x = g\cdot x$$

Now that the map is injective.

$$\begin{aligned}\Phi(g\operatorname{Stab}(x)) = \Phi(h\operatorname{Stab}(x)) &\implies g\cdot x = h\cdot x \implies g^{-1}\cdot(g\cdot x) = g^{-1}\cdot(h\cdot x) \implies (g^{-1}g)\cdot x = (g^{-1}h)\cdot x\\ &\implies x = (g^{-1}h)\cdot x \implies g^{-1}h \in \operatorname{Stab}(x)\\ &\implies g\operatorname{Stab}(x) = h\operatorname{Stab}(x)\end{aligned}$$

And surjective:

$$y \in O(x) \implies y = g\cdot x \quad (\text{for some } g \in G) \implies y = \Phi(g\operatorname{Stab}(x))$$

If G is finite, then we have:

$$\begin{aligned}|G| &= [G:\operatorname{Stab}(x)]\,|\operatorname{Stab}(x)| && \text{(by Lagrange's theorem)}\\ &= |O(x)|\,|\operatorname{Stab}(x)| && \text{(since } \Phi \text{ is a bijection)}\end{aligned}$$

□

We'll now look at some consequences of the orbit-stabiliser relation. The first are contained in the following exercises.

Exercise 154. Let G be the subgroup of S_{15} given by

$$G = \langle (1,12)(3,10)(5,13)(11,15), (2,7)(4,14)(6,10)(9,13), (4,8)(6,10)(7,12)(9,11)\rangle$$

Find the orbits in $X = \{1,\dots,15\}$ under the action of G. Deduce that the order of G is a multiple of 60.

Exercise 155. 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?

The next result is a result of applying the orbit-stabiliser relation to the conjugacy action of a group on itself. First a definition.

Definition 5.10. Let G be a group. The **centre** of G, denoted $Z(G)$, is the set of elements that commute with all elements of G. That is, $Z(G) = \{g \in G \mid \forall h \in G,\ gh = hg\}$.

Remark. The centre of G consists of all fixed points of the action of G on itself by conjugation.

Example 5.11.

1. $Z(\mathbb{Z}) = \mathbb{Z}$
2. $Z(D_4) = \{e, r^2\}$
3. $Z(S_3) = \{e\}$

Exercise 156. Show that $Z(G)$ is a normal subgroup of G.

Exercise 157. Suppose that G is a group with centre Z and is such that G/Z is a cyclic group. Show that there exists an element $h \in G$ such that every element of G can be written in the form $g = h^i z$ with $i \in \mathbb{Z}$ and $z \in Z$. Deduce that G is commutative.

Theorem 5.12

Let G be a group of size p^n where $p \in \mathbb{N}$ is prime and $n \in \mathbb{N}$. Then $|Z(G)| \geqslant p$.

Proof. Consider G acting on itself by conjugation. The orbits partition G and $Z(G)$ is the union of all orbits having size 1. Therefore, G is a disjoint union

$$G = Z(G) \cup C_1 \cup C_2 \ldots C_k \qquad (*)$$

where the C_i are the orbits having size at least 2. By the orbit-stabiliser relation we have that for all i, $|C_i| \mid |G|$. Therefore $p \mid |C_i|$ for all i, and hence $p \mid |Z(G)|$ by $(*)$. □

Theorem 5.13

Let G be a group of size p^n where $p \in \mathbb{N}$ is prime and $n \in \mathbb{N}$. Suppose that G acts on a finite set X. If p does not divide $|X|$, then the action has a fixed point.

Proof. Denote the orbits of the action as O_1, O_2,. . . , O_k. By the orbit-stabiliser relation $|O_i| \mid |G| = p^n$. Therefore $\forall i$, $|O_i| = 1$ or $p \mid |O_i|$. Suppose, for a contradiction, that there are no orbits of size 1. Then we would have $p \mid |X|$ since $|X| = |O_1| + \cdots + |O_k|$. □

Example 5.14. Let $p \in \mathbb{N}$ be a prime. Recall that $\mathbb{F}_p$ denotes the filed with p elements. Let $G \leqslant GL(3, \mathbb{F}_p)$ be given by

$$G = \left\{ \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \mid a, b, c \in \mathbb{F}_p \right\}$$

Note the $|G| = p^3$. Let X be the set of all 1-dimensional subspaces of $\mathbb{F}_p^3$. Then G acts on X (since $GL(3, \mathbb{F}_p)$ does). Explicitly, after fixing a basis $\mathcal{B}$ for $\mathbb{F}_p^3$ we identify $\mathbb{F}_p^3$ with $M_{3\times1}(\mathbb{F}_p)$ and define $g \cdot \operatorname{span}(u) = \operatorname{span}(gu)$. The number of 1-dimensional subspaces is given by

$$|X| = \frac{p^3 - 1}{p - 1} = p^2 + p + 1$$

Since p does not divide $p^2 + p + 1$ we conclude (from the above theorem) that there is a 1-dimensional subspace that is fixed by G.

Theorem 5.15

Let $p \in \mathbb{N}$ be prime and G a group. If $|G| = p^2$, then either $G \cong \mathbb{Z}/p^2\mathbb{Z}$ or $G \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.

Remark. As a consequence, if $|G| = p^2$ then G is abelian.

Proof. Suppose that G is not cyclic. We need to show that $G \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. By Theorem 5.12, $|Z(G)| > 1$. Let $g \in Z(G) \setminus \{e\}$. Since G is not cyclic and $g \neq e$, we have $|g| = p$. Let $H = \langle g \rangle$. Then $H \lhd G$ since $g \in Z(G)$. By Lagrange's Theorem, $|G/H| = |G|/|H| = p$. Hence G/H is cyclic. Let $x \in G$ be such that xH generates G/H. Then

$$G/H = \{eH, xH, x^2H, \dots, x^{p-1}H\}$$

It follows that $\langle x, g \rangle = G$.

Define a map $\varphi : \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \to G$ by $\varphi([a]_p, [b]_p) = x^a g^b$. Since both x and g have order p, this map is well-defined. It is a homomorphism since

$$\begin{aligned}
\varphi(([a_1]_p, [b_1]_p) + ([a_2]_p, [b_2]_p)) &= \varphi(([a_1 + a_2]_p, [b_1 + b_2]_p)) \\
&= x^{a+1+a_2} g^{b_1+b_2} = x^{a_1} x^{a_2} g^{b_1} g^{b_2} \\
&= x^{a_1} g^{b_1} x^{a_2} g^{b_2} && \text{(since } xg = gx\text{)} \\
&= \varphi([a_1]_p, [b_1]_p)\varphi([a_2]_p, [b_2]_p)
\end{aligned}$$

Since $x, g \in \operatorname{im}(\varphi)$ and $\langle x, g \rangle = G$, the homomorphism is surjective, It is therefore also injective since $|G| = |\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}| = p^2$. □

Exercise 158. Describe the finite groups having exactly one or exactly two or exactly three conjugacy classes.

4 Cauchy's Theorem

We know from Lagrange's theorem that if $g \in G$, then $|g|$ divides $|G|$. The converse is in general false, that is, $m \mid |G|$ does not imply that there exists an element in G of order m. But it does hold for prime divisors.

Theorem 5.16: Cauchy's theorem

Let G be a finite group and $p \in \mathbb{N}$ a prime. If p divides |G| , then there exists $g \in G$ with $|g| = p$.

Proof. Let $X = \{(x_1, \dots, x_p) \in G^p \mid x_1 x_2 \dots x_p = e\}$. Note that $|X| = |G|^{p-1}$ and therefore $p \mid |G|$. The group $\mathbb{Z}/p\mathbb{Z}$ acts on X by cyclic permutation, that is:

$$[1]_p \cdot (x_1, \dots, x_p) = (x_p, x_1, \dots, x_{p-1}) \qquad [2]_p \cdot (x_1, \dots, x_p) = (x_{p-1}, x_p, x_1, \dots, x_{p-2}) \quad \text{etc}$$

Note that a fixed point of this action is of the form $(x, x, \dots, x)$ with $x^p = 1$. One such fixed point is $(e, \dots, e)$. Our goal is to show that there exists at least one other orbit of size 1. By the orbit stabiliser relation, all orbits have size that divides $|\mathbb{Z}/p\mathbb{Z}| = p$. If there were only one orbit of size 1, we would have $|X| = 1 + kp$ for some $k \in \mathbb{N}$ which contradicts the fact that $p \mid |X|$. □

Exercise 159. Show that if p is a prime number, then any group of order $2p$ must have a subgroup of order p and that this subgroup must be normal.

Exercise 160. Let $p \in \mathbb{N}$ be prime. Show that, up to isomorphism, there are exactly two groups of order $2p$.

5 Burnside orbit counting lemma

Definition 5.17. Given an action $G \curvearrowright X$ and an element $g \in G$, the **fixed point set** of g is

$$X^g = \{x \in X \mid g \cdot x = x\}$$

Lemma 5.18: Burnside counting lemma

Let G be a finite group acting on a finite set X. Let N be the number of orbits of the action. Then

$$N = \frac{1}{|G|} \sum_{g \in G} |X^g|$$

Proof. Consider the set $S = \{(g,x) \in G \times X \mid g \cdot x = x\}$. We will count the elements on S in two ways. Firstly,

$$|S| = \sum_{g \in G} |\{x \in X \mid g \cdot x = x\}| = \sum_{g \in G} |X^g| \tag{1}$$

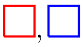
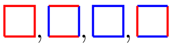
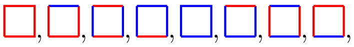
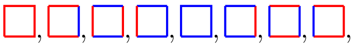
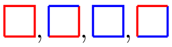
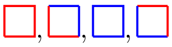
For the second count denote the orbits of the action by $O_1, \dots, O_N$. We have

$$\begin{aligned} |S| &= \sum_{x \in X} |\{g \in G \mid g \cdot x = x\}| = \sum_{x \in X} |\operatorname{Stab}(x)| \\ &= \sum_{i=1}^{N} \sum_{x \in O_i} |\operatorname{Stab}(x)| && \text{(since the orbits partition } X\text{)} \\ &= \sum_{i=1}^{N} \sum_{x \in O_i} \frac{|G|}{|O_i|} && \text{(by the orbit-stabiliser relation)} \\ &= |G| \sum_{i=1}^{N} \sum_{x \in O_i} \frac{1}{|O_i|} = |G| \sum_{i=1}^{N} 1 = N|G| && (2) \end{aligned}$$

Equating (1) and (2) gives the desired result. □

Example 5.19. How many ways are there to colour the sides of a square using two colours? There are a total of 2^4 different colourings, but some are equivalent in the sense that one can be obtained from the other by applying a reflection or a rotation.

More precisely, if we let X denote the set of all colourings, then $|X| = 16$ and D_4 acts on X. The number of "different" (i.e., non-equivalent) colourings is given by the number of orbits. To find the number of orbits, we can apply the Burnside Lemma. For that we need to consider the set X^g.

$g \in D_4$	X^g	$\lvert X^g \rvert$
e	all colourings	16
r, r^3		2
r^2		4
s		8
$r^2 s$		8
rs		4
$r^3 s$		4

The number of colourings (up to symmetry) is given by the number of orbits, which by Burnside's lemma is:

$$\begin{aligned} \frac{1}{|D_4|} \sum_{g \in D_4} |X^g| &= \frac{1}{8}(16 + 2 + 2 + 4 + 8 + 8 + 4 + 4) \\ &= \frac{48}{8} = 6 \end{aligned}$$

Up to symmetry, there are six different colourings of the square.

Exercise 161. There are 70 (which is $\binom{8}{4}$) ways to colour the edges of an octagon so that four edges are green and four edges are red. Let X be the set of such coloured octagons (so $|X| = 70$). The group D_8 acts on X and two colourings are considered to be equivalent if they are in the same orbit. Use Burnside's orbit counting lemma to find the number of equivalence classes (i.e., orbits).

6 Sylow Theorems

The Sylow theorems are an important tool for understanding finite groups. We know from Cauchy's theorem that if the order of a group G is divisible by a prime p, then G contains a subgroup of order p. The first Sylow theorem generalises this to subgroups of size that is a power of p.

Theorem 5.20: First Sylow theorem

Let G be a finite group, $p \in \mathbb{N}$ a prime and $s \in \mathbb{N}$. If p^s divides $|G|$, then G has a subgroup of size p^s.

Proof. We proceed by induction on $|G|$. If $|G| < p$, then there is nothing to prove, so we assume that $|G| > p$. The inductive hypothesis is that for all groups H with $|H| < |G|$ we have that if $p^t \mid |H|$ (for some $t \in \mathbb{N}$), then there exists a subgroup of H having size p^t. We split into two cases.

Case 1: Suppose first that G contains a proper subgroup $H \subsetneq G$ such that $p \nmid [G : H]$. Since $p^s \mid |G| = [G : H]|H|$ it follows that $p^s \mid |H|$. By the induction hypothesis H (hence G) contains a subgroup $K \leqslant H$ with $|K| = p^s$.

Case 2: Suppose that every proper subgroup of G has index divisible by p. We first show that $|Z(G)|$ is divisible by p. Considering the action of G on itself by conjugation we have

$$|G| = |Z(G)| + |C_1| + |C_2| + \cdots + |C_k| \qquad (*)$$

where the C_i are the conjugacy classes of size at least 2. For each i, fix some $g_i \in C_i$. From the orbit-stabiliser relation and Lagrange's theorem we have that

$$|C_i| = |G|/|C_G(g_i)| = [G : C_G(g_i)]$$

Since this index is at least 2, $C_G(g_i)$ is a proper subgroup of G and therefore $[G : C_G(g_i)]$ is divisible by p. Therefore, from $(*)$, $|Z(G)|$ is divisible by p.

By Cauchy's theorem there is an element $z \in Z(G)$ with $|z| = p$. Let $N = \langle z\rangle \leqslant Z(G)$. Then $|N| = p$ and N is a normal subgroup of G. Let $H = G/N$. Then $|H| = |G|/p$ and therefore $|H| < |G|$ and $p^{s-1} \mid |H|$. By the inductive hypothesis there is a subgroup $K \leqslant H$ with $|K| = p^{s-1}$. Denote by π the natural projection homomorphism $\pi : G \to H = G/N$, $\pi(g) = gN$. Let $L = \pi^{-1}(K) = \{g \in G \mid \pi(g) \in K\}$. Then L is a subgroup of G and has order p^s.

Exercise 162. Use the first isomorphism theorem to prove that L has size p^s.

□

Definition 5.21. A group of order p^s for some prime p and some $s \in \mathbb{N}$ is called a ***p*-group**. A **Sylow *p*-subgroup** of a finite group G is a subgroup $H \leqslant G$ such that

1) H is a p-group

2) $[G : H]$ is not divisible by p

Remark. 1. The condition that $[G : H]$ be not divisible by p is equivalent to the condition that if $|H| = p^s$ then s is the largest element in $\mathbb{N}$ for which $p^s \mid |G|$.

2. The first Sylow theorem shows that p-Sylow subgroups exist for all primes p that divide $|G|$.

Example 5.22. For the group $G = D_6$ we have $|G| = 12 = 2^2 \times 3$. The subgroup $H = \langle s, r^3 s\rangle \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ is a Sylow 2-subgroup.

Theorem 5.23: Second Sylow theorem

Let G be a finite group. Any two Sylow p-subgroups of G are conjugate.

Theorem 5.24: Third Sylow theorem

Let $p \in \mathbb{N}$ be prime and let G be a finite group such that $p \mid |G|$. Denote by n_p the number of Sylow p-subgroups of G. Then

1) $n_p \mid |G|$

2) $n_p \equiv 1 \pmod{p}$

Theorem 5.25: Fourth Sylow theorem

Let G be a finite group and $H \leqslant G$ a subgroup. If H is a p-group, then H is contained in a Sylow p-subgroup.

6.1 Groups of size 12

As an application we consider the possibilities for a group G of size 12.

For $p = 2$ a Sylow 2-subgroup H has size 4 and therefore $H \cong \mathbb{Z}/4\mathbb{Z}$ or $H \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. The number n_2 of Sylow 2-subgroups divides 12 and is odd. Therefore $n_2 \in \{1, 3\}$.

For $p = 3$, a Sylow 3-subgroup K has size 3 and therefore $K \cong \mathbb{Z}/3\mathbb{Z}$. The number n_3 of Sylow 3-subgroups divides 12 and $n_3 \equiv 1 \pmod 3$. Therefore $n_3 \in \{1, 4\}$.

We claim that $n_2 = 1$ or $n_3 = 1$ (or both). To see this, suppose that $n_3 = 4$. Then there are 8 elements of order 3 in G. If H is a Sylow 2-subgroup, then its 4 elements make up all the remaining elements of G. Therefore $n_2 = 1$.

We now consider the three possible cases.

Case 1: $n_2 = n_3 = 1$
Let H be the Sylow 2-subgroup and K the Sylow 3-subgroup. Both are normal in G since $n_2 = n_3 = 1$. It follows that for every $h \in H$ and $k \in K$, $hk = kh$. Therefore the map $H \times K \to G$, $(h, k) \mapsto hk$ is an isomorphism. Therefore

$$\begin{aligned} & G \cong \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} && (1)\\ \text{or} \quad & G \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} && (2) \end{aligned}$$

Case 2: $n_2 = 1, n_3 = 4$
Since $n_3 > 1$ it follows from the second Sylow theorem that G is not abelian. Let $F_1, F_2, F_3, F_4 \leqslant G$ be the four Sylow 3-subgroups and let $X = \{1, 2, 3, 4\}$. Noting that the conjugate of a Sylow p-subgroup is again a Sylow p-subgroup, we define an action of G on X as follows

$$g \cdot i = j \iff gF_ig^{-1} = F_j$$

By the second Sylow theorem there is only one orbit. Then by the orbit-stabiliser relation we have $|\operatorname{Stab}(i)| = |G|/4 = 3$. Since $F_i \subseteq \operatorname{Stab}(F_i)$ and $|F_i| = 3$, we have that $\operatorname{Stab}(i) = F_i$. The action of G on X corresponds to a homomorphism $\varphi : G \to S_4$ (see Exercise 145) whose kernel is given by

$$\ker(\varphi) = \bigcap_{i=1}^{4} \operatorname{Stab}(F_i) = \bigcap_{i=1}^{4} F_i = \{e\}$$

Therefore $G \cong \operatorname{im}(\varphi)$ and $[S_4 : \operatorname{im}(\varphi)] = |S_4|/|G| = 2$. The only index 2 subgroup in S_4 is

$$\begin{aligned} A_4 = \{&e, (12)(34), (13)(24), (14)(23)\\ &(123), (132), (124), (142), (134), (143), (234), (243)\} \end{aligned}$$

Therefore, in this case we have

$$G \cong A_4 \qquad (3)$$

The Sylow 2-subgroup of A_4 is $\{e, (12)(34), (13)(24), (14)(23)\}$ which is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Case 3: $n_2 = 3, n_3 = 1$
Let $T = \{e, t, t^2\} \leqslant G$ be the Sylow 3-subgroup of G and let F, F', F'' be the Sylow 2-subgroups of G. Let $F = \{e, x, y, z\}$. Since $F \cap T = \{e\}$ we have that

$$G = \{e, x, y, z, t, xt, yt, zt, t^2, xt^2, yt^2, zt^2\}$$

(The point being that the twelve listed elements are distinct.) Since $n_2 > 1$ it follows from the second Sylow theorem that G is not abelian. It follows that there exists an element in F that does not commute with t. There is no loss in generality in assuming that $xt \neq tx$, which implies that $xtx^{-1} = t^2$.

We know that $F \cong F' \cong F''$ by the second Sylow theorem. There are two possibilities; either $F \cong \mathbb{Z}/4\mathbb{Z}$ or $F \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Subcase 3(a): $n_2 = 3, n_3 = 1$, $F \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$
In this case we have that $z = xy$. It is easy to check that $S = \{e, t, t^2, x, xt, xt^2\}$ is a subgroup of G. Since S has size 6 and is not abelian, we know that $S \cong S_3$. Since it has index 2, the subgroup S is normal in G. We therefore have that $yty^{-1} \in \{t, t^2\}$. If $yty^{-1} = t^2$, then we have $ztz^{-1} = xyty^{-1}z^{-1} = xt^2x^{-1} = t$. Swapping the roles of y and z if necessary we can therefore assume that $yty^{-1} = t$. We have that $y \in Z(G)$ and $y \notin S$. Therefore $G \cong \langle y \rangle \times S$. Since $\langle y \rangle \cong \mathbb{Z}/2\mathbb{Z}$, $S \cong S_3$ we conclude that

$$G \cong \mathbb{Z}/2\mathbb{Z} \times S_3 \cong D_6 \qquad (4)$$

Subcase 3(b): $n_2 = 3, n_3 = 1$, $F \cong \mathbb{Z}/4\mathbb{Z}$
In this case we have $F = \{e, x, x^2, x^3\}$ and $G = \{e, x, x^2, x^3, t, xt, x^2t, x^3t, t^2, xt^2, x^2t^2, x^3t^2\}$. Since G is not abelian and T is normal, we have $xtx^{-1} = t^2$. With this knowledge (together with $|x| = 4$ and $|t| = 3$) we know the product of any pair of elements: $(x^mt^i)(x^nt^j) = x^{m+n}t^{j-i}$. It remains to show that there exists a group that has the corresponding multiplication table. Let $\epsilon \in \mathbb{C}$ be a non-real cube root of unity (say $\epsilon = -1/2 + i\sqrt{3}/2$). Let

$A = \left[\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right] \in GL(2, \mathbb{C})$ and $B = \left[\begin{smallmatrix} \epsilon & 0 \\ 0 & \epsilon^2 \end{smallmatrix}\right] \in GL(2, \mathbb{C})$. Then $Dic_3 = \langle A, B \rangle \leqslant GL(2, \mathbb{C})$ has the specified multiplication table:

$$G \cong Dic_3 \tag{5}$$

There is an isomorphism $\varphi : \langle A, B \rangle \to G$ with $\varphi(A) = x$ and $\varphi(B) = t$.

Combining the above cases, we have shown that (up to isomorphism) there are exactly five groups of size 12:

$$\mathbb{Z}/12\mathbb{Z}, \quad \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}, \quad A_4, \quad D_6, \quad Dic_3$$

6.2 Proof of the second and fourth Sylow theorems

Let G be a finite group and $p \in \mathbb{N}$ a prime such that $p \mid |G|$. Suppose that $H \leqslant G$ be a p-group and let $K \leqslant G$ ba a Sylow p-subgroup of G. We want to show that there exists an element $g \in G$ such that $gHg^{-1} \subseteq K$.

Let X be the set of left cosets of K, $X = \{gK \mid g \in G\}$. So $|X| = [G : K]$. Therefore, because K is a Sylow p-subgroup, p does not divide $|X|$. Consider the action of H on X by left multiplication:

$$h \cdot (gK) = (hg)K$$

For any $x \in X$, from the orbit stabiliser relation, we have that $|O(x)| = |H|/|\operatorname{Stab}(x)|$. Therefore, since H is a p-group, if $|O(x)| > 1$, then $|O(x)|$ is divisible by p. Since $|X|$ is the sum of the sizes of the orbits and $|X|$ is not divisible by p, there must be an orbit of size 1. Let $g \in G$ be such that $gK \in X$ is fixed by every element of H. Then we have that for all $h \in H$ $hgK = gK$. Therefore $g^{-1}hg \in K$ for all $h \in H$, and hence $g^{-1}Hg \leqslant K$.

For the second Sylow theorem, note that if H is a Sylow p-subgroup, then $gHg^{-1} = K$ since $|g^{-1}Hg| = |H| = |K|$.

For the fourth Sylow theorem, note that $H \leqslant gKg^{-1}$ and that gKg^{-1} is a Sylow p-subgroup because it has the same size as K.

6.3 Proof of the third Sylow theorem

Let $p \in \mathbb{N}$ be prime and let G be a finite group such that $p \mid |G|$. Denote by n_p the number of Sylow p-subgroups of G. Let X denote the set of Sylow p-subgroups and consider the action of G on X by conjugation. Note that $|X| = n_p$ and that, by the second Sylow theorem, there is a single orbit. Let $H \in X$. By the orbit-stabiliser relation, we have

$$|G| = |X| \times |\operatorname{Stab} H| = n_p \times |\operatorname{Stab} H|$$

Therefore $n_p \mid |G|$.

Now consider H acting on X by conjugation. Note that, from the orbit stabiliser relation and the fact that H is a p-group, all orbits of this action have size that is a power of p. It is clear that (since H is a subgroup of G) that one of the orbits of this action is $\{H\}$. We claim that all other orbits have size strictly larger than 1. Since the orbits partition X, we would then have that

$$n_p = |X| = 1 + |C_1| + \cdots + |C_k| \equiv 1 \pmod{p}$$

All that remains is to show that $\{H\}$ is the only orbit (of H acting on X by conjugation) having size 1. Suppose that $\{K\}$ is an orbit of size 1. Then $hKh^{-1} = K$ for all $h \in H$. Let $N = \{g \in G \mid gKg^{-1} = K\}$. Then $N \leqslant G$ and $H, K \leqslant N$. Since $[G : H] = [G : N] \times [N : H]$, p does not divide $[N : H]$ and therefore H is a Sylow p-subgroup of N. Similarly, K is a Sylow p-subgroup of N. By the second Sylow theorem (applied to N), there exists an element $n \in N$ such that $nKn^{-1} = H$. However, from the definition of N, we know that $nKn^{-1} = K$. Therefore $K = H$.

6.4 Exercises

Exercise 163. Let G be a non-trivial finite group. Prove that G is a p-group if and only if every element of G has order a power of p.

Exercise 164. (a) Show that if H is a Sylow p-subgroup of G and $g \in G$, then gHg^{-1} is also a Sylow p-subgroup.

(b) Show that if G has only one Sylow p-subgroup H, then H is normal.

Exercise 165. Let G be a group with $|G| = pq$ where $p, q \in \mathbb{N}$ are primes and $p < q$. Show that G has exactly one subgroup of order q.

Exercise 166. Let G be a group of size $255 = 3 \times 5 \times 17$. Show that the Sylow 17-subgroup is normal in G.

Appendices

Appendix A

Linear algebra revision

1 Vector spaces and subspaces

We begin with the formal definition of vector spaces.

Definition A.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 $u, v, w \in V$

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

(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 $v + (-v) = (-v) + v = 0$

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 A.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 $V = \{(a_1, a_2, \dots, a_n) \mid a_1, \dots a_n \in K\}$ with addition and scalar multiplication defined by:

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, , a_n + b_n)$$
$$c(a_1, \dots, a_n) = (ca_1, \dots, ca_n).$$

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 any 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 $K[X]$. It is also denoted $\mathcal{P}(K)$.

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 $K[X]_{\leq d}$. It is also denoted $\mathcal{P}_d(K)$.

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 A.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 A.4 (Subspace theorem)**.** *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 A.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 Spanning, linear dependence, bases

Definition A.6. If S is a subset of a vector space V then a **linear combination** of S is a finite sum of the form

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

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 A.7. *If S is a non-empty subset of V, then $\langle S\rangle$ is a subspace of V.*

Example A.8. 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 A.9. We say that a subset S of a vector space V is **linearly dependent** if some non-zero linear combination gives the zero vector:

$$\exists a_1, a_2, \ldots, a_n \in F \;\; \exists v_1, v_2, \ldots, v_n \in S, \qquad \sum_{i=1}^{n} a_i v_i = 0 \quad \text{and not all } a_i \text{ are zero.}$$

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

Example A.10. 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 $\mathbb{R}[X]$.

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 A.11. *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 A.12. A **basis** for a vector space V is a linearly independent spanning set.

Example A.13. 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 A.14. *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 A.15. *If $\mathcal{B}_1$ and $\mathcal{B}_2$ are two bases of a vector space then they have the same cardinality. (This is, there exists a bijective function $f\colon \mathcal{B}_1 \to \mathcal{B}_2$.)*

Definition A.16. The *dimension* of a vector space V is the number of elements in a basis. We usually write this as $\dim V$.

By Theorem A.15, we know that this number will not depend on the particular choice of basis.

Example A.17. For the examples after Definition A.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. In fact $U + W$ is the smallest subspace containing both U and W.

Lemma A.18. *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 $\{v_1, \dots, v_l, u_1, \dots, u_m, w_1, \dots, w_n\}$. 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.

□

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 A.19. 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 $a \in F, v \in V$.

Example A.20.
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 A.21. 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 $\{f(v) : v \in V\}$.

Example A.22.
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 A.23. *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 A.24. 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 A.25. *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 $\mathcal{B} = \{v_1, \dots, v_m\}$. 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 $\{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$ of V. So the dimension of V is n.

We claim that $\{f(v_{m+1}), \dots, f(v_n)\}$ is a basis of the range of V. We must show that $\{f(v_{m+1}), \dots, f(v_n)\}$ is linearly independent and that every element of the range of V can be expressed as a linear combination of $\{f(v_{m+1}), \dots, f(v_n)\}$. We leave the details as Exercise A.16.

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

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 $\mathcal{B}_V = \{v_1, v_2, \dots, v_m\}$ be an ordered basis for V and $\mathcal{B}_W = \{w_1, w_2, \dots, w_n\}$ be an ordered basis for W. Then $f(v_i) \in W$ for each $i = 1, \dots, m$ 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 A.26. This matrix A is called the **matrix of** f with respect to the bases $\mathcal{B}_V$ and $\mathcal{B}_W$.

Explicitly, if

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

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 $\mathcal{B}_V$.

Example A.27. 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 $\mathcal{B}_V = \{v_1, v_2, \dots, v_m\}$ for a vector space V, each vector $v \in V$ can be written *uniquely* as a linear combination

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

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

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

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 $\mathcal{B}_V = \{v_1, v_2, \dots, v_m\}$ be a basis for V and $\mathcal{B}_W = \{w_1, w_2, \dots, w_n\}$ be a basis for W. Suppose that $\mathcal{B}'_V = \{v'_1, v'_2, \dots, v'_m\}$ is a new basis for V and $\mathcal{B}'_W = \{w'_1, w'_2, \dots, w'_n\}$ 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 A.28. *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 A.29. 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 A.30. 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.

6 Exercises

Exercise A.1. 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.

Exercise A.2. 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.

Exercise A.3. 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 $U_2 = V$.

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

Exercise A.5. 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) $\left\{\begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 4 & -6 \\ 3 & 8 \end{bmatrix}\right\}$

Exercise A.6. Show that any subset of a linearly independent set is also linearly independent.

Exercise A.7. 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 $\{E_{i,j} : 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n\}$ is a basis of $M_{m\times n}(F)$.

Exercise A.8. Show that the space $\mathcal{P}(F)$ does not have finite dimension.

Exercise A.9. What is the dimension of the space $M_{3\times 3}(\mathbb{F}_5)$?

Exercise A.10. 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.

Exercise A.11. Find the matrix of the linear transformation of Exercise A.10 with respect to the basis

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

found in Exercise A.7.

Exercise A.12. Find the matrix, with respect to the standard basis of $\mathbb{R}^2$, of the reflection across the x-axis. Now let $\mathcal{B}$ be the basis $\{(a, b), (c, d)\}$, where $a, b, c, d \in \mathbb{R}$ satisfy $ad - bc \neq 0$ 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.

Exercise A.13. Calculate the nullity and rank of the linear transformation f on $\mathbb{R}^3$ determined by (here $\{e_1, e_2, e_3\}$ is the standard basis)

$$\begin{aligned} f(e_1) &= e_1 - e_2 \\ f(e_2) &= e_2 - e_3 \\ f(e_3) &= e_1 - e_3 \end{aligned}$$

Exercise A.14. Calculate the nullity and rank of the linear transformation f on $\mathbb{F}_7^3$ determined by

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

Exercise A.15. 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.

Exercise A.16. Complete the proof of Lemma A.25.

Appendix B

Answers and hints to some exercises

Modular arithmetic and fields

1) Let $E \subset \mathbb{Z}$ be a subset of $\mathbb{Z}$ that is bounded below. We need to show the following:

$$\forall F \subseteq E \qquad F \neq \emptyset \implies F \text{ has a minimal element}$$

Proof. Let $k \in \mathbb{Z}$ be such that $\forall e \in E, e \geqslant k$. Let $F \subseteq E$ be a non-empty subset of E. Note that

$$f \in F \implies f \in E \implies f \geqslant k \implies f - k + 1 \geqslant 1 \implies f - k + 1 \in \mathbb{N}$$

Letting $F' = \{f - k + 1 \mid f \in F\}$ we have that $F' \subseteq \mathbb{N}$ and $F' \neq \emptyset$. Therefore F' has a minimal element, $m' \in F'$ say. Let $m = m' + k - 1$. Then $m \in F$ and for any $f \in F$ we have

$$f - k + 1 \geqslant m' \implies f \geqslant m$$

That is, m is a minimal element of F. □

2) We want to show

$$\forall a, b \in \mathbb{Z} \quad (\gcd(a, b) = 1 \iff (\exists x, y \in \mathbb{Z}, xa + yb = 1))$$

Let $a, b \in \mathbb{Z}$. That $\gcd(a, b) = 1 \implies (\exists x, y \in \mathbb{Z}, xa + yb = 1)$ is Bézout's Theorem (1.6).
For the converse, suppose that $x, y \in \mathbb{Z}$ are such that $xa + yb = 1$ and let $d = \gcd(a, b)$. Since d is a common divisor of a and b we have that $d \mid (xa + yb)$ (Lemma 1.3). But then we have that $d \in \mathbb{N}$ and $d \mid 1$, which implies that $d = 1$.

3) (a) $q = 8, r = 1$ (b) $q = 9, r = 5$ (c) $q = -5, r = 2$

5) Suppose that $qd + r = q'd + r'$ with $0 \leqslant r, r' < d$. Then

$$\begin{aligned} qd + r = q'd + r' &\implies (q - q')d = r' - r && (*) \\ &\implies |(q - q')d| < d \\ &\implies |q - q'| < 1 \\ &\implies q = q' \\ &\implies r = r' && \text{(from } *) \end{aligned}$$

6) Let $x, y, \alpha, \beta \in \mathbb{Z}$ be such that $c = \alpha a = \beta b$ and $xa + yb = 1$. Then

$$\begin{aligned} xa + yb = 1 &\implies x\alpha a + y\alpha b = \alpha \\ &\implies \alpha = x\beta b + y\alpha b \\ &\implies \alpha = b(x\beta + y\alpha) \\ &\implies c = b(x\beta + y\alpha)a \\ &\implies ab \mid c \end{aligned}$$

8) (a) 7 (b) 15 (c) 143 (d) 8 (e) 1

9)

(a) $\gcd(27, 33) = 3 = 5 \times 27 + (-4) \times 33$

(b) $\gcd(27, 32) = 1 = 11 \times 32 + (-13) \times 27$

(c) $\gcd(312, 317) = 13 = 5 \times 377 - 6 \times 312$

13)

(a) For the forward implication

$$[a]_m = [b]_m \implies a \in [b]_m \implies a \equiv b \pmod{m}$$

For the converse, suppose that $a \equiv b \pmod{m}$. Then

$$\begin{aligned} x \in [a]_m &\iff x \equiv a \pmod{m} \\ &\iff x \equiv b \pmod{m} && \text{(transitivity, Lemma 1.16)} \\ &\iff x \in [b]_m \end{aligned}$$

Therefore $[a]_m = [b]_m$.

(b) Suppose that $[a]_m \cap [b]_m \neq \emptyset$ and let $x \in [a]_m \cap [b]_m$. Then $x \equiv a \pmod{m}$ and $x \equiv b \pmod{m}$. Since the congruence relation is transitive (Lemma 1.16), we have $a \equiv b \pmod{m}$ and hence $[a]_m = [b]_m$ by part (a).

(c) Let $a \in \mathbb{Z}$. By Theorem 1.1 there exists $q, r \in \mathbb{Z}$ such that $a = qm + r$ and $r \in \{0, 1, \ldots, m-1\}$. Then note that $a \in [r]_m$ since $m \mid (a - r)$.

14) Suppose that $b, c \in \mathbb{Z}$ are such that $[b]_m[a]_m = [1]_m$ and $[c]_m[a]_m = [1]_m$. Then we have

$$[b]_m = [b]_m[1]_m = [b]_m[c]_m[a]_m = [b]_m[a]_m[c]_m = [1]_m[c]_m = [c]_m$$

15)

$(\mathbb{Z}/7\mathbb{Z}, \times)$

	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
3	0	3	6	2	5	1	4
4	0	4	1	5	2	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

$(\mathbb{Z}/8\mathbb{Z}, \times)$

	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7
2	0	2	4	6	0	2	4	6
3	0	3	6	1	4	7	2	5
4	0	4	0	4	0	4	0	4
5	0	5	2	7	4	1	6	3
6	0	6	4	2	0	6	4	2
7	0	7	6	5	4	3	2	1

16)

(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 933 is a multiple of 5 and 11, it is a multiple of 55, hence $-2 \equiv 933 \pmod{55}$.

17)

(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) $57102725 \equiv 5+7+1+0+2+7+2+5 \equiv 29 \equiv 2 \pmod 9$

18)

(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) $1 \times 2 \times 3 \times \ldots \times 20 \times 21 \equiv (2 \times 11) \times (3 \times \ldots \times 10) \times (12 \times \ldots \times 21) \equiv 0 \times (\ldots) \times (\ldots) \equiv 0 \pmod{22}$

19) 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 $326 \times 4471 \neq 1357546$.

20) Consider the possible values of $[x]_m^2 + [y]_m^2 + [z]_m^2$

21)

(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 $\{1,2,4,7,8,11,13,14\}$

22)

(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 $\mathbb{Z}/911\mathbb{Z}$.

23) 52

25) We need to show that

$$(x+y=0) \wedge (x+z=0) \implies y=z$$

Suppose that $x+y=x+z=0$. We have

$$\begin{aligned} x+y=0 &\implies (x+y)+z=0+z \\ &\implies (y+x)+z=z && \text{(addition is commutative, property of additive identity)} \\ &\implies y+(x+z)=z && \text{(addition is associative)} \\ &\implies y+0=z && (x+z=0) \\ &\implies y=z && \text{(property of additive identity)} \end{aligned}$$

27) Suppose that $1=0$, and let $a \in R$. Then $a = 1 \times a = 0 \times a = 0$. Therefore $R = \{0\}$.

29) 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.

30)

$$[1]_7 = [3]_7^6 \quad [2]_7 = [3]_7^2 \quad [3]_7 = [3]_7^1 \quad [4]_7 = [3]_7^4 \quad [5]_7 = [3]_7^5 \quad [6]_7 = [3]_7^3$$

31) The polynomial X, for example, does not have a multiplicative inverse.

32) Showing closure under addition and subtraction is relatively straightforward. It is reasonably easy to convince yourself of closure under multiplication. The problem is with multiplicative inverses. (I do not recommend attempting detailed proofs of all of the axioms!)
To see that

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

has an inverse, assume that $c_{-k} \neq 0$ and write the above as $c_k t^{-k} g$ 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.

33) The polynomial $X^2 + 1$ has no root in the field $\mathbb{Q}[\sqrt{2}]$.

35) (a) $m = 57, n = 36$ (b) 5 is relatively prime to 36 (c) 32 15 24 18 (d) $d = 29$ (e) 49 8

36) (a) $n = 40$ (b) (c) 17 14 48 25 17 15 2 15 (d) $d = 27$, rosebud

Linear algebra I

37) Suppose that $A, B, C \in M_n(K)$. Suppose that $A \sim B$ and $B \sim C$. Let $P, Q \in \mathrm{GL}(K)$ be such that $B = P^{-1}AP$ and $C = Q^{-1}BQ$. Then we have

(a) $A \sim A$ since $A = I^{-1}AI$

(b) $B \sim A$ since $B = P^{-1}AP \implies (P^{-1})^{-1}BP^{-1} = A$

(c) $A \sim C$ since $C = Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ)$

38) Let $P \in \mathrm{GL}_n(K)$ and fix a basis $\mathcal{B}'$ for V. Since P is invertible, its columns form a basis for $M_{n,1}$. Let $b_j \in V$ be such that $[b_j]_{\mathcal{B}'}$ is the j-th column of P. Then $\mathcal{B} = \{b_1, \dots, b_n\}$ is linearly independent and therefore a basis for V. Finally note that $[\mathrm{Id}_V]_{\mathcal{B}',\mathcal{B}} = P$.

39) Let $\{u_1, \dots, u_k\}$ be a basis for $\ker(f)$. Extend to a basis $\mathcal{B} = \{u_1, \dots, u_k, v_1, \dots, v_m\}$ of V. For each $1 \leqslant i \leqslant m$ define $w_i = f(v_i)$. Then $\{w_1, \dots, w_m\}$ is linearly independent since

$$\begin{aligned}
\sum_{i=1}^m \alpha_i w_i = 0 &\implies \sum_{i=1}^m \alpha_i f(v_i) = 0 \implies f(\sum_{i=1}^m \alpha_i v_i) = 0 \implies \sum_{i=1}^m \alpha_i v_i \in \ker(f) \\
&\implies \sum_{i=1}^m \alpha_i v_i = \sum_{j=1}^k \beta_j u_j \implies \sum_{i=1}^m \alpha_i v_i + \sum_{j=1}^k (-\beta_j) u_j = 0 \\
&\implies \forall i, \quad \alpha_i = 0 \quad \text{(since } \mathcal{B} \text{ is linearly independent)}
\end{aligned}$$

Extend to a basis $\mathcal{B}' = \{b_1, \dots, b_k, w_1, \dots, w_m\}$ of V. Then $[f]_{\mathcal{B}',\mathcal{B}}$ is a diagonal matrix with the diagonal entries being k zeros followed by m ones.

40) Note that, from the definition of an eigenvalue, there exists $v \in V_\lambda \setminus \{0\}$. In particular, $V_\lambda \neq \emptyset$. Now let $u, v \in V_\lambda$ and $k \in K$. Then $f(u + v) = f(u) + f(v) = \lambda u + \lambda v = \lambda(u + v)$ and $f(ku) = kf(u) = k\lambda u = \lambda ku$. Therefore $u + v \in V_\lambda$ and $ku \in V_\lambda$. It follows that V_λ is a subspace. Since $V_\lambda \setminus \{0\} \neq \emptyset$, $V \neq \{0\}$ and therefore $\dim(V_\lambda) \geqslant 1$.

41) Let P be an invertible matrix such that $B = P^{-1}AP$. Then

$$Bu = \lambda u \implies PBu = \lambda Pu \implies APu = \lambda Pu$$

Since P is invertible, $u \neq 0 \implies Pu \neq 0$. Therefore, if u is an eigenvector for B having eigenvalue λ, then Pu is an eigenvector for A having eigenvalue λ. Reversing the roles of A and B, we can conclude that λ is an eigenvalue for A iff λ is an eigenvalue for B.

42) We need to show that $f(f(v)) = \lambda f(v)$. Note that $f(f(v)) = f(\lambda v) = \lambda f(v)$. Therefore $f(V_\lambda) \subseteq V_\lambda$.

44) The proofs given for $1 \implies 2$ and $2 \implies 3$ did not assume that V is finite dimensional. We need only prove $3 \implies 1$.

Assume that $\mathcal{B}$ and $\mathcal{C}$ are as in 3. Since $\mathcal{B} \cup \mathcal{C}$ is a spanning set, we have that $V = U + W$. It remains to show that $U \cap W = \{0\}$.

Let $v \in U \cap W$. Then, since $v \in U$ we have $v = \sum_{i=1}^m \beta_i b_i$ for some $m \in \mathbb{N}$, $\beta_i \in K$ and $b_i \in \mathcal{B}$. Similarly, $v = \sum_{j=1}^n \gamma_j c_i$ for some $n \in \mathbb{N}$, $\gamma_j \in K$ and $c_j \in \mathcal{C}$. Then

$$\begin{aligned}\sum_{i=1}^m \beta_i b_i = \sum_{j=1}^n \gamma_j c_i &\implies \sum_{i=1}^m \beta_i b_i + \sum_{j=1}^n (-\gamma_j) c_i = 0\\ &\implies \forall i \forall j \quad (\beta_i = 0 \text{ and } \gamma_j = 0) \qquad \text{(since } \mathcal{B} \cup \mathcal{C} \text{ is linearly indepemdent)}\\ &\implies v = 0\end{aligned}$$

45) Let $V_1 = \{v \in V : f(v) = -v\}$ and $V_{-1} = \{v \in V : f(v) = -v\}$. We need to show that $V = V_1 + V_{-1}$ and that $V_1 \cap V_{-1} = \{0\}$. For the second note that

$$v \in V_1 \cap V_{-1} \implies v = -v \implies 2v = 0 \implies v = 0$$

The last implication above requires the hypothesis that $2 \neq 0$ in the field of scalars.

We need now to show that $V = V_1 + V_{-1}$. Note that since $X^2 - 1$ is the minimal polynomial of f, we have that $f^2 = \mathrm{Id}_V$. Given any $v \in V$ we have that $v = 2^{-1}(v + f(v)) + 2^{-1}(v - f(v))$ (we have again used that $2 \neq 0$). Finally, note that $2^{-1}(v + f(v)) \in V_1$ and $2^{-1}(v - f(v)) \in V_{-1}$ since

$$f(2^{-1}(v + f(v))) = 2^{-1} f((v + f(v))) = 2^{-1}(f(v) + f^2(v)) = 2^{-1}(f(v) + v) = 2^{-1}(v + f(v))$$

and

$$f(2^{-1}(v - f(v))) = 2^{-1} f((v - f(v))) = 2^{-1}(f(v) - f^2(v)) = 2^{-1}(f(v) - v) = -2^{-1}(v - f(v))$$

If $\mathcal{B}$ is a basis for V_1 and $\mathcal{C}$ is a basis for V_{-1}, then $[f]_{\mathcal{B}\cup\mathcal{C}}$ will be diagonal with all diagonal entries ±1.

47) From Theorem 2.18, there are $q(X), r(X) \in K[X]$ such that $p(X) = q(X)(X - k) + r(X)$ and $\deg(r(X)) = 0$. We need to show that $r(X) = 0$. Let $a \in K$ be such that $r(X) = a$. Then

$$p(k) = q(k)(k - k) + a \implies 0 = q(k)0 + a = a$$

Therefore $r(X) = 0$ and $p(X) = q(X)(X - k)$.

48)

(a) $\frac{1}{b-a}(X - a) + \frac{1}{a-b}(X - b) = 1$

(b) First we will show that $(X - a)$ is prime. That is, we show that

$$\forall p(X), q(X) \in K[X] \quad (X - a) \mid p(X)q(X) \implies (X - a) \mid p(X) \vee (X - a) \mid q(X) \qquad (*)$$

To prove this we have

$$\begin{aligned}(X - a) \mid p(X)q(X) &\implies p(X)q(X) = (X - a)r(X) && \text{(for some } r(X) \in K[X])\\ &\implies p(a)q(a) = 0\\ &\implies p(a) = 0 \vee q(a) = 0 && \text{(using that } K \text{ is a field)}\\ &\implies (X - a) \mid p(X) \vee (X - a) \mid q(X) && \text{(by Exercise 47)}\end{aligned}$$

Using $(*)$, the desired result can then be established using induction on m. Here's the outline.

$$\begin{aligned}d(X) \mid (X - a)^m &\implies d(X)e(X) = (X - a)^m && \text{(some } e(X) \in K[X])\\ &\implies (X - a) \mid d(X)e(X)\\ &\implies (X - a) \mid d(X) \vee (X - a) \mid e(X)\end{aligned}$$

and

$$\begin{aligned}
(X-a) \mid d(X) &\implies d(X) = d'(X)(X-a) && \text{(some } d'(X) \in K[X]\text{)}\\
&\implies d'(X)e(X) = (X-a)^{m-1} &&\\
&\implies d'(X) = (X-a)^{k'} && \text{(for some } 1 \leqslant k' \leqslant m-1\text{)}\\
&\implies d(X) = (X-a)^{k} && \text{(with } 2 \leqslant k \leqslant m\text{)}
\end{aligned}$$

and

$$\begin{aligned}
(X-a) \mid e(X) &\implies e(X) = e'(X)(X-a) && \text{(some } e'(X) \in K[X]\text{)}\\
&\implies d(X)e'(X) = (X-a)^{m-1} &&\\
&\implies d(X) = (X-a)^{k} && \text{(for some } 1 \leqslant k \leqslant m-1\text{)}
\end{aligned}$$

50) $(X-2)(X+1)$, X^2+X-1, X^3-1, $(X-1)^3$

51) 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.

52) 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. Note that

$$A^2 - 2A - 8I = 0 \implies A\frac{1}{8}(A-2I) = I$$

The above calculation shows that A^{-1} exists and is equal to $\frac{1}{8}(A-2I)$.

53) If the minimal polynomial has non-zero constant term, use the idea of the previous 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.

54) 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 $\{e_1, \dots, e_n\}$. 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 $f^n(e_i)$.

55)

(a) The characteristic polynomial is $(X-2)^3$. The only eigenvalue is 2.

(b) The minimal polynomial is $(X-2)^2 \in \mathbb{F}_5[X]$

(c) $\mathcal{B} = \{(1,0,1), (-2,1,0), (1,0,0)\}$

$$[f]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 & 3\\ 0 & 2 & 2\\ 0 & 0 & 2 \end{bmatrix}$$

56) Let P be an invertible matrix such that $B = P^{-1}AP$. Using standard properties of the determinant, we have

$$\begin{aligned}
c_A(X) &= \det(XI - A) = \det(P^{-1}P)\det(XI-A) = \det(P^{-1})\det(XI-A)\det(P)\\
&= \det(P^{-1}(XI-A)P) = \det(XP^{-1}IP - P^{-1}AP) = \det(XI - P^{-1}AP)\\
&= \det(XI - B)\\
&= c_B(X)
\end{aligned}$$

58) Suppose that $T : V \to V$ is nilpotent, and let $n \in \mathbb{N}$ be minimal with the property that that $T^n = 0$. Since $T^{n-1} \neq 0$, there exists $v \in V$ such that $w = T^{n-1}(v) \neq 0$. However, $T(w) = T^n(v) = 0$.

63)

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

64) 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 $J(-1,1)$

65) (a) no (b) yes (c) yes

66) Remember that the minimal polynomial divides the characteristic polynomial and has the same roots (possibly with different multiplicity). Then use Exercise 59. Note that 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. In the following list, a, b, c, d are distinct scalars.

characteristic polynomial	minimal polynomial	JNF
$(X-a)(X-b)(X-c)(X-d)$	$(X-a)(X-b)(X-c)(X-d)$	$J(a,1)\oplus J(b,1)\oplus J(c,1)\oplus J(d,1)$
$(X-a)^2(X-b)(X-c)$	$(X-a)(X-b)(X-c)$	$J(a,1)\oplus J(a,1)\oplus J(b,1)\oplus J(c,1)$
	$(X-a)^2(X-b)(X-c)$	$J(a,2)\oplus J(b,1)\oplus J(c,1)$
$(X-a)^2(X-b)^2$	$(X-a)(X-b)$	$J(a,1)\oplus J(a,1)\oplus J(b,1)\oplus J(b,1)$
	$(X-a)^2(X-b)$	$J(a,2)\oplus J(b,1)\oplus J(b,1)$
	$(X-a)(X-b)^2$	$J(a,1)\oplus J(a,1)\oplus J(b,2)$
	$(X-a)^2(X-b)^2$	$J(a,2)\oplus J(b,2)$
$(X-a)^3(X-b)$	$(X-a)(X-b)$	$J(a,1)\oplus J(a,1)\oplus J(a,1)\oplus J(b,1)$
	$(X-a)^2(X-b)$	$J(a,1)\oplus J(a,2)\oplus J(b,1)$
	$(X-a)^3(X-b)$	$J(a,3)\oplus J(b,1)$
$(X-a)^4$	$(X-a)$	$J(a,1)\oplus J(a,1)\oplus J(a,1)\oplus J(a,1)$
	$(X-a)^2$	$J(a,1)\oplus J(a,1)\oplus J(a,2)$ OR $J(a,2)\oplus J(a,2)$
	$(X-a)^3$	$J(a,1)\oplus J(a,3)$
	$(X-a)^4$	$J(a,4)$

67) Set D to be the diagonal part of J and set $N = J - D$. Then Exercise 54 shows that N is nilpotent. For the second part, choose a basis so that f is represented by a JNF 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.

68) 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.

Groups

70)

(a)

$$\begin{aligned} e' &= e * e' && \text{(since } \forall g \in G, e * g = g) \\ &= e && \text{(since } \forall g \in G, g * e' = g) \end{aligned}$$

(b)

$$\begin{aligned} h' &= (h * g) * h' && \text{(since } h * g = e) \\ &= h * (g * h') \\ &= h && \text{(since } g * h' = e) \end{aligned}$$

(c) Note that

$$(g * h) * (h^{-1} * g^{-1}) = g * (h * h^{-1}) * g^{-1} = g * g^{-1} = e$$
$$(h^{-1} * g^{-1}) * (g * h) = h^{-1} * (g^{-1} * g) * h = h^{-1} * h = e$$

and apply the previous part.

72) Let $I = \left[\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right]$, $-I = \left[\begin{smallmatrix}-1 & 0\\ 0 & -1\end{smallmatrix}\right]$, $A = \left[\begin{smallmatrix}-1 & 0\\ 0 & 1\end{smallmatrix}\right]$, $-A = \left[\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right]$, $B = \left[\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right]$, and $-B = \left[\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right]$.

V	I	A	$-I$	$-A$
I	I	A	$-I$	$-A$
A	A	I	$-A$	$-I$
$-I$	$-I$	$-A$	I	A
$-A$	$-A$	$-I$	A	I

C_4	I	B	$-I$	$-B$
I	I	A	$-I$	$-B$
B	B	$-I$	$-B$	I
$-I$	$-I$	$-B$	I	B
$-B$	$-B$	I	B	$-I$

73) For the second part, note that the product of two reflections having the same centre, is a rotation.

74) Use $w = yx^{-1}$ and $z = x^{-1}y$. For uniqueness, suppose that, also, $w_1x = y$. Then $wx = w_1x$ and so $wx(x^{-1}) = w_1x(x^{-1})$. Then $w(xx^{-1}) = w_1(xx^{-1})$ and so $we_G = w_1e_G$. 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.

75) 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.

76) Let $g, h \in G$. Then

$$ghgh = g^2h^2 \implies g^{-1}ghghh^{-1} = g^{-1}g^2h^2h^{-1} \implies hg = gh$$

77)

(a) $(264)(35)$

(b) (1356724)

(c) (1456)

78) Let $H = \cap_{i \in I} H_i$.

$$\begin{aligned} h, k \in H &\implies \forall i \in I,\ h \in H_i \wedge k \in H_i \\ &\implies \forall i \in I,\ hk^{-1} \in H_i && \text{(since } H_i \text{ is a subgroup)} \\ &\implies hk^{-1} \in H \end{aligned}$$

79) $\{0\}$, $\{0, 6\}$, $\{0, 4, 8\}$, $\{0, 3, 6, 9\}$, $\{0, 2, 4, 6, 8, 10\}$, $\mathbb{Z}/12\mathbb{Z}$.

80)

(a) No

(b) No

(c) Yes

81) Let $H = \{z \in \mathbb{C} \mid \exists n \in \mathbb{N},\ z^n = 1\}$ and let $h, k \in H$. Note that $H \neq \emptyset$ since $1 \in H$. Let $m, n \in \mathbb{N}$ be such that $h^m = 1$ and $k^n = 1$. Then $hk^{-1} \in H$ since $(hk^{-1})^{mn} = (h^m)^n(k^n)^{-m} = 1^n1^{-m} = 1$. Apply Lemma 3.12.

84)

(a) 12

(b) 10

(c) 2

(d) infinite order

(e) 10, 5, 20, 10

(f) 12, 2, 4

85) Note that for any $m \in \mathbb{N}$.

$$g^m = e \implies (g^{-1})^m = (g^m)^{-1} = e^{-1} = e$$

Similarly, $(g^{-1})^m = e \implies g^m = e$. That $|g| = |g^{-1}|$ is then immediate from the definition of order.

86) Let $g, h \in G$ and $m, n \in \mathbb{N}$ with $g^m = h^n = e$. Then $(gh)^{mn} = g^{mn}h^{mn} = e^n e^m = e$.

87)

$$\begin{aligned}
&A \neq I \qquad A^2 = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \neq I \qquad A^3 = I \\
&B \neq I \qquad B^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \neq I \qquad B^3 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \neq I \qquad B^4 = I \\
&AB = \begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix} \neq I \qquad (AB)^2 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \neq I \qquad (AB)^3 = \begin{bmatrix} -1 & 0 \\ 3 & -1 \end{bmatrix} \neq I \\
&(AB)^m = (-1)^m \begin{bmatrix} 1 & 0 \\ -m & 1 \end{bmatrix} \neq I \qquad \text{(for all } m \in \mathbb{N})
\end{aligned}$$

89) $|e| = 1$. For $1 \leqslant i \leqslant n$ we have: $|r^i s| = 2$, $|r^i| = n/\gcd(i,n)$.
To show that $|r^i| = n/\gcd(i,n)$ we can argue as follows. Let $d = \gcd(i,n)$ and let $i', n' \in \mathbb{N}$ be such that $i = di'$ and $n = dn'$. Note that i' and n' are relatively prime since

$$\begin{aligned}
d &= xi + yn && \text{(for some } x, y \in \mathbb{Z}) \\
&= xdi' + ydn' \\
\implies 1 &= xi' + yn'
\end{aligned}$$

Note that

$$(r^i)^{n'} = r^{di'n'} = (r^n)^{i'} = e^{i'} = e$$

and for $m \in \mathbb{N}$

$$\begin{aligned}
(r^i)^m = e &\implies r^{im} = e \implies n \mid im && (|r| = n) \\
&\implies n' \mid i'm \\
&\implies n' \mid m && (i' \text{ and } n' \text{ are relatively prime})
\end{aligned}$$

Therefore $|r^i| = n'$.

90)

(a) See Lemma 3.30.

(b) From Lemma 3.30 we have that $|\varphi(g)| \mid |g|$ and $|g| = |\varphi^{-1}(\varphi(g))| \mid |\varphi(g)|$.

91) By definition $SO(2) = \{A \in M_2(\mathbb{R}) \mid A^T A = I\}$. Each element of $SO(2)$ is of the form $A = \begin{bmatrix} \cos(\theta_A) & -\sin(\theta_A) \\ \sin(\theta_A) & \cos(\theta_A) \end{bmatrix}$ for some $\theta_A \in (-\pi, \pi]$. Show that the map $\varphi : SO(2) \to S^1$, $\varphi(A) = \theta_A$ is an isomorphism.

92) 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 (i.e., is isomorphic to) the symmetries of a regular m-gon.

93)

(a) The element $-1 \in \mathbb{R}^\times$ has order 2. No element in $(\mathbb{R}, +)$ has order 2.

(b) Suppose that $\varphi : \mathbb{Z} \to \mathbb{Q}$ is an isomorphism. Consider the element $z = \varphi^{-1}(\varphi(1)/2) \in \mathbb{Z}$. Then $\varphi(z + z) = \varphi(z) + \varphi(z) = \varphi(1)$, which implies that $2z = 1$. Contradiction.

(c) Suppose that $\varphi : (\mathbb{Q}, +) \to (\mathbb{Q}^+, \times)$ is an isomorphism. Consider the element $q = \varphi(\varphi^{-1}(2)/2) \in (\mathbb{Q}^+, \times)$. Then $q^2 = \varphi(\varphi^{-1}(2)/2)\varphi(\varphi^{-1}(2)/2) = \varphi(\varphi^{-1}(2)/2 + \varphi^{-1}(2)/2) = \varphi(\varphi^{-1}(2)) = 2$. Contradiction.

96) The order of $H \cap K$ must divide both 7 and 29.

97) Each right coset

$$H \begin{bmatrix} x_0 & y_0 \\ 0 & 1 \end{bmatrix} = \left\{ \begin{bmatrix} zx_0 & zy_0 \\ 0 & 1 \end{bmatrix} \mid z > 0 \right\}$$

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

$$\begin{bmatrix} x_0 & y_0 \\ 0 & 1 \end{bmatrix} H = \left\{ \begin{bmatrix} x_0 z & y_0 \\ 0 & 1 \end{bmatrix} \mid z > 0 \right\}$$

can be identified with a horizontal half-line.

98) 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).

99)

(a) The cosets of H are H and Hb. Now show that if $ab \in Hb$, then $a \in H$.

(b) Consider cosets H, Hx, Hy with $H \neq Hx$ and $H \neq Hy$. So $x, y \notin H$ and therefore $x, y^{-1} \notin H$. Thus $xy^{-1} \in H$ and so $Hx = Hy$. Thus there can be at most one coset different from H.

100) Suppose that r is a rotation through $2\pi/5$ and s is a reflection. Then the subgroups are

$$\{e\}, \langle r \rangle, \langle s \rangle, \langle rs \rangle, \langle r^2 s \rangle, \langle r^3 s \rangle, \langle r^4 s \rangle, D_5$$

101)

(a) $D_4 = \{e, r, r^2, r^3, s, rs, r^2 s, r^3 s\}$. The cyclic subgroups are: $\langle e \rangle, \langle r \rangle = \langle r^3 \rangle, \langle r^2 \rangle, \langle s \rangle, \langle rs \rangle, \langle r^2 s \rangle, \langle r^3 s \rangle$.

(b) The idea is to find two reflections that commute.

One such pair is s and $r^2 s$. This subgroup is $\langle s, r^2 s \rangle = \{e, s, r^2 s, r^2\}$. Observe that no element has order 4.

Another pair of reflections that commute is rs and $r^3 s$. This subgroup is $\langle rs, r^3 s \rangle = \{e, rs, r^3 s, r^2\}$. Observe that no element has 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.

102) For the subgroups of orders 2,3, take cyclic subgroups generated by rotations about the midpoint of an edge and a vertex (respectively). For the subgroup of order 4, consider the set of all rotations about axes connecting the midpoints of opposite sides (you need to show it gives a subgroup). For the last part, first establish that there is no element of order 6 and so no cyclic subgroup of order 6.

103) Let $p = 29$ and note that p is prime. An element of G must have order dividing p^2 and so must have order 1, p or p^2. If there is an element of order p^2, then G is cyclic.

105) Check explicitly that $\forall \sigma \in S_4 \; \forall h \in H, \; \sigma h \sigma^{-1} \in H$. Alternatively, show that for $a, b, c, d \in \{1, 2, 3, 4\}$ distinct, we have

$$\sigma(a, b)(c, d)\sigma^{-1} = (\sigma(a), \sigma(b))(\sigma(c), \sigma(d))$$

106) Let $g \in G \backslash H$. The two left cosets are $\{H, gH\}$. The two right cosets are $\{H, Hg\}$. Since the left cosets partition G and the right cosets partition G, it must be the case that $gH = Hg$. Therefore H is normal.

110) Let $H \leqslant G$ have order n. 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 $gHg^{-1} = H$.

111) The normal subgroups of D_4 are:

$$\{e\}, \langle r^2 \rangle, \langle r \rangle, \langle s, r^2 s \rangle, \langle rs, r^3 s \rangle, D_4$$

112)

(a) This is essentially straight calculation.

(b) There are five cyclic subgroups. 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) From the pervious parts we know that all subgroups aside from $\langle U\rangle$, $\langle -U\rangle$ and Q_8 have size 4. They are therefore of index 2 and hence normal. The subgroups $\langle U\rangle$ and Q_8 are obviously normal. That $\langle -U\rangle$ is normal is an easy check.

(e) No. *All* proper subgroups of Q_8 are cyclic but this is not true for D_4. (Alternatively, all subgroups of Q_8 are normal but this is not true for D_4.)

114)

$$\mathbb{Q}/\mathbb{Z} = \{a + \mathbb{Z} : a \in \mathbb{Q}\} \leqslant \{a + \mathbb{Z} : a \in \mathbb{R}\} = \mathbb{R}/\mathbb{Z}.$$

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 $a \in \mathbb{Q}$.

115)

(a) It's enough to note that $sr^4s^{-1} = sr^4s = r^4 \in H$ and $rr^4r^{-1} = r^4 \in H$.

(b)

	H	Hr	Hr^2	Hr^3	Hs	Hrs	Hr^2s	Hr^3s
H	H	Hr	Hr^2	Hr^3	Hs	Hrs	Hr^2s	Hr^3s
Hr	Hr	Hr^2	Hr^3	H	Hrs	Hr^2s	Hr^3s	Hs
Hr^2	Hr^2	Hr^3	H	Hr	Hr^2s	Hr^3s	Hs	Hrs
Hr^3	Hr^3	H	Hr	Hr^2	Hr^3s	Hs	Hrs	Hr^2s
Hs	Hs	Hr^3s	Hr^2s	Hrs	H	Hr^3	Hr^2	Hr
Hrs	Hrs	Hs	Hr^3s	Hr^2s	Hr	H	Hr^3	Hr^2
Hr^2s	Hr^2s	Hrs	Hs	Hr^3s	Hr^2	Hr	H	Hr^3
Hr^3s	Hr^3s	Hr^2s	Hrs	Hs	Hr^3	Hr^2	Hr	H

117) From the first isomorphism theorem, we know that $\mathrm{im}(\varphi) \cong (\mathbb{Z}/8\mathbb{Z})/\ker(\varphi)$. The possibilities for $\ker(\varphi)$ are:

$$\{e\}, \langle 4\rangle, \langle 2\rangle, \langle 1\rangle$$

(These are the only subgroups. All subgroups are normal because the group is abelian.) Then note that $(\mathbb{Z}/8\mathbb{Z})/\{e\} \cong (\mathbb{Z}/8\mathbb{Z})$, $(\mathbb{Z}/8\mathbb{Z})/\langle 4\rangle \cong (\mathbb{Z}/4\mathbb{Z})$, $(\mathbb{Z}/8\mathbb{Z})/\langle 2\rangle \cong (\mathbb{Z}/4\mathbb{Z})$, and $(\mathbb{Z}/8\mathbb{Z})/\langle 1\rangle \cong \{e\}$.

Linear algebra II

120)

(a) $\sqrt{19}$ (b) $\sqrt{\frac{11}{30}}$ (c) $\sqrt{30}$

121) If $u = v$ then the claim gives $\|2u\| = 4\|u\|$, which is false if $u \neq 0$. Try proving $\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$ by expanding into inner products.

122) For the third part, note that

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle = \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \overline{\langle x, y\rangle} + \langle y, y\rangle \\ &= \|x\|^2 + 2\Re(\langle x, y\rangle) + \|y\|^2 \leqslant \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2\end{aligned}$$

The required inequality then follows by taking $x = u - w$ and $y = w - v$.

124) We know from Exercise 123 that $W \subseteq (W^\perp)^\perp$ and that $\dim V = \dim W + \dim W^\perp = \dim W^\perp + \dim(W^\perp)^\perp$. Therefore $\dim W = \dim(W^\perp)^\perp$ and $W \subseteq (W^\perp)^\perp$. It follows that $W = (W^\perp)^\perp$.

126) Suppose that $A = [\mathrm{Id}]_{\mathcal{C},\mathcal{B}}$ for orthonormal bases $\mathcal{C}$ and $\mathcal{B} = \{b_1, \ldots, b_n\}$. The j-th column of A is equal to $[b_j]_\mathcal{C}$. The i-th row of A^* is equal to $([b_i]_\mathcal{C})^*$. Therefore the ij-th entry of A^*A is equal to $([b_i]_\mathcal{C})^*[b_j]_\mathcal{C}$. Then note that $([b_i]_\mathcal{C})^*[b_j]_\mathcal{C} = \overline{\langle b_i, b_j\rangle}$ because $\mathcal{C}$ is orthonormal.

127) Let $\lambda \in \mathbb{C}$ and $v \in V \setminus \{0\}$ be such that $f(v) = \lambda v$.

(a) If $f^* = f$, then
$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle f(v), v\rangle = \langle v, f(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$$
Therefore $\lambda = \overline{\lambda}$.

(b) If $f^*f = \mathrm{Id}$, then
$$\lambda\overline{\lambda}\langle v, v\rangle = \langle \lambda v, \lambda v\rangle = \langle f(v), f(v)\rangle = \langle v, f^*f(v)\rangle = \langle v, v\rangle$$
Therefore $\lambda\overline{\lambda} = 1$.

128) Let $K = \ker(f)$. Let $k \in \ker(f)$ and $u \in V$. Then $\langle f^*(u), k\rangle = \langle u, (f^*)^*(k)\rangle = \langle u, f(k)\rangle = \langle u, 0\rangle = 0$. Therefore $\mathrm{im}(f^*) \subseteq K^\perp$. Further, $(\mathrm{im}(f^*))^\perp \subseteq K$ since
$$\begin{aligned} w \in \mathrm{im}(f^*)^\perp &\implies \forall v \in V, \quad \langle w, f^*(v)\rangle = 0 \\ &\implies \forall v \in V, \quad \langle f(w), v\rangle = 0 \\ &\implies \langle f(w), f(w)\rangle = 0 \\ &\implies f(w) = 0 \end{aligned}$$
It follows that $\mathrm{im}(f^*) = (\mathrm{im}(f^*)^\perp)^\perp \supseteq K^\perp$.
$$\mathrm{rank}(f^*) = \dim(\mathrm{im}(f^*)) = \dim(K^\perp) = V - \dim(K) = \dim(\mathrm{im}(f)) = \mathrm{rank}(f)$$

131) Suppose that $\{v_1, \ldots, v_n\}$ is an orthonormal basis of V. Suppose that $w = \sum_i a_i v_i$ and that $f(v_j) = \sum_i b_{ij} v_i$. Then $\langle f(v_j), w\rangle = \sum_i b_{ij} a_i$. Set $w_1 = \sum_k c_k v_k$ where $c_k = \sum_i b_{ik} a_i$. Note that $\langle v_j, w_1\rangle = c_j = \langle f(v_j), w\rangle$.
For the uniqueness, note that if w_2 also satisfies the conditions, then $\langle v, w_1\rangle = \langle v, w_2\rangle$ for all $v \in V$.

132)

(a) $\langle g(u+w), u+w\rangle = 0 \implies \langle g(u) + g(w), u+w\rangle = 0 \implies \langle g(u), w\rangle + \langle g(w), u\rangle = 0$

(b) Observe that $\langle g(w), u\rangle = \langle w, g^*(u)\rangle = \langle w, g(u)\rangle = \overline{\langle g(u), w\rangle} = \langle g(u), w\rangle$ to show that $2\langle g(u), w\rangle = 0$. Now deduce that g is zero.

(c) As in the previous part but deduce that the real part of $\langle g(u), w\rangle$ is 0.

(d) If $\langle g(iu), w\rangle$ is imaginary for all $u, w \in V$, then $\langle g(u), w\rangle = i\langle g(u), w\rangle$ is both real and imaginary and so zero.

(e) Take $w = g(u)$.

135) You can solve this by writing the matrix as $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. Then multiply this by its transpose and equate the result to the identity. It will be useful to observe that if $x^2 + y^2 = 1$, then there is an angle θ so that $x = \cos\theta$ and $y = \sin\theta$.

136) Show that $AA^* = UDD^*U^* = UD^*DU^* = A^*A$.

137)
$$\begin{aligned} ff^* = f^*f &\iff \forall v \in V, \quad ff^*(v) - f^*f(v) = 0 \\ &\iff \forall u, v \in V, \quad \langle u, ff^*(v) - f^*f(v)\rangle = 0 \\ &\iff \forall u, v \in V, \quad \langle u, ff^*(v)\rangle - \langle u, f^*f(v)\rangle = 0 \\ &\iff \forall u, v \in V, \quad \langle f^*(u), f^*(v)\rangle - \langle f(u), f(v)\rangle = 0 \end{aligned}$$

138) Find a diagonal matrix similar to A, take the square root of that and use that to find a square root of A.
The matrix $\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ (which is not normal) has no square root.

139)

(a) No; different eigenvalues (b) No; different eigenvalues (c) Yes

140) No; they may have different eigenvalues.

142) 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. Then $p(A) = A^*$ and so $p(f) = f^*$.

143) 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 $f^*g = gf^*$.

144)

(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.

Groups II

145)

(a) (i) For injectivity:

$$\varphi_g(x) = \varphi_g(y) \implies g{\cdot}x = g{\cdot}y \implies g^{-1}{\cdot}(g{\cdot}x) = g^{-1}{\cdot}(g{\cdot}y) \implies (g^{-1}g){\cdot}x = (g^{-1}g){\cdot}y \implies e{\cdot}x = e{\cdot}y \implies x = y$$

For surjectivity: given $y \in X$ let $x = g^{-1} \cdot x$. Then $\varphi_g(x) = g \cdot (g^{-1} \cdot y) = (gg^{-1}) \cdot y = e \cdot y = y$

(ii) Let $g, h \in G$. For all $x \in X$ we have

$$\varphi_{gh}(x) = (gh) \cdot x = g \cdot (h \cdot x) = g \cdot \varphi_h(x) = \varphi_g(\varphi_h(x)) = \varphi_g \circ \varphi_h(x)$$

Since this holds for all $x \in X$, we have $\Phi(gh) = \varphi_{gh} = \varphi_g \circ \varphi_h = \Phi(g)\Phi(h)$

(b) Since Ψ is a homomorphism, $\Psi(e_G) = e_{S_X} = \mathrm{Id}_X$. Therefore, for all $x \in X$ we have

$$e_G \cdot x = \Psi(e_G)(x) = \mathrm{Id}_X(x) = x$$

Let $g, h \in G$. Then

$$(gh) \cdot x = \Psi(gh)(x) = \Psi(g)\Psi(h)(x) = \Psi(g)(h \cdot x) = g \cdot (h \cdot x)$$

147)

(a) Orbits: $\{1, 2, 3\}, \{4\}$; stabilisers, $\mathrm{Stab}(1) = \mathrm{Stab}(2) = \mathrm{Stab}(3) = \{e\}, \mathrm{Stab}(4) = G$

(b) Orbits: $\{1, 2, 3, 4\}$; stabilisers, $\mathrm{Stab}(1) = \mathrm{Stab}(2) = \mathrm{Stab}(3) = \mathrm{Stab}(4) = \{e\}$

(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 (which is isomorphic to S_3)

(e) Orbits: $\{1, 2, 3, 4\}$; stabilisers, $\mathrm{Stab}(1) = \mathrm{Stab}(3) = \langle(24)\rangle, \mathrm{Stab}(2) = \mathrm{Stab}(4) = \langle(13)\rangle$. (It's convenient to think of $\{1, 2, 3, 4\}$ as the vertices of a square.)

149) $\{U\}, \{-U\}, \{I, -I\}, \{J, -J\}, \{K, -K\}$

150)

(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

151) 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.

152) Show that if $g = khk^{-1}$, then $C_G(g) = kC_G(h)k^{-1}$. In more detail,

$$x \in C_G(g) \iff xg = gx \iff xkhk^{-1} = khk^{-1}x \iff k^{-1}xkh = hk^{-1}xk \iff k^{-1}xk \in C_G(h)$$

153) This can be 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) $\{A \in GL(3, \mathbb{R}) \mid A$ is diagonal$\}$

(b) $\{A \in GL(3, \mathbb{R}) \mid \exists B \in GL(2, \mathbb{R})\, \exists C \in GL(1, \mathbb{R}),\ A = B \oplus C\}$

(c) $\{A \in GL(3, \mathbb{R}) \mid \exists a, b, e \in \mathbb{R}, A = \left[\begin{smallmatrix} a & b & 0 \\ 0 & a & 0 \\ 0 & 0 & e \end{smallmatrix}\right]\}$

(d) $\{A \in GL(3, \mathbb{R}) \mid \exists a, b, c, d, e \in \mathbb{R}, A = \left[\begin{smallmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{smallmatrix}\right]\}$

(e) $\{A \in GL(3, \mathbb{R}) \mid \exists a, b, c \in \mathbb{R}, A = \left[\begin{smallmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{smallmatrix}\right]\}$

154) 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.

155) 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 $G = \langle(123)(45678)\rangle \leqslant S_8$ 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).

157) Choose h so that G/Z is generated by hZ. Then each element of G can be written in the form $h^i z$ for some $z \in Z$ and some $i \in \mathbb{Z}$.

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 $|G|$ by the orbit-stabiliser relation. Therefore $|C| = 1$ and $|G| = 2$. Thus G is (isomorphic to) the cyclic group of order 2.

If there are three classes, $\{e_G\}$, C and D say with $|C| \leqslant |D|$, then $|G| = 1 + |C| + |D|$ and both $|C|$ and $|D|$ divide $|G|$. Show that the only solutions to this equation are $|C| = |D| = 1$ or $|C| = 1, |D| = 2$ or $|C| = 2, |D| = 3$. The first possibility corresponds to the cyclic group of order 3. The third possibility corresponds to S_3. The second possibility does not occur because if $|G| = 4$, then G is abelian and therefore the number of conjugacy classes if $|G| = 4$.

159) 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.

160) Use the previous exercise to show that there is a normal subgroup of order p, generated by x say. By Cauchy's themorem there is an elment of order 2. Let y be an element of order 2. Since $\langle x \rangle$ is normal, $yxy^{-1} \in \langle x \rangle$. 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.

161)

$g \in D_8$	e	r, r^3, r^5, r^7	r^2, r^6	r^4	s, r^2s, r^4s, r^6s	rs, r^3s, r^5s, r^7s
$\lvert X^g \rvert$	70	0	2	6	6	6

There are $\frac{1}{16}(70 + 2 \times 2 + 6 + 4 \times 6 + 4 \times 6) = 8$ orbits.

162) Consider the homomorphism $\varphi : L \to K$ given by $\varphi(g) = \pi(g)$. Then $\operatorname{im}(\varphi) = K$ and $\ker(\varphi) = N$. By the first isomorphism theorem we have that $K \cong L/N$ and therefore, using Lagrange, $|L| = |K| \times |N| = p^{s-1}p$.

163) If $|G| = p^n$, then it follows from Lagrange's theorem that for all $g \in G$ $|g|$ divides p^n and is therefore a power of p.

For the converse, suppose that all elements of G have order that is a power of p. If G were not a p-group then there is a prime $q \in \mathbb{N}$ such that $q \mid |G|$ and $q \neq p$. But then by Cauchy's theorem (or the first Sylow theorem), there would be an element $g \in G$ of order q.

164)

(a) This follows from the fact that gHg^{-1} is a subgroup of G and has the same size as H.

(b) If H is the only Sylow p-subgroup, then from the previous part we have that $gHg^{-1} = H$ for all $g \in G$.

165) Let H be a Sylow q-subgroup of G. Then $|H| = p$. By the third Sylow theorem we have that $n_q \mid pq$ and $n_q \equiv 1 \pmod{q}$. The only divisors of pq are 1, p, q, and pq. Since $p < q$, $p \not\equiv 1 \pmod{q}$. Also, $pq \equiv q \equiv 0 \not\equiv 1 \pmod{q}$. The only possibility is therefore that $n_q = 1$.

166) This is very similar to the previous exercise. We have that $n_{17} \mid (3 \times 5 \times 17)$ and $n_{17} \equiv 1 \pmod{17}$. The divisors of 255 that are not divisible by 17 are: 1,3,5, and 15. Of these, the only value that is congruent to 1 modulo 17 is 1. Since $n_{17} = 1$, the Sylow 17-subgroup is normal (see Exercise 164).

Linear algebra revision

A.1) 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.

A.2) Use the definition of subspace, as in the previous question.

A.3) 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?.

A.4)

(a) linearly independent, not a basis;

(b) linearly independent, not a basis;

(c) linearly dependent, not a basis.

A.5) (a) yes (b) yes (c) no

A.8) $\{1, X, X^2, X^3, \dots\}$ is an infinite linearly independent set in $F[X]$. In a finite dimensional vector space every linearly independent set is finite.

A.9) 9

A.10) Given $A_1, A_2 \in M_{2\times 2}$ and $\alpha \in \mathbb{R}$ we have

$$\begin{aligned} g(A_1+A_2) &= (A_1+A_2)B = A_1B + A_2B = g(A_1)+g(A_2) \\ g(\alpha A_1) &= (\alpha A_1)B = \alpha(A_1B) = \alpha g(A_1) \end{aligned}$$

(Notice that it doesn't matter what matrix B is.)

A.11) $\begin{bmatrix} 2 & 3 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$

A.12) The matrix is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. The matrix with respect to the new basis is

$$\frac{1}{ad-bc}\begin{bmatrix} ad+bc & 2cd \\ -2ab & -(ad+bc) \end{bmatrix}$$

A.13) The nullity is 1; the rank is 2.

A.14) The nullity is 1; the rank is 2.

A.15) 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.

Index