Groups

Copier’s Message

These notes may contain errors. In fact, they almost certainly do since they were just copied down by me during lectures and everyone makes mistakes when they do that. The fact that I had to type pretty fast to keep up with the lecturer didn’t help. So obviously don’t rely on these notes.

If you do spot mistakes, I’m only too happy to fix them if you email me at mdj27@cam.ac.uk with a message about them. Messages of gratitude, chocolates and job offers will also be gratefully received.

Whatever you do, don’t start using these notes instead of going to the lectures, because the lecturers don’t just write (and these notes are, or should be, a copy of what went on the blackboard) – they talk as well, and they will explain the concepts and processes much, much better than these notes will. Also beware of using these notes at the expense of copying the stuff down yourself during lectures – it really makes you concentrate and stops your mind wandering if you’re having to write the material down all the time. However, hopefully these notes should help in the following ways;

- you can catch up on material from the odd lecture you’re too ill/drunken/lazy to go to;
- you can find out in advance what’s coming up next time (if you’re that sort of person) and the general structure of the course;
- you can compare them with your current notes if you’re worried you’ve copied something down wrong or if you write so badly you can’t read your own handwriting. Although if there is a difference, it might not be your notes that are wrong!

These notes were taken from the course lectured by Dr Saxl in Michaelmas 2009. If you get a different lecturer (increasingly likely as time goes on) the stuff may be rearranged or the concepts may be introduced in a different order, but hopefully the material should be pretty much the same. If they start to mess around with what goes in what course, you may have to start consulting the notes from other courses. And I won’t be updating these notes (beyond fixing mistakes) – I’ll be far too busy trying not to fail my second/third/nth year courses.

Good luck – Mark Jackson

Schedules

These are the schedules for the year 2009/10, i.e. everything in these notes that was examinable in that year. The numbers in brackets after each topic give the subsection of these notes where that topic may be found, to help you look stuff up quickly.

Examples of groups

- Axioms for groups (1.1). Examples from geometry: symmetry groups of regular polygons (1.4.4), cube, tetrahedron (4.6). Permutations on a set; the symmetric group (1.4). Subgroups (1.3.2, 2.1) and homomorphisms (1.3.4, 5.1). Symmetry groups as subgroups of general permutation groups (1.4.4). The Möbius group (7.2); cross-ratios (7.7), preservation of circles, the point at infinity (7.6). Conjugation (4.3). Fixed points of Möbius maps and iteration (7.4).

Lagrange’s theorem

- Cosets (2.2). Lagrange’s theorem (2.3). Groups of small order (up to order 8) (3.4). Quaternions (3.4.4). Fermat-Euler theorem from the group-theoretic point of view (2.4).
Contents

1. Introduction to groups; permutations and symmetry groups ............................................ 3
  1.1 Introduction .................................................................................................................. 3
  1.2 Functions .................................................................................................................... 4
  1.3 Some preliminaries ..................................................................................................... 4
  1.4 The symmetric and dihedral groups ........................................................................... 5
  1.5 More on permutations ............................................................................................... 7
  1.6 Signs ......................................................................................................................... 8
2. Subgroups, cosets and Lagrange’s theorem ................................................................. 9
  2.1 Subgroups ................................................................................................................ 9
  2.2 Cosets ...................................................................................................................... 10
  2.3 Lagrange’s theorem .............................................................................................. 11
  2.4 An application of Lagrange’s theorem: Fermat’s little theorem ....................... 11
3. Small groups ............................................................................................................... 12
  3.1 Cyclic groups ......................................................................................................... 12
  3.2 Direct (Cartesian) products of groups .................................................................... 13
  3.3 Dihedral groups ..................................................................................................... 13
  3.4 The groups of order \( \leq 8 \) .................................................................................... 14
4. Group actions .......................................................................................................... 15
  4.1 Introduction to actions ......................................................................................... 15
  4.2 The orbit-stabiliser theorem .................................................................................. 16
  4.3 Conjugation ........................................................................................................... 17
  4.4 Cayley’s theorem .................................................................................................. 18
  4.5 Cauchy’s theorem ................................................................................................. 19
  4.6 Groups of symmetries of regular solids .............................................................. 19
5. Homomorphisms, normal subgroups and quotient groups ..................................... 20
  5.1 Homomorphisms ................................................................................................... 20
  5.2 Normal subgroups ................................................................................................ 21
  5.3 Quotient group ...................................................................................................... 21
  5.4 Isomorphism theorem ......................................................................................... 22
  5.5 Simple groups ....................................................................................................... 23
6. Groups of matrices I ............................................................................................. 24
1. Introduction to groups; permutations and symmetry groups

1.1 Introduction

1.1.1 Definition

Definition. We say that \((G, \ast)\) is a group if \(G\) is a set and \(\ast\) is a binary operation (one that combines two elements) with the following properties holding:

- if \(a, b \in G\) then \(a \ast b \in G\) (closure axiom)
- if \(a, b, c \in G\) then \((a \ast b) \ast c = a \ast (b \ast c)\) (associative law)
- there exists an element \(e \in G\) such that \(a \ast e = e \ast a = a\) for all \(a \in G\) (existence of the identity element)
- for any \(a \in G\), there exists an element \(a^{-1} \in G\) such that \(a \ast a^{-1} = e = a^{-1} \ast a\) (existence of inverses).

1.1.2 Examples and remarks

Examples. (i) \((\mathbb{Z}, +)\), the integers under addition, where \(e = 0\) and \(a^{-1} = -a\).
(ii) \((\mathbb{Q}, +), (\mathbb{R}, +)\) etc. are all groups as well.
(iii) \((\mathbb{Q} - \{0\}, \times)\), the rationals under multiplication, where \(e = 1\) and \(a^{-1} = 1/a\).

However \((\mathbb{Z}, -)\) is not a group because \(-\) is not associative, and \((\mathbb{Z}, \times)\) lacks inverses.

Remarks.

- Groups may or may not be abelian (i.e. \(a \ast b = b \ast a\) for all \(a, b \in G\))
- \(((\pm 1), \times)\) is a group.
- Associativity means no brackets are needed, i.e. \(a \ast b \ast c\) is unambiguous.
- We often omit the \(\ast\) and write e.g. \(ab = a \ast b\).

1.1.3 Basic properties and algebra

Lemma 1.1. (i) Let \((G, \ast)\) be a group. Then the identity element is unique, that is to say, if \(a \ast e = a = e \ast a\) for all \(a, b \in G\), and \(a \ast e' = a = e' \ast a\) for all \(a, b \in G\), then \(e = e'\).

(ii) Inverses are also unique.

Proof (i). \(e = e \ast e' = e'\) using the fact that \(e\) is a right identity and \(e'\) is a left identity.
Proof (ii). Assume \( a \cdot b = e = b' \cdot a \). Then \( b = b' \), since
\[
b' = b'^{\cdot} e = b'^{\cdot} (a \cdot b) = (b'^{\cdot} a) \cdot b = e \cdot b = b.
\]

**Lemma 1.2.** Let \((G,\ast)\) be a group and let \(a, b \in G\). Then (i) \((ab)^{-1} = b^{-1}a^{-1}\) (not \(a^{-1}b^{-1}\)) and (ii) \((a^{-1})^{-1} = a\).

Proof (i). \((ab)(b^{-1}a^{-1}) = aea^{-1} = aa^{-1} = e\) and similarly, \(b^{-1}a^{-1}ab = e\). Therefore \((ab)^{-1} = b^{-1}a^{-1}\).

Proof (ii). Exercise.

**Lemma 1.3.** If \((G,\ast)\) is a group, then \(ax = bx \Rightarrow a = b\).

Proof. \(axx^{-1} = bxx^{-1}\) so \(ae = be\) so \(a = b\), and similarly in reverse.

1.2 Functions

If \(A, B\) are sets, a function (or mapping) \(f : A \to B\) is a rule that assigns to each element \(a \in A\) exactly one element \(f(a) \in B\).

Our functions are single-valued, and obey the rule that \(f(a) \in B\) for all \(a \in A\).

If we have \(f_1, f_2 : A \to B\), then \(f_1 = f_2\) if \(f_1(a) = f_2(a)\) for all \(a \in A\).

### 1.2.1 Injective, surjective and bijective functions

**Definition.** \(f : A \to B\) is **injective** (or one-to-one) if \(f(a_1) = f(a_2) \Rightarrow a_1 = a_2\), so that each element in \(A\) is assigned a **different** element in \(B\).

**Definition.** \(f : A \to B\) is **surjective** (or onto \(B\)) if each \(b \in B\) has an \(a \in A\) with \(f(a) = b\), so that everything in \(B\) is ‘covered’ by something in \(A\). Injectivity implies surjectivity for **finite** sets.

**Definition.** \(f : A \to B\) is **bijective** if it’s injective and surjective.

### 1.2.2 Composition of functions

If \(g : A \to B\) and \(f : B \to C\), then we define \(f \circ g : A \to C\) by, for all \(a \in A\), \(f \circ g(a) = f(g(a))\).

**Lemma 1.4.** If \(g : A \to B\) and \(f : B \to C\) are both injective, both surjective or both bijective, then so is \(f \circ g\).

Proof (injective case). If \(f \circ g(a_1) = f \circ g(a_2)\) then \(a_1 = a_2\), because \(f(g(a_1)) = f(g(a_2))\). But \(f\) is injective, thus \(g(a_1) = g(a_2)\). Also \(g\) is injective, thus \(a_1 = a_2\). QED.

Similar arguments work with surjectivity and bijectivity.

1.3 Some preliminaries

Before we dive into the study of the symmetric and dihedral groups, there are a few things we should get our heads around.

1.3.1 Order

The order of the group \(G\) is written \(|G|\) and defined to be the size of the set \(G\).

1.3.2 Subgroups

Let \((G,\ast)\) be a group. The subset \(H \subseteq G\) (i.e. \(x \in H \Rightarrow x \in G\)) is a **subgroup** if

(i) \(h_1, h_2 \in H \Rightarrow h_1 \ast h_2 \in H\)

(ii) \(e \in H\)

(iii) \(h \in H \Rightarrow h^{-1} \in H\)

(note that associativity is inherited). We write \(H \leq G\) to indicate that \(H\) is a subgroup. E.g.

\((\mathbb{Z}, +) \leq (\mathbb{Q}, +) \leq (\mathbb{R}, +)\)

\(((\pm 1), \times) \leq (\mathbb{Q} \setminus \{0\}, \times)\)
1.3.3 Cyclic groups

A group \( H \) is cyclic if there exists \( h \in H \) such that \( x = h^n \forall x \in H \) (\( h \) is called the generator of the group). Thus

\[
x^n = \begin{cases} 
  x \ast x \ast \ldots \ast x, & n > 0 \\
  e, & n = 0 \\
  x^{-1} \ast x^{-1} \ast \ldots \ast x^{-1}, & n < 0
\end{cases}
\]

E.g. \((\mathbb{Z}, +)\) is cyclic with generator 1. Also the group of rotations of a regular \( n \)-gon is cyclic (see below).

1.3.4 Homomorphisms and isomorphisms

Definition. Let \((G_1, \ast_1)\) and \((G_2, \ast_2)\) be groups. The function \( \theta : G_1 \rightarrow G_2 \) is a homomorphism if

\[
\theta(a \ast_1 b) = \theta(a) \ast_2 \theta(b)
\]

for any \( a, b \in G_1 \). An isomorphism is a bijective homomorphism.

We write \( G_1 \cong G_2 \) if there is an isomorphism \( G_1 \rightarrow G_2 \).

Examples. (i) \( \mathbb{Z}_n = \langle x | x^n = e \rangle \), i.e. the group generated by an \( x \) with this property.

(ii) Rotations of an \( n \)-gon \( \cong \left( \mathbb{Z} / n\mathbb{Z}, + \mod n \right) \) where \( \mathbb{Z} / n\mathbb{Z} = \{0, 1, \ldots, n - 1\} \), i.e. the group of rotations of a regular \( n \)-gon has an isomorphism to addition modulo \( n \). The isomorphism can be taken to be \( x^j \mapsto j \), which is a bijection, and then \( \theta(x^i, x^k) = \theta(x^{i+k}) = j + k = \theta(x^i) + \theta(x^k) \) which satisfies the definition.

Lemma 1.7. (i) The composition of homomorphisms is a homomorphism, and the same with isomorphisms.

(ii) If \( \theta : G_1 \rightarrow G_2 \) is an isomorphism, its inverse \( \theta^{-1} : G_2 \rightarrow G_1 \) is also one.

Hence;

\[
G \cong G_1 \cong G_2 \Rightarrow G_2 \cong G_1, \quad G_1 \cong G_2 \cong G_3 \Rightarrow G_1 \cong G_3
\]

Proof. (i). Let \( \theta : G_1 \rightarrow G_2 \) and \( \phi : G_2 \rightarrow G_3 \). Then \( \phi \circ \theta \) is a homomorphism, since

\[
(\phi \circ \theta)(a \ast_1 b) = \phi(\theta(a \ast_1 b)) = \phi(\theta(a) \ast_2 \theta(b)) = \phi(\theta(a)) \ast_2 \phi(\theta(b)).
\]

Lemma 1.8. If \( \theta : G \rightarrow H \) is a homomorphism then (i) \( \theta(e_G) = e_H \) and (ii) \( \theta(g^{-1}) = (\theta(g))^{-1} \).

Proof. (i). Let \( g \in G \). Then \( \theta(g) \ast_{e_H} e_H = \theta(g) = \theta(g \ast_{e_G} e_G) = \theta(g) \ast_{e_H} \theta(e_G) \) by the definition of a homomorphism. So, cancelling, \( \theta(e_G) = e_H \).

1.4 The symmetric and dihedral groups

For \( f : A \rightarrow B \), let \( A = B = X \), i.e. we are mapping from a set to itself. Bijective functions \( X \rightarrow X \) are called permutations of \( X \) (think shuffling).

1.4.1 The symmetric group

We define \( \text{Sym}(X) \) to be the set of all permutations. If \( X \) is finite of size \( |X| = n \), often we take \( X \) to be the set \( \{1, 2, \ldots, n\} \) and write \( S_n \) for \( \text{Sym}(X) \).

Theorem 1.5. \( \text{Sym}(X) \) is a group under composition. (It is called the symmetric group on \( X \).)

Proof. We prove each of the axioms in turn;

(i) If \( f, g \in \text{Sym}(X) \) then \( f \circ g \in \text{Sym}(X) \); this follows from 1.4.

(ii) Associativity; if \( f, g, h \in \text{Sym}(X) \), then for any \( x \in X \),

\[
(f \circ (g \circ h))(x) = (f \circ g)(h(x)) = f((g \circ h)(x)).
\]

However

\[
(f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x))
\]
\[(f \circ g) \circ h)(x) = ((f \circ g) \circ h(x)) = f\left(g(h(x))\right)\]

so the two are identical.

(iii) \(\iota: X \to X\) does nothing, thus \(f \circ \iota = f = \iota \circ f\)

(iv) Define \(f^{-1}: X \to X\) as follows, for \(y \in X\) let \(f^{-1}(y)\) be the unique pre-image of \(y\) in \(X\). Then \(f^{-1} \in \text{Sym}(X)\) and \((f^{-1} \circ f)(x) = x = \iota(x)\) for \(x \in X\). Also \((f \circ f^{-1})(x) = x = \iota(x)\), so the result follows since \(f, f^{-1}\) are bijective.

1.4.2 Permutation notation

It is customary to use small Greek letters to denote permutations; \(\iota\) represents the permutation which maps every element to itself (the 'identity'). For \(\sigma \in S_n\), we write

\[
\sigma = \begin{pmatrix}
1 & 2 & \cdots & n \\
\sigma(1) & \sigma(2) & \cdots & \sigma(n)
\end{pmatrix}
\]

It is easy to see inverses;

\[
\sigma^{-1} = \begin{pmatrix}
1 & 2 & \cdots & n \\
\sigma(1) & \sigma(2) & \cdots & \sigma(n)
\end{pmatrix}
\]

1.4.3 The elements of \(S_n\)

Examples. We can now write

\[
S_1 = \{\iota\} \quad S_2 = \{(1 2), (1 2)\} \quad S_3 = \{(1 2 3), (1 3 2), (1 2 3), (2 1 3), (2 3 1), (3 1 2), (3 2 1)\}
\]

In \(S_3\), if we let

\[
\sigma = \begin{pmatrix}
1 & 2 & 3 \\
1 & 2 & 3
\end{pmatrix}, \quad \tau = \begin{pmatrix}
1 & 2 & 3 \\
2 & 3 & 1
\end{pmatrix}
\]

then \(\sigma^2 = \iota, \tau^3 = \iota, \tau^2 = \tau^{-1}\), and in particular note that

\[
\sigma \circ \tau = \begin{pmatrix}
1 & 2 & 3 \\
1 & 2 & 1
\end{pmatrix}, \quad \tau \circ \sigma = \begin{pmatrix}
1 & 2 & 3 \\
2 & 1 & 3
\end{pmatrix}
\]

so \(\sigma \circ \tau \neq \tau \circ \sigma\), which means \(S_3\) is not abelian (this is true of \(S_n\) in general with \(n \geq 3\)).

We can write \(S_3 = \{\iota, \tau, \tau^2, \sigma, \sigma \tau, \sigma \tau^2\}\) so \(|S_3| = 6\). In general, \(|S_n| = n!\)

1.4.4 The dihedral group

Consider the group of symmetries of a regular \(n\)-gon, with \(n\) vertices (for example, the \(n\)th roots of unity in \(\mathbb{C}\)). This is a subgroup of \(S_n\), since

- the composition of symmetries is a symmetry
- the identity is a symmetry
- if \(\alpha\) is a symmetry then \(\alpha^{-1}\) is a symmetry

This is denoted \(D_{2n}\) (for \(n \geq 3\), \(D_{2n} \leq S_n\), for example, \(D_6 \leq S_3\) (as a matter of fact, \(D_6 = S_3\)).

**Theorem 1.6.** The group \(D_{2n}\) of symmetries of a regular \(n\)-gon \((n \geq 3)\) is a non-abelian subgroup of \(S_n\), of order \(2n\). It is generated by the elements \(\tau\) and \(\sigma\), of orders \(n\) and 2 respectively, representing a rotation and a reflection, subject to \(\tau^n = \iota, \sigma^2 = \iota, \sigma \tau \sigma = \tau^{-1}\).

**Proof.** First look at the \(n\) rotations; these are the set \(\{\iota, \tau, \tau^2, \ldots, \tau^{n-1}\}\) where

\[
\tau = \begin{pmatrix}
1 & 2 & \cdots & n \\
2 & 3 & \cdots & n
\end{pmatrix}
\]

so that \(\tau^n = \iota\). These form a subgroup of order \(n\), which is a cyclic group \(C_n\) of order \(n\), since each element is a power of \(\tau\) and \(\tau^i \tau^j = \tau^{i+j \pmod{n}}\).

Let \(\sigma\) be a reflection through the real axis. There are \(n\) reflections in the \(n\) different axes of symmetry. Therefore \(|D_{2n}| \geq 2n|.
We now need to show that any symmetry \( g \in D_{2n} \) is either of the form \( \tau^j \) for some \( j \) or of the form \( \sigma \tau^j \). Let \( j = g(1) \). Then as \( \tau^{j-1}(1) = j \), we have \( x = (\tau^{j-1})^{-1} g \) fixes 1. Since only two vertices, 2 and \( n \), are next to 1, either \( x \) fixes 2 and \( n \) as well (and therefore it fixes everything, by induction), so that \( g = \tau^{j-1} \), or \( x \) swaps 2 and \( n \). Now the above applies to \( \sigma^{-1} x \), so \( g = \tau^{j-1} \sigma = \sigma \tau^{j-1} \).

So any symmetry of our \( n \)-gon is either \( \tau^j \) for some \( j \), or \( \sigma \tau^j \) for some \( j \), so \( |D_{2n}| \leq 2n \Rightarrow |D_{2n}| = 2n \).

Algebraically, we write
\[
D_{2n} = \langle t, s | t^n = 1, s^2 = 1, sts = t^{-1} \rangle = \{t^j, st^j | 0 \leq j < n \}.
\]

1.5 More on permutations

1.5.1 Cycles

**Definition.** \( \sigma \in S_n \) is a \( k \)-cycle (notation: \( \sigma = (a_1 \ a_2 \ ... \ a_k) \)) if the following hold; \( \sigma(a_i) = a_{i+1} \); \( \sigma(a_k) = a_1 \); and \( \sigma(x) = x \) for all \( x \in S_n \) not one of the \( a_j \)’s.

**Remarks.** (i) \( (a_1 \ a_2 \ ... \ a_k) = (a_2 \ a_3 \ ... \ a_k \ a_1) = \cdots \) so we can cycle cycles and these cycles are not unique.

(ii) \( (a_1 \ a_2 \ ... \ a_k)^{-1} = (a_1 \ a_k \ a_{k-1} \ ... \ a_2) = (a_k \ a_{k-1} \ ... \ a_1) \)

(iii) \( \sigma^k = \iota \), but \( \sigma^l \neq \iota \) if \( 0 < l < k \), so \( k \) is the order of the \( k \)-cycle \( \sigma \).

1.5.2 Disjoint cycles

**Definition.** Two cycles \( \sigma = (a_1 \ ... \ a_k) \) and \( \tau = (b_1 \ ... \ b_l) \) are disjoint if all the \( a_j \)’s are different from all the \( b_j \)’s, i.e. the sets are non-intersecting.

**Lemma 1.9.** If \( \sigma, \tau \) are disjoint cycles as above, then \( \sigma \tau = \tau \sigma \). E.g. \( (1 \ 2)(3 \ 4 \ 5) = (3 \ 4 \ 5)(1 \ 2) \).

**Note.** This is only true of disjoint cycles, e.g. \( (2 \ 3)(1 \ 2 \ 3) = (1 \ 3) \neq (1 \ 2 \ 3)(2 \ 3) = (1 \ 2) \).

**Proof.** If \( x \in X \) is one of the \( a_j \)’s, then \( \tau \sigma(a_j) = \tau(a_j + 1) = a_j + 1 \) and \( \sigma \tau(a_j) = \sigma(a_j) = a_j + 1 \), and similarly if \( x \) is one of the \( b_j \)’s. If \( x \) is in neither set then \( \tau \sigma(x) = \tau(x) = x = \sigma \tau(x) \) by the definition. So \( \tau \sigma(x) = \sigma \tau(x) \) for all \( x \in X \). Thus \( \tau \sigma = \sigma \tau \).

1.5.3 Disjoint cycle decomposition

**Theorem 1.10.** Any permutation of a finite set \( X \) can be written as a product of disjoint cycles (in an essentially unique way). E.g.
\[
\begin{pmatrix}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
2 & 7 & 6 & 8 & 5 & 4 & 1 & 3
\end{pmatrix} = (1 \ 2 \ 7)(3 \ 6 \ 4 \ 8)(5)
\]

**Proof.** Let \( \pi \in S_n \). Start with any point \( a_1 \), and ‘chase it’ to get a sequence \( a_1, \pi(a_1), \pi^2(a_1), \ldots \). \( X \) is finite, so we have to get \( \pi^l(a_1) \) being equal to one of the previous entries, say \( \pi_j(a_1) \). If \( l \) is the smallest such that we get repetition and \( \pi^l(a_1) = \pi^j(a_1) \), then \( j = 0 \), so \( \pi^l(a_1) = a_1 \).

If not all points of \( X \) are covered, choose \( b_1 \) to be a point in \( X \) not yet reached. Chase it, to obtain \( b_2 = \pi(b_1) \), and in general, \( b_k = \pi^{k-1}(b_1) \) with \( \pi^k(b_1) = b_1 \).

Note that all the \( b_j \)’s are different from all the \( a_j \)’s, since \( \pi \) is injective. Continue until finished.

**Exercise.** Prove that, if \( (a_1 \ a_2 \ ... \ a_k)(b_1 \ b_2 \ ... \ b_k) \ldots (z_1 \ z_2 \ ... \ z_k) = \pi \), then the order of \( \pi \) is the least common multiple of all the \( k \)’s.

1.5.4 Transposition decomposition

**Definition.** A transposition on \( X \) is a 2-cycle on \( X \).

**Theorem 1.11.** Any permutation of finite \( X \) can be written as a product of transpositions.
Proof. Let \( \pi \in S_n \), and write it as a product of disjoint cycles. Now note that \((a_1 \ a_2 \ldots \ a_k) = (a_1 \ a_2)(a_3 \ a_4) \ldots (a_{k-1} \ a_k)\), so a \( k \)-cycle is a product of \( k - 1 \) transpositions.

**Example.** \((1 \ 2 \ 3 \ 4 \ 5) = (1 \ 2)(2 \ 3)(3 \ 4)(4 \ 5)\)

Note \( i = (1 \ 2)(1 \ 2) = (1 \ 2)(2 \ 3)(2 \ 3)(1 \ 2) \) so there are many representations of each permutation. Note that for \( i \) there must be an even number of factors; this is explained below.

### 1.6 Signs

#### 1.6.1 The sgn function

**Definition.** The sign of a permutation \( \sigma \) is \((-1)^k\) where \( k \) is the number of factors in some expression of \( \sigma \) as a product of transpositions.

**Lemma 1.12.** The function \( \text{sgn} : S_n \to \{\pm 1\} \) taking \( \sigma \mapsto \text{sgn} \sigma \) is well defined; that is, if \( \sigma = \tau_1 \tau_2 \ldots \tau_a = \tau'_1 \tau'_2 \ldots \tau'_b \), two lots of transpositions, then \((-1)^a = (-1)^b\).

**Proof.** Let \( c(\sigma) \) be the number of cycles in a disjoint cycle decomposition of \( \sigma \), including 1-cycles. E.g. \( c(i) = n \).

We claim that if \( \tau \) is any transposition then \( c(\sigma \tau) = c(\sigma) \pm 1 = c(\sigma) + 1 \text{ (mod 2)} \).

Write \( \tau = (k \ l) \). Then the \( \sigma \tau \) cycles are the same as the \( \sigma \) cycles, except for the one other involving \( k \) and \( l \). If \( k, l \) are in two different cycles of \( \sigma \), these two become one cycle of \( \sigma \tau \); if they are in the same cycle, then the cycle splits into two cycles of \( \sigma \tau \).

Now assume that \( \sigma = \tau_1 \tau_2 \ldots \tau_a = \tau'_1 \tau'_2 \ldots \tau'_b \). Since \( c(i) = n \), using the above result gives \( c(\sigma) = n + a \text{ (mod 2)} = n + b, \) so \((-1)^a = (-1)^b\).

**Theorem 1.13.** \( \text{sgn} : (S_n, \circ) \to (\{\pm 1\}, \times) \) taking \( \sigma \mapsto \text{sgn} \sigma \) is a well-defined, non-trivial homomorphism.

**Proof.** Well-defined by 1.13 and non-trivial due to the sign of any transposition being \(-1\).

If \( \text{sgn} \alpha = (-1)^k \) and \( \text{sgn} \beta = (-1)^l \), then write \( \alpha = \tau_1 \tau_2 \ldots \tau_k \) and \( \beta = \tau'_1 \tau'_2 \ldots \tau'_l \) where all the \( \tau \)s are transpositions. Then \( \alpha \beta = \tau_1 \tau_2 \ldots \tau_k \tau'_1 \tau'_2 \ldots \tau'_l \), so \( \text{sgn} \alpha \beta = (-1)^{k+l} = (-1)^k(-1)^l = \text{sgn} \alpha \text{ sgn} \beta \), satisfying the definition of a homomorphism.

#### 1.6.2 Even and odd permutations and the alternating group

**Definition.** \( \alpha \in S_n \) is an even permutation if its sign is \(+1\), and an odd permutation if its sign is \(-1\). Note that even \( \circ \) even = even, odd \( \circ \) odd = even, and the other compositions as expected.

**Corollary 1.14.** The even permutations in \( S_n \) form a subgroup of \( S_n \), called the **alternating group** and written \( A_n \). E.g.

\[
A_4 = \{i, (1 \ 2)(3 \ 4), (1 \ 3)(2 \ 4), (1 \ 4)(2 \ 3), (1 \ 2 \ 3), (1 \ 3 \ 2), (1 \ 2 \ 4), (1 \ 4 \ 2), (1 \ 3 \ 4), (1 \ 4 \ 3), (2 \ 3 \ 4), (2 \ 4 \ 3)\}.
\]

**Exercise.** Show that exactly half the permutations in \( S_n \) are even and thus are in \( A_n \), i.e. \( |A_n| = n!/2 \).

In fact, if \( H \leq S_n \), and \( H \) contains some odd permutation, then \( |A_n \cap H| = |H|/2 \).

**Remark 1.15.** Cycles of even length are odd permutations whereas cycles of odd length are even permutations (see 1.11).

**Lemma 1.16.** \( \alpha \) is an odd permutation iff the number of cycles of **even length** in its disjoint-cycle representation is odd.

#### 1.6.3 Alternative treatment of sgn (non-examinable)

**Definition.** Given \( \alpha \in S_n \), let \( x_1, x_2, \ldots, x_n \) be distinct integers. Then the sign of \( \alpha \) is

\[
\varepsilon(\alpha) = \prod_{1 \leq i < j \leq n} \frac{x_{\alpha(j)} - x_{\alpha(i)}}{x_j - x_i}.
\]
Example. Let $\sigma = (1\ 2\ 3)$ and $x_j = j$, then
\[
\varepsilon(\sigma) = \frac{3 - 2}{2 - 1} \times \frac{1 - 2}{3 - 1} \times \frac{1 - 3}{3 - 2} = +1.
\]

**Lemma 1.17.** $\varepsilon(\sigma) = \pm 1$, independent of the actual numbers $x_j$. In fact, $\varepsilon(\sigma) = (-1)^{N(\sigma)}$, where $N(\sigma) = \{i < j \mid \sigma(i) > \sigma(j)\}$.

**Proof.** For each $r \neq s$, either $x_r - x_s$ or $x_s - x_r$ appears on the top of the fraction. If $i = \sigma^{-1}(r) < \sigma^{-1}(s) = j$, then $x_s - x_r$ appears; if vice versa, then $x_r - x_s$ appears.

**Lemma 1.18.** For any permutations $\sigma, \tau \in S_n$, we have $\varepsilon(\sigma \tau) = \varepsilon(\sigma) \varepsilon(\tau)$.

**Proof.**
\[
\varepsilon(\sigma \tau) = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i} = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i} \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\sigma(j) - \sigma(i)} = \varepsilon(\sigma) \varepsilon(\tau)
\]

since the product on the right, taking $x_i = \sigma(i)$, gives us $x_{\tau(i)} = \sigma(\tau(i))$.

**Lemma 1.19.** If $\tau$ is a transposition then $\varepsilon(\tau) = -1$.

**Proof.** Let $\tau = (1\ 2)$, $x_j = j$. Then
\[
\varepsilon(\tau) = \frac{1 - 2}{2 - 1} \times \frac{3 - 2}{3 - 1} \times \ldots = -1
\]

If $\tau = (1\ 1)$ then write it as $(2\ 1)(1\ 2)(2\ 1)$ and cancel the end terms. If $\tau = (k\ l)$ it can be written in the form $(1\ k)(1\ l)(1\ k)$. Thus $\varepsilon(k\ l) = \varepsilon(1\ 2)$ and the lemma is proved.

**Corollary 1.20.** $\varepsilon(\sigma) = \text{sgn}\ \sigma$ for all $\sigma \in S_n$.

### 2. Subgroups, cosets and Lagrange’s theorem

#### 2.1 Subgroups

**Recall.** Let $G$ be a group. The subset $H \subseteq G$ (i.e. $x \in H \Rightarrow x \in G$) is a **subgroup** if

(i) $h_1, h_2 \in H \Rightarrow h_1 h_2 \in H$

(ii) $e_G \in H$

(iii) $h \in H \Rightarrow h^{-1} \in H$

**Example.** $A_n \leq S_n$.

**Remark.** If $G$ is finite and $H$ is non-empty, then condition (i) is sufficient.

**2.1.1 The single subgroup condition**

**Lemma 2.1.** The subset $H$ of $G$ is a subgroup if $H$ is non-empty and $a, b \in H \Rightarrow a^{-1}b \in H$.

**Proof.** Let $h, k \in H$. Then letting $a = b = h$, we obtain $h^{-1}h \in H$, and this is the identity, so $e \in H$. Next, letting $a = h, b = e$, we obtain $h^{-1}e = h^{-1} \in H$. Finally, letting $a = h^{-1}, b = k$, we obtain $(h^{-1})^{-1}k = hk \in H$. Thus all three conditions are satisfied and $H$ is a subgroup.

**2.1.2 Generation**

**Definition.** If $g_1, \ldots, g_k \in G$, the subgroup generated by the elements $g_1, \ldots, g_k$ is the smallest subgroup of $G$ containing all the $g_j$s. It is denoted by $\langle g_1, \ldots, g_k \rangle$ and is the intersection of all the subgroups of $G$ that contain all the $g_j$s.

If there is no proper subgroup of $G$ containing all the $g_j$s, then the subgroup generated by $g_1, \ldots, g_k \in G$ is $G$ itself.

**Exercise.** Show that $S_n$ is generated by each of
2.1.3 Example

Example. $S_3$ has order 6. Its subgroups are:

- $\{i\}$, of order 1
- $\langle(1 2)\rangle = \{i, (1 2)\}$, of order 2, and two others which are similar; $\langle(1 3)\rangle$ and $\langle(2 3)\rangle$
- $\langle(1 2 3)\rangle = \{i, (1 2 3), (1 3 2)\}$, of order 3
- $S_3$, of order 6

2.2 Cosets

Definition. Let $g \in G$; the left coset $gH$ of $H$ in $G$ is defined as

$$gH = \{gh | h \in H\}.$$

Example. $A_n \leq S_n$ has precisely two different cosets, $H$ and $\tau H$, with $\tau$ any odd permutation, such as $(1 2)$.

2.2.1 The crucial coset characteristic

Let $H \leq G$. Then $H = eH = h_1H$ for some $h_1 \in H$, so there are many ways of writing a coset. In fact, $gH = H$ iff $g \in H$. Also, $[aH = bH \iff a^{-1}b \in H]$ (see 2.1), and this is the defining characteristic of a coset.

Proof. If $aH = bH$, then $be = ah$ for some $h \in H$, so $a^{-1}b \in H$. Conversely, if $a^{-1}b \in H$, say $b = ah_1$ for some $h_1 \in H$, then $b \in bH \cap aH$ and $aH = bH$.

Example. $G = S_3$ and $H = \langle(1 2)\rangle$. Now $|H| = 2$ and $|G| = 6$, so we are expecting 3 left cosets.

$$H = \{i, (1 2)\}, \quad (1 2 3)H = \{(1 2 3), (1 3)\}, \quad (1 3 2)H = \{(1 3 2), (2 3)\}$$

2.2.2 Equal size and partitioning of cosets

Lemma 2.3. All the left cosets of $H$ in $G$ have size equal to $|H|$.

Proof. If $gH$ is a coset, then consider the map $H \rightarrow gH$ taking $h \mapsto gh$. Now if $gh_1 = gh_2$ then $h_1 = h_2$ so this map is injective. If $y \in gH$ then $y = gh$ for some $h \in H$, so $y$ is the image of $h$; therefore the map is surjective. Hence the map is a bijection, so $|H| = |gH|$.

Lemma 2.4. The left cosets form a partition of $G$, which means;

(i) each $g \in G$ lies in some left coset
(ii) if $aH \cap bH$ is non-empty, then $aH = bH$ (i.e. any two cosets overlap totally or not at all).

Proof (i). $g = ge \in gH$.

Proof (ii). It is sufficient to prove that if $c \in aH$ then $cH = aH$. Now, if $c \in aH$, then $c = ah_1$ for some $h_1 \in H$. For any $h \in H$, we have $ch = (ah_1)h \in aH$, because $h_1h \in H$. Therefore $cH \subseteq aH$.

However, also, $a = ch_1^{-1} \in cH$, so by the same reasoning, $aH \subseteq cH$. Therefore $aH = cH$.

2.2.3 Right cosets

The right coset $Hg$ of $H$ in $G$ is defined by $Hg = \{hg | h \in H\}$. The distinct right cosets of $H$ all have size $|H|$, and form a partition of $G$.

Remark. If $H \leq G$, precisely one of the cosets is a subgroup, because only one can contain the identity.

Corollary 2.8. The number of distinct left cosets of $H$ in $G$ is equal to the number of distinct right cosets of $H$ in $G$.

Example. $G = S_3$ and $H = \langle(1 2)\rangle$ as above.
\[ H = \{i, (1 2)\}, \quad H(1 2 3) = \{(1 2 3), (2 3)\}, \quad H(1 3 2) = \{(1 3 2), (1 3)\} \]

**Exercise.** Write down a bijection between the left and right cosets of \( H \leq G \).

### 2.3 Lagrange’s theorem

**Theorem 2.2 (Lagrange’s theorem).** If \( H \) is a subgroup of the finite group \( G \), then the order of \( H \) divides the order of \( G \).

**Proof.** We have done all the hard work already. From Lemma 2.4, if \( c \in (aH \cap bH) \), then \( aH = cH = bH \). Hence \( G \) is covered by the distinct cosets of \( H \), which are all of size \( |H| \). Therefore \( |H| | |G| \).

#### 2.3.1 Index of a subgroup

**Definition.** The index of \( H \) in \( G \) is the number of distinct left cosets of \( H \) in \( G \).

Note that if \( G \) is finite, the index is \( [G] / |H| \) by Lagrange’s theorem. (This is also written as \( |G : H| \), with \( (G : H) \) used to denote the set of all left cosets of \( H \) in \( G \)).

Note. \( A_4 \) has order 12 but no subgroup of order 6, so the converse is not true. A later exercise will be to show that \( A_5 \) has no subgroup of index 2, 3 or 4.

#### 2.3.2 The order of an element

**Definition.** The order of an element \( g \in G \), written \( o(g) \), is the least positive integer \( n \) (if one exists) such that \( g^n = e_G \). If no such integer exists, then we say that \( g \) has infinite order.

**Lemma 2.5.** If \( o(g) = n \) then \( g^m = e \) iff \( n|m \).

**Proof (\( \Rightarrow \)).** If \( g^m = e \), then write \( m = qn + r \) with \( q, r \in \mathbb{Z}, 0 \leq r < n \). Then \( g^r = g^{m((g^{-1})^n)^q} = e \), but \( r < n \) and \( n \) is the smallest positive integer such that \( g^n = e \). Therefore \( r = 0 \), which means \( n|m \).

**Proof (\( \Leftarrow \)).** If \( m = qn \), then \( g^m = (g^n)^q = e \).

#### 2.3.3 Lagrange’s corollary

**Theorem 2.6 (Lagrange’s corollary).** If \( G \) is a finite group, and \( g \in G \), then \( o(g) | |G| \). Thus \( g^{|G|} = e \) for all \( g \in G \).

**Proof.** Consider the subgroup \( \langle g \rangle = \{e, g, g^2, ..., g^{n-1}\} \) where \( n = o(g) \). So \( o(g) = |\langle g \rangle| | |G| \) by Lagrange’s theorem.

**Corollary 2.7.** If \( G \) is a group of prime order \( p \), then \( G \) is cyclic. In fact, any element \( g \in G \setminus \{e\} \) generates \( G \).

**Proof.** Take \( g \in G \setminus \{e\} \). Then \( \langle g \rangle \leq G \) of an order dividing \( p \), so \( G = \langle g \rangle \). Thus \( G \simeq C_p \).

#### 2.3.4 Equivalence relations and Lagrange’s theorem

Let \( G \) be a group and \( H \leq G \). Note that \( [a \equiv b \text{ if } a^{-1}b \in H \text{ for } a, b \in G] \) is an equivalence relation on \( G \). Check this:

- \( a \equiv a \), since \( a^{-1}a = e \in H \)
- \( a \equiv b \Rightarrow b \equiv a \), since \( a^{-1}b \in H \Rightarrow (a^{-1}b)^{-1} \in H \Rightarrow b^{-1}a \in H \)
- \( a \equiv b, b \equiv c \Rightarrow a \equiv c \), since \( a^{-1}b \in H \) and \( b^{-1}c \in H \) imply \( a^{-1}c \in H \) by Lemma 2.4.

Now if \( [a] = \{b : a \equiv b\} \) is the equivalence class of \( a \), then \( [a] = aH \), the left coset, since \( [a] = \{b : a^{-1}b \in H\} = \{b : b \in aH\} \). Therefore the left cosets form a partition of \( G \).

### 2.4 An application of Lagrange’s theorem: Fermat’s little theorem

#### 2.4.1 Working mod \( n \)

Let \( n \in \mathbb{N} \). Then \( \mathbb{Z}_n \) (note the lack of \(*\)) = \{0, 1, ..., n -1\} = a cyclic group under + mod \( n \).
For $a \in \mathbb{Z}$, define $r_n(a)$ to be the remainder in $R_n$ (also known as the residue mod $n$); that is, if $\exists q \in \mathbb{Z}, r \in R_n$ with $a = qn + r$, then $r_n(a) = r$.

We see that there is a function $r_n : \mathbb{Z} \rightarrow R_n$ which takes $a \mapsto r_n(a)$. This is a surjective function, and in fact, a homomorphism from $(\mathbb{Z}, +) \rightarrow (R_n, +_n)$.

2.4.2 Working with numbers coprime to $n$

Let $n \in \mathbb{N}$. Define $R_n^* = \{a \in R_n \mid \text{hcf}(a, n) = 1\}$. Define the operation $\otimes_n$ to be multiplication mod $n$, that is, $a \otimes_n b = r_n(ab)$ (see above for the definition of this).

**Lemma 2.11.** $R_n^*$ is an abelian group with operation $\otimes_n$.

**Proof.**
- Closure; if $\text{hcf}(a, n) = 1 = \text{hcf}(b, n)$, then $\text{hcf}(ab, n) = 1$, so $(a \otimes_n b) \in R_n^*$.
- Abelian; multiplication in $\mathbb{Z}$ is abelian, so $a \otimes_n b = b \otimes_n a$.
- Identity is 1, since $a \otimes_n 1 = a \forall a \in R_n^*$.
- Associativity; multiplication in $\mathbb{Z}$ is associative.
- Inverses; given $a \in R_n^*$, $\exists u, v \in \mathbb{Z}$ with $au + nv = 1$ (see Numbers and Sets). So $a \otimes_n u = 1 \mod n$. Then $r_n(u)$ is the inverse of $a$ under $\otimes_n$.

**Remark.** The order of $R_n^*$ is $\phi(n)$, Euler’s totient function, defined by

$$\phi(n) = \{1 \leq a \leq n \mid \text{hcf}(a, n) = 1\}$$

Note that $\phi(p) = p - 1$ for prime $p$.

2.4.3 The Fermat-Euler theorem

**Theorem 2.12 (Fermat-Euler Theorem).** If $n \in \mathbb{N}, a \in \mathbb{Z}$ with $a$ coprime to $n$, then $a^{\phi(n)} \equiv 1 \mod n$.

**Proof.** $r_n(a^{\phi(n)}) = 1$ in $R_n^*$ by theorem 2.6, because $|R_n^*| = \phi(n)$. Hence $a^{\phi(n)} \equiv 1 \mod n$.

(There is another proof in Numbers and Sets.)

**Theorem 2.9 (Fermat’s Little Theorem).** If $p$ is a prime and $a \in \mathbb{Z}$ with $p \nmid a$, then $a^{p-1} \equiv 1 \mod p$.

**Proof.** This is just the special case of 2.12 when $n$ is prime.

3. Small groups

3.1 Cyclic groups

**Recall.** The group $G$ is cyclic if $\exists x \in G$ such that every element of $G$ is a power of $x$.

3.1.1 Elements of cyclic groups

**Lemma 3.1.** Let $G$ be cyclic and generated by $x$. Then either

- (i) $G$ is infinite, all the $x^i$s are distinct, and $G = (\mathbb{Z}, +)$; or
- (ii) $G$ is finite with order $n$ for some $n \in \mathbb{N}$, and $G = \{e, x, x^2, ..., x^{n-1}\}$ with $G = C_n = (R_n, +_n) \simeq (\mathbb{Z}/n, +)$.

**Proof (i).** Assume $G$ is infinite. Note that if $x^k = x^l$ for some $k, l \in \mathbb{Z}$, then $x^{k-l} = e$. If $k > l$, we get $G = \{e, x, x^2, ..., x^{k-l-1}\}$, which is not infinite, and similarly if $k < l$. So if $G$ is infinite and $x^k = x$ then $k = l$, so all the $x^i$s are distinct. Also, the mapping $\theta : G \rightarrow (\mathbb{Z}/n, +)$ taking $x^i \mapsto j$ is an isomorphism; it is bijective, and $\theta(x^ix^k) = \theta(x^{i+k}) = j + k = \theta(x^j) + \theta(x^k)$.

**Proof (ii).** If $G$ is finite and $|G| = n$, then $\phi(x) = n$, so $G = \{e, x, ..., x^{n-1}\}$. Also, the map $\theta : G \rightarrow (\mathbb{Z}/n, +)$ taking $x^j \mapsto j$ is an isomorphism, as follows. If $x^{j_1} = x^{j_2}$, then $x^{j_1-j_2} = e$, so $n \mid j_1 - j_2$ by Lemma 2.5. Thus $[j_1] = [j_2]$ in $\mathbb{Z}/n$, so the map is well-defined. It is also bijective and a homomorphism (check it!)
3.1.2 Subgroups of cyclic groups

**Lemma 3.2.** Any subgroup $H$ of a cyclic group $G = \langle x \rangle$ is cyclic. In fact, if $H \neq \{e\}$, then $H = \langle x^k \rangle$ with $k$ the smallest positive integer such that $x^k \in H$.

**Proof.** See example sheet 2, question 6.

**Remark 3.3.** Let $a, b \in \mathbb{Z}$. By 3.2, the subgroup of $\mathbb{Z}$ generated by $a, b$ is cyclic, say $\langle c \rangle$ for some $c \in \mathbb{Z}$. Note that $c = \text{hcf} (a, b)$. Then $c \in \langle a, b \rangle$, so $c = a^u b^v$ in $\mathbb{Z}$ (since $\mathbb{Z}$ is abelian). Note; multiplication here is actually addition since the operation in $\mathbb{Z}$ is $+$, so we can write $c = ua + vb$.

3.2 Direct (Cartesian) products of groups

Given groups $H$ and $K$, we can construct the direct product $H \times K = \{h, k \mid h \in H, k \in K\}$ with operation $(h_1, k_1)(h_2, k_2) = (h_1 h_2, k_1 k_2)$.

**Lemma 3.4.** If $H$ and $K$ are groups then $H \times K$ is a group.

**Proof.**
- Closed under the given operation, as $h_1 h_2 \in H$ and $k_1 k_2 \in K$.
- Associative since multiplication in $H$ and $K$ is associative.
- Identity $e = (e_H, e_K)$.
- Inverse of $(h, k) = (h^{-1}, k^{-1})$.

3.2.1 Examples and notes

**Example (Klein 4-group).** $C_2 \times C_2 = \{e, x\} \times \{e, y\} = \{(e, e), (x, e), (y, e), (x, y)\}$ where all non-identity elements have order 2.

Alternatively, $C_2 \times C_2 \cong \mathbb{Z}/2 \times \mathbb{Z}/2 = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$ under addition mod 2.

Now $C_2 \times C_3 \cong C_6$ since $C_2$ is generated by $x$ of order 2 and $C_3$ is generated by $y$ of order 3, so $o(x, y) = 6$. In general, $C_m \times C_n \cong C_{mn}$ iff $\text{hcf}(m, n) = 1$.

It follows that $\mathbb{R}^n$ for $n \in \mathbb{N}$ is a group under component-wise addition, since $\mathbb{R}^n = \mathbb{R} \times ... \times \mathbb{R}$. $n$ times

**Notes.**
1. If $H, K$ are finite, then so is $H \times K$, and $|H \times K| = |H||K|$.
2. $H, K$ are abelian iff $H \times K$ is abelian.
3. $H \times K$ contains a subgroup $\{(h, e) : h \in H\}$ isomorphic to $H$ and a subgroup $\{(e, k) : k \in K\}$ isomorphic to $K$.

3.2.2 Full product conditions

**Proposition 3.5.** Let $G$ be a group with subgroups $H \times K$ such that
1. each element can be written as $hk$ with $h \in H, k \in K$
2. $H \cap K = \{e\}$
3. $hk = kh \forall h \in H, k \in K$

Then $G \cong H \times K$.

**Proof.** If $h_1 k_1 = h_2 k_2$ with $h_i \in H, k_i \in K$, then $h_1^{-1} h_2 = k_2 k_1^{-1} \in H \cap K = \{e\}$, so expressions $hk$ for elements of $G$ are unique. Thus we can define $\theta : G \to H \times K$, taking $g = hk \mapsto (h, k)$. This is well defined, by uniqueness of expressions $hk$, and is bijective. By condition (iii),

$$\theta(h_1 k_1 h_2 k_2) = \theta(h_1 h_2 k_1 k_2) = (h_1 h_2, k_1 k_2) = (h_1, k_1)(h_2, k_2) = \theta(h_1 k_1)\theta(h_2 k_2)$$

so the map $\theta$ is a homomorphism.

3.3 Dihedral groups

**Proposition 3.6.** Let $G$ be a group generated by two elements, $s$ of order $n$ and $t$ of order 2, with $tst = s^{-1}$. Algebraically, $G = \langle s, t \mid s^n = e, t^2 = e, tst = s^{-1} \rangle$. Then $G$ is dihedral of order $2n$, that is, $G$ is isomorphic to the group of symmetries of a regular $n$-gon.
Proof. \( G = \{e, s, s^2, \ldots, s^{n-1}, t, ts, ts^2, \ldots, ts^{n-1}\} \) and these are precisely the elements of \( G \).
(Note Saxl’s Theorem, the very useful fact that \( s^l t = ts^{-l} \)).

Now the group of symmetries of a regular \( n \)-gon contains \( \sigma \), a rotation of order \( n \), and \( \tau \), a reflection satisfying \( \tau^2 = e \) and \( \tau \sigma = \sigma^{-1} \). So the mapping \( \theta : \tau^k s^l \mapsto \tau^k \sigma^l \) is a well-defined bijective homomorphism, i.e. an isomorphism, since both groups satisfy the same conditions.

### 3.4 The groups of order \( \leq 8 \)

Here is a table of them. For groups of prime order we use 2.7. For other orders see the sections below.

<table>
<thead>
<tr>
<th>order</th>
<th>groups (up to isomorphism)</th>
<th>number of these</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( {e} )</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>( C_2 )</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>( C_3 )</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>( C_4, C_2 \times C_2 ) (see 3.7)</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>( C_5 )</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>( C_6 ; D_6 ) (see 3.8)</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>( C_7 )</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>( C_8, C_4 \times C_2, C_2 \times C_2 \times C_2 ; D_8, Q_8 ) (see 3.9)</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>( C_9, C_3 \times C_3 ) (method similar to 3.7)</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>( C_{10} ; D_{10} ) (see example sheet 2, Q9)</td>
<td>2</td>
</tr>
</tbody>
</table>

Note. There are 10 groups of order 16, and \( \approx 50,000,000,000 \) groups of order \( 2^{10} \).

#### 3.4.1 Groups of order 4

**Lemma 3.7.** A group of order 4 is abelian, and is isomorphic to \( C_4 \) or \( C_2 \times C_2 \).

**Proof.** Non-identity elements have order 2 or 4. If \( G \) contains an element of order 4, then \( G \cong C_4 \). So assume all non-identity elements have order 2. Let \( a \in G \setminus \{e\} \) and let \( b \in G \setminus \{a\} \). Then \( G = \{e, a, b, ab\} \). So \( G \cong \langle a \rangle \times \langle b \rangle = C_2 \times C_2 \) by 3.5.

#### 3.4.2 Groups of order 6

**Lemma 3.8.** A group \( G \) of order 6 is isomorphic to either \( C_6 \) or \( D_6 \).

**Proof.** If all the elements of \( G \) have order 2, then take \( a, b \) of order 2 and observe that \( |\langle a, b \rangle| = 4 \mid 6 \). If all the elements have order 3, then consider \( a, b \in G \) with \( \langle a \rangle \neq \langle b \rangle \). Thus \( G \cong \{e, a, a^2, b, b^2, ab, ab^2, \ldots\} \) which is too big. Therefore \( G \) contains an element \( s \) of order 3 and an element \( t \) of order 2.

Thus \( G = \{e, s, s^2, t, ts, ts^2\} \). Now \( st \) must be \( ts \) or \( ts^2 \) (the other possibilities lead to contradictions). If \( ts = st \), then \( G \cong C_6 \) as \( o(st) = 6 \). If \( st = ts^2 \), then \( G \) is generated by \( t, s \) of orders 2 and 3 with \( tst = s^{-1} \), so \( G \cong D_6 \).

#### 3.4.3 Groups of order 8

**Lemma 3.9.** If \( G \) is a group of order 8, then \( G \) is isomorphic to exactly one of \( C_8, C_4 \times C_2, C_2 \times C_2 \times C_2, D_8 \) and \( Q_8 \).

**Proof.** The orders of non-identity elements are 2, 4 and 8 by Lagrange’s corollary.

- If \( G \) contains an order-8 element, then \( C \cong C_8 \). So assume none such.
- If all the elements of \( G \setminus \{e\} \) have order 2, then \( G \) is abelian since \( ba = (ba)^{-1} = a^{-1}b^{-1} = ab \). Take \( a \in G \setminus \{e\} \), \( b \in G \setminus \{a\} \) and \( c \in G \setminus \{b\} \). Then \( |\langle a, b \rangle| = 4 \), and is isomorphic to \( C_2 \times C_2 \), and \( G \cong \langle a \times b \rangle \times \langle c \rangle = C_2 \times C_2 \times C_2 \) by 3.5.
So there is an element \( a \) of order 4 in \( G \) but none of order 8. Let \( b \in G \setminus \langle a \rangle \) with order 2 or 4. Hence \( G = \{ e, a, a^2, a^3, b, ab, a^2b, a^3b \} \) with no more elements. What are \( b^2 \) and \( ba \)?

- Now \( b^2 \in \langle a \rangle \). If \( b^2 \) is \( a \) or \( a^3 \), then \( o(b) = 8 \) which cannot be. So \( b^2 = e \) or \( b^2 = a^2 \).
- Now \( ba = a^j b \) for some \( j \in \{ 1, 2, 3 \} \). So \( bab^{-1} = a^j \). Note that \( a = b^2 ab^{-2} = b(bab^{-1})b^{-1} = ba b^{-1} = (bab^{-1})^j = (a^j)^j = a^{j^2} \). Therefore \( a^{j^2 - 1} = e \), so 4 divides \( j^2 - 1 \) (using 2.5), so \( j \) is odd. So \( j \in \{ 1, 3 \} \) and \( bab^{-1} = a \) or \( a^3 \).

- Case 1; \( bab^{-1} = a \). Then \( ba = ab \) so \( G \) is abelian.
  - If \( b^2 = e \), then \( G \simeq \langle a \rangle \times \langle b \rangle \) by 3.5 \( \simeq C_4 \times C_2 \).
  - If \( b^2 = a^2 \), replace \( b \) by \( b' = ab^{-1} \). Then \( (ab^{-1})^2 = a^2b^{-2} = e \). So \( G \simeq \langle a \rangle \times \langle b' \rangle \simeq C_2 \times C_4 \).

- Case 2; \( bab^{-1} = a^{-1} \).
  - If \( b^2 = e \), then \( G = D_8 \) by 3.6. \( G \) then has two elements of order 4, \( a \) and \( a^{-1} \), and five elements of order 2, \( a^2, ab, b, a^2b \) and \( a^3b \).
  - If \( b^2 = a^2 \), then define \( G \).

3.4.4 The quaternion group \( Q_8 \)

This is defined as \( \langle a, b \mid a^4 = e, b^2 = a^2, bab^{-1} = a^{-1} \rangle \). It has one element of order 2, \( a^2 \), and six elements of order 4, all the others.

The customary notation for \( Q_8 \) is \( \{ \pm 1, \pm i, \pm j, \pm k \} \) with the operation as the vector cross-product. So \( ij = k, jk = i, ki = j \) but \( ji = k, ik = j, ik = j \) and \( i^2 = j^2 = k^2 = -1 \).

Another realisation is 2 \( \times \) 2 matrices over \( \mathbb{C} \);

\[
\pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \pm \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \pm \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}
\]

corresponding to the \( 1, i, j, k \) from above.

It follows that there is a 4-dimensional, non-commutative, associative algebra over \( \mathbb{R} \). \( \mathbb{H} = \{ \alpha 1 + \beta i + \gamma j + \delta k \} \) where \( \alpha, \beta, \gamma, \delta \in \mathbb{R} \).

4. Group actions

4.1 Introduction to actions

4.1.1 Definition

Definition. Let \( G \) be a group and \( X \) a non-empty set. We say that \( G \) acts on \( X \) if there is a mapping \( \rho : G \times X \to X \), taking \( (g, x) \mapsto \rho(g, x) = g(x) \), such that

0) if \( g \in G, x \in X \) then \( g(x) \in X \)

(i) \( (gh)(x) = g(h(x)) \) \( \forall h, g \in G, x \in X \)

(ii) \( e(x) = x \) \( \forall x \in X \)

4.1.2 Actions and permutations

Lemma 4.1. Fix \( g \in G \); then we have a map \( \varphi_g : X \to X \) such that \( x \mapsto g(x) \). Then \( \varphi_g \) is a permutation of \( X \), so \( \varphi_g \in \text{Sym}(X) \).

Proof. This is certainly a mapping \( X \to X \). For \( x \in X \), \( \varphi_{g^{-1}} \circ \varphi_g = \varphi_{g^{-1}}(g(x)) = g^{-1}(g(x)) = (g^{-1}g)(x) \) by condition (i) = \( e(x) = x \) by condition (ii), and similarly in reverse. Thus we have inverses, which means the map is bijective, so it is a permutation.

Lemma 4.2. Let \( \rho \) be an action of \( G \) on a set \( X \). Then the mapping \( \varphi : G \to \text{Sym}(X) \) taking \( g \mapsto \varphi_g \), with \( \varphi_g(x) = g(x) \), is a homomorphism – called a permutation representation of \( G \).
Proof. \( \varphi_g \) is a permutation of \( X \) by the previous lemma, so \( \varphi : G \to \text{Sym}(X) \). For \( g_1, g_2 \in G \) and for all \( x \), \( \varphi_{g_1 g_2}(x) = (g_1 g_2)(x) = g_1( g_2(x) ) = \varphi_{g_1} \left( \varphi_{g_2}(x) \right) = (\varphi_{g_1} \circ \varphi_{g_2})(x) \). Thus the map is a homomorphism.

Note. \( G \)-actions on \( X \) $\leftrightarrow$ homomorphisms from \( G \to \text{Sym}(X) \).

Example. \( G = D_8 \) = the group of symmetries of a square. Then if \( X \) is the set of vertices and \( Y \) the set of edges, \( G \) acts on both \( X \) and \( Y \). Both give an injective homomorphism into \( S_4 \).

4.2 The orbit-stabiliser theorem

4.2.1 Orbits and transitivity

Definition. \( G \) acting on \( X \) is transitive if for every pair \( x_1, x_2 \in X \) there is an element \( g \in G \) with \( g(x_1) = x_2 \).

Definition. Let \( G \) act on \( X \) and let \( x \in X \). The orbit of \( G \) on \( X \) containing \( x \) is \( G(x) = \{ g(x) \mid g \in G \} \). This is a subset of \( X \) on which \( g \) acts transitively.

Lemma 4.4. Each \( G \)-orbit on \( X \) is \( G \)-invariant, with \( G \) transitive in its action on \( G(x) \), for \( x \in X \). The distinct \( G \)-orbits form a partition of \( X \).

Proof. The orbit \( G(x) \) is \( G \)-invariant, so that \( g(G(x)) = G(x) \) for \( g \in G \). So the action of \( G \) on \( X \) induces an action on \( G(x) \), which is transitive. It follows that we can define an equivalence relation on \( X \) given by \( x_1 \sim x_2 \) if \( \exists g \in G : x_2 = g(x_1) \).

Example. The left regular action (see section 4.5) is transitive on \( X = G \), because if \( x_1, x_2 \in G \) then we can set \( g = x_2 x_1^{-1} \) to get \( (x_2 x_1^{-1}) x_1 = x_2 \).

4.2.2 Stabilisers

Definition. If \( G \) acts on \( X \), and \( x \in X \), the stabiliser of \( x \) in \( G \) is \( G_x = \{ g \in G : g(x) = x \} \).

Lemma 4.5. \( G_x \) is a subgroup of \( G \).

Proof. \( e \in G_x \), obviously. If \( g_1, g_2 \in G_x \) then so is \( g_1^{-1} g_2 \), because \( (g_1^{-1} g_2)(x) = g_1^{-1}(g_2(x)) = g_1^{-1}(x) = x \).

4.2.3 The orbit-stabiliser theorem

Theorem 4.6 (Orbit-stabiliser theorem). Let a finite group \( G \) act on \( X \) and let \( x \in X \). Then \( G_x \subseteq G \), and \( |G| = |G_x||G(x)| \); that is, \( |G : G_x| \), the number of left cosets of \( G_x \) in \( G \), is equal to \( |G(x)| \).

Proof. Define \( (G : G_x) \) to be the set of left cosets of \( G_x \) in \( G \). Then all that is needed is to prove that the mapping \( G(x) \to (G : G_x) \) taking \( g(x) \to gG_x \) is a well-defined bijection.

This can be proved as follows. If \( g_1(x) = g_2(x) \) then \( g_2^{-1} g_1(x) = x \), so \( g_2^{-1} g_1 \in G_x \), so \( g_2 G_x = g_1 G_x \), so the map is well-defined. To prove injectivity, reverse the steps (carefully). For surjectivity, note that any coset of \( G_x \) is \( g G_x \) for some \( g \in G \), so \( g G_x \) is the image of \( g(x) \). QED.

4.2.4 Examples including the left coset action

Example. If \( G \) is the group of all symmetries of a regular \( n \)-gon (see 1.6), then \( |G| \geq 2n \) because there are \( n \) rotations and \( n \) reflections. Let \( X = \{ \text{vertices of the } n \text{-gon}\} \). \( G \) acts on \( X \) and is transitive (e.g. by rotations). Then considering the point 1, we get \( |G| = |X||G_1| = n |G_1| = n \times 2 \) because the stabiliser of the point 1 contains just a reflection and the identity.

Example. Let \( G \) be a group, \( H \) a subgroup of \( G \) and \( X = (G : H) \) the set of left cosets of \( G_x \) in \( G \). Then \( (G : H) \) acts on \( X \) as follows; \( g(xH) = (gx)H \) with \( g, x \in G \) so that \( xH \in X \). This is called the left coset action of \( G \).

This action is transitive; \( (x_2 x_1^{-1})(x_1 H) = x_2 H \) where \( x_2 x_1^{-1} g = g \in G \).

The stabiliser of the element \( H \) in \( X \) is \( H \) itself; \( g(H) = gH \), so \( g(H) = H \) iff \( g \in H \).
By the orbit-stabiliser theorem, $|G|/|G_x| = |G : H|$.

### 4.2.5 The kernel

Let $G$ act on $X$. Then the kernel of the action is $\cap_{x \in X} G_x$, that is, the set consisting of all the elements of $G$ which act trivially on $X$.

An action is faithful if its kernel is trivial $= \{e_G\}$.

**Example.** $G = S_4$ and $X = \text{the set of all partitions of } \{1,2,3,4\} \text{ into two parts of size } 2 = \{12|34, 13|24, 14|23\}$. Then $G$ acts on $X$ in the way you would expect. What is the kernel? Investigate this action.

**Exercise.** Show that if $G$ is any group acting on a set $X$, with $x \in X$ and $g \in G$, then $G_g(x) = gG_xg^{-1} = \{ghg^{-1} : h \in G_x\}$.

### 4.3 Conjugation

**Lemma 4.9.** Let $G$ be a group, $X = G$ with $g(h)$ defined as $ghg^{-1}$ for $g \in G, h \in X$. Then $ghg^{-1}$ is an element of $G$ (which is $X$). This is an action, called the conjugation action.

**Proof.** $(g_1g_2)h(g_1g_2)^{-1} = g_1(g_2 h g_2^{-1}) g_1^{-1} = g_1(g_2(h))$ so it satisfies the first axiom.

#### 4.3.1 Conjugacy classes and centralisers

The orbit of $x \in G$ is called its conjugacy class, written $ccl_G(x) = \{gxg^{-1} | g \in G\}$. Note that $e$ is the only element in its conjugacy class, and others may be the only element in their classes as well. If two elements are in the same orbit, they are said to be conjugate.

The stabiliser of $x$ in $G$ is called its centraliser, written $C_G(x) = \{g \in G \mid gx = xg\}$. This is the set of all elements which commute with $x$.

**Remark 4.10.** If $G$ is a finite group and $x \in G$, then $|G : C_G(x)| = |ccl_G(x)|$ by the orbit-stabiliser theorem.

#### 4.3.2 The centre of a group

The kernel of the conjugation action is called the centre of $G$, written $Z(G) = \cap_{x \in G} C_G(x)$. This is the intersection of all the centralisers, so consists of the elements which commute with everything, i.e. $\{g \in G : gx = xg \forall x \in G\}$.

**Exercise.** Find the conjugacy classes of elements of $D_8$ and find its centre. Ditto for $Q_8$. (The centre of $D_8$ is $\{a^2\}$.)

**Remark 4.11.**

(i) $o(gxg^{-1}) = o(x)$, i.e. conjugate elements have the same order.

(ii) $z \in Z(G)$ if $ccl_G(z) = \{z\}$, i.e. $z$ is the only element in its conjugacy class, i.e. $G = C_G(z)$.

**Proposition 4.12.** If $G$ is a finite group of order $p^n$ for some prime $p$, then the centre of $G$ is not just the identity but contains at least one other element.

**Proof.** Exercise – consider the conjugation action of $G$ and its orbit sizes, which must divide the order of $G$.

#### 4.3.3 Cycle type and conjugacy classes in $S_n$

Consider the permutation $\pi \in S_n$, written in disjoint-cycle notation (including fixed points or 1-cycles). The cycle type of $\pi$ is $(n_1, n_2, \ldots, n_k)$ where $n_1 \geq n_2 \geq \cdots \geq n_k \geq 1$, and the cycles in $\pi$ have length $n_i$. It follows that $n = n_1 + \cdots + n_k$.

**Examples.** $(1 \ 2 \ 3)(4 \ 5)(6) \in S_6$ has cycle type $(3,2,1)$.

$e \in S_6$ has cycle type $(1,1,1,1,1,1) = (1^6)$.

**Theorem 4.13.** The permutations $\pi, \sigma \in S_n$ are conjugate in $S_n$ iff they have the same cycle type.
Proof. If \( \sigma = (a_{11}a_{12} \ldots a_{1n_1})(a_{21}a_{22} \ldots a_{2n_2}) \ldots (a_{k1}a_{k2} \ldots a_{kn_k}) \) and \( \tau \in S_n \) (any element), then \((4.14) \tau \sigma \tau^{-1} = \left( \tau(a_{11}) \tau(a_{12}) \ldots \tau(a_{1n_1}) \right) \left( \tau(a_{21}) \tau(a_{22}) \ldots \tau(a_{2n_2}) \right) \ldots \left( \tau(a_{kn_k}) \right) \). Now note that \( \tau \sigma \tau^{-1}(\tau(a_{11})) = \tau \sigma(a_{11}) = \tau(a_{12}) \), and the same pattern works for all \( a_{ij} \). Therefore if elements are conjugate in \( S_n \) then they have the same cycle type.

Conversely, if \( \sigma \) is as before and \( \pi \) has the same cycle type, say \( \pi = (b_{11}b_{12} \ldots b_{1n_1})(b_{21}b_{22} \ldots b_{2n_2}) \ldots (b_{kn_k}) \), then let \( \tau \) be a permutation with \( \tau : a_{ij} \mapsto b_{ij} \). Then by \( 4.14, \pi = \tau \sigma \tau^{-1} \), so \( \pi, \sigma \) are conjugate.

Examples. \((1 \; k) = (2 \; k)(1 \; 2)(2 \; k)^{-1} \) and \((k \; l) = (1 \; l)(1 \; k)(1 \; l)^{-1} \), see 1.19. In fact, all transpositions are conjugate as they have the same cycle type.

Example. The following is a table of the conjugacy classes in \( S_4 \):

<table>
<thead>
<tr>
<th>example member</th>
<th>cycle type</th>
<th>size</th>
<th>sign</th>
<th>centraliser</th>
</tr>
</thead>
<tbody>
<tr>
<td>((1 ; 2)(3 ; 4))</td>
<td>(1,1,1,1)</td>
<td>6</td>
<td>+</td>
<td>( S_4 )</td>
</tr>
<tr>
<td>((1 ; 2 ; 3)(4))</td>
<td>(3,1)</td>
<td>8</td>
<td>+</td>
<td>self-centralising</td>
</tr>
<tr>
<td>((1 ; 2)(3 ; 4))</td>
<td>(2,2)</td>
<td>3</td>
<td>+</td>
<td>( D_8 )</td>
</tr>
<tr>
<td>((1 ; 2 ; 3 ; 4))</td>
<td>(4)</td>
<td>6</td>
<td>-</td>
<td>( \langle 1 ; 2 ; 3 ; 4 \rangle )</td>
</tr>
</tbody>
</table>

Corollary 4.15. The number of conjugacy classes in \( S_n \) is \( p(n) \), the number of partitions of \( n \) into \( n = n_1 + \cdots + n_k \) for some \( k \) with \( n_1 \geq \cdots \geq n_k \geq 1 \).

4.3.4 Conjugacy classes in \( A_n \)

The following is a table of the conjugacy classes in \( A_4 \):

<table>
<thead>
<tr>
<th>example member</th>
<th>cycle type</th>
<th>size</th>
<th>centraliser</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e )</td>
<td>(1^4)</td>
<td>1</td>
<td>( A_4 )</td>
</tr>
<tr>
<td>((1 ; 2)(3 ; 4))</td>
<td>(2^2)</td>
<td>3</td>
<td>( \langle (1 ; 2)(3 ; 4), (1 ; 3)(2 ; 4) \rangle )</td>
</tr>
<tr>
<td>((1 ; 2 ; 3)(4))</td>
<td>(3,1)</td>
<td>4</td>
<td>( \langle (1 ; 2 ; 3) \rangle )</td>
</tr>
<tr>
<td>((1 ; 3 ; 2)(4))</td>
<td>(3,1)</td>
<td>4</td>
<td>( \langle (1 ; 2 ; 3) \rangle )</td>
</tr>
</tbody>
</table>

For general \( n \), let \( x \in A_n \). Then \( \text{ccl}_{A_n}(x) \subseteq \text{ccl}_{S_n}(x) \), because \( A_n \subseteq S_n \). Now \( C_{S_n}(x) \) either consists of even permutations, in which case \( C_{A_n}(x) = C_{S_n}(x) \), or it contains an odd permutation (i.e. precisely half of its elements are odd), in which case \( C_{A_n}(x) \) has index 2 in \( C_{S_n}(x) \).

Now \( |\text{ccl}_{S_n}(x)| = |S_n : C_{S_n}(x)| \) so \( |\text{ccl}_{A_n}(x)| = |A_n : C_{A_n}(x)| \). Thus

**Theorem 4.16.** Let \( x \in A_n \). Then \( \text{ccl}_{A_n}(x) \subseteq \text{ccl}_{S_n}(x) \), with equality iff the centraliser in \( S_n \) of \( x \) is half odd. Moreover, if inequality holds, then \( \text{ccl}_{S_n}(x) \) splits into two classes in \( A_n \), both of the same size, \( |\text{ccl}_{S_n}(x)|/2 \).

4.4 Cayley’s theorem

Let \( G \) be a group and let \( X = G \). Then (Lemma 4.3) the map \( \rho : G \times X \to X \) taking \( (g, x) \mapsto gx \) is an action (because it satisfies all the axioms). This is called the left regular action of \( G \).

Hence we get a homomorphism from \( G \to \text{Sym}(X) \). If, for some \( x \in X, g_1(x) = g_2(x) \), then \( g_1x = g_2x \) so \( g_1 = g_2 \). So the homomorphism is injective.

**Theorem 4.4 (Cayley’s theorem).** Any group \( G \) is isomorphic to a subgroup of \( \text{Sym}(X) \) for some non-empty set \( X \). E.g. we can take \( X \) to be the set of elements of \( G \).

**Proof.** We have just found an injective homomorphism \( \varphi : G \to \text{Sym}(G) \) arising from the left regular action. Hence \( G \) is isomorphic to the subgroup \( \varphi(G) = \{ \varphi(g) : g \in G \} \subseteq \text{Sym}(G) \); we write this as \( G \leq \text{Sym}(G) \).
Remark. Let \( \varphi : G \to H \) be a homomorphism. Then \( \varphi(G) = \{ \varphi(g) : g \in G \} \leq H \). This is because \( e_H = \varphi(e_G) \in \varphi(G) \), and if \( a, b \in \varphi(G) \), say \( a = \varphi(g) \) and \( b = \varphi(h) \) with \( g, h \in G \), then we have \( a^{-1}b = (\varphi(g))^{-1}\varphi(h) = \varphi(g^{-1}h) \in \varphi(G) \).

4.5 Cauchy’s theorem

**Theorem 4.17 (Cauchy’s theorem).** If \( G \) is a finite group of order divisible by the prime \( p \), then \( G \) contains an element of order \( p \).

**Proof.** Let \( X = \{ (x_1, x_2, ..., x_p) \mid x_1 x_2 ... x_p = e \} \) with all \( x_i \in G \). Let the cyclic group \( C_p = \langle g \mid g^p = e \rangle \) act on the set \( X \) such that \( g : (x_1, x_2, ..., x_p) \mapsto (x_{1+j}, x_{2+j}, ..., x_j) \), so that \( g \) cycles the coordinates. (Note; we are working \( \mod{p} \).)

This is an action of \( C_p \); the image is in \( X \) since if \( x_1 x_2 ... x_p = e \), then \( x_{1+j} x_{2+j} ... x_j = e \).

Now look at the orbits of this action; they have sizes \( 1 \) or \( p \). Any orbit of size \( 1 \) is \( \{ (x, x, ..., x) \mid x^p = e \} \), i.e. \( (e, e, ..., e) \). Finally, \( |X| = |G|^{p-1} \), since the first \( p - 1 \) coordinates can be chosen independently, but the last is forced; \( x_p = (x_1 x_2 ... x_{p-1})^{-1} \). Therefore \( p | |X| \). Since \( X \) is partitioned into \( C_p \)-orbits of sizes \( 1 \) or \( p \), there are at least \( p \) of them of size \( 1 \).

4.6 Groups of symmetries of regular solids

4.6.1 Tetrahedron

Let \( G \) be the group of all symmetries of a tetrahedron and let \( G^+ \) be the group of rotations (rigid motions), which is a subgroup of \( G \). Let \( X = \{ 1, 2, 3, 4 \} \) be the set of vertices. Then \( G \) acts on \( X \) transitively.

Now \( G \leq Sym(X) = S_4 \). If all vertices are fixed by a symmetry, it is \( i \). By the orbit-stabiliser theorem, \( |G| = |G(1)| |G_1| \), but \( |G_1| = 6 \) because \( G_1 = S_3 \), and \( |G(1)| = 4 \) obviously. Thus \( G = S_4 \).

The same argument works for \( G^+ \), because \( |G^+| = 4 \times |G^+_f| = 4 \times 2 \times 2 = 12 \) and in fact \( G^+ \cong A_4 \). In fact, \( G^+ \) contains all the \( 3 \)-cycles, and thus also \( (1 2 3)(1 2 4) = (1 3)(2 4) \) and the other elements of cycle type \( 2^2 \).

4.6.2 Cube (or octahedron)

Put vertices in the middle of a cube’s faces and you get an octahedron, so the two are dual.

Let \( G \) and \( G^+ \) be as before, and let \( G^+ \) act on the set \( F \) of all faces, where \( |F| = 6 \). Then \( G^+ \) is transitive on \( F \) by rotations. So \( |G^+_f| = 6 \times |G_f^+| = 6 \times 4 = 24 \) for any face \( f \) (stabilise one face and you also stabilise the opposite face; the others can rotate into four positions).

Now let \( D \) be the set of all diagonals of the cube;
\[
D = \{ d_1, d_2, d_3, d_4 \} = \text{diagonals on } i \text{ and } i'.
\]

\( G^+ \) and \( G \) act on \( D \), so we have a homomorphism of \( G \) and \( G^+ \) to the symmetric group \( S_4 \) on \( D \).

The kernel of the action of \( G \) on the set \( D \) is the set of permutations which act trivially on all the diagonals.

- Claim; it is the group \( \langle g \rangle \) of order \( 2 \) where \( g = (1 1')(2 2')(3 3')(4 4') \). In fact, placing the vertices in \( \mathbb{R}^3 \) at points \((\pm 1, \pm 1, \pm 1)\), \( g = -I \) because it sends \(+\) to \(-\) and vice versa.
- **Proof:** if an element of the kernel on the set \( D \) moves \( 1 \) to \( 1' \), then \( 1' \to 1, 2 \to 2' \) (as \( 2' \) is next to \( 1' \) and \( 2 \) isn’t), and so on, by cube-waving. So \( g \) is in the kernel. Also, if an element of the kernel of \( G \) on \( D \) stabilises \( 1 \), then it stabilises \( 1' \) and hence (by cube-waving) it stabilises everything else. Hence the kernel is \( \langle g \rangle \).
What about $G^+$? Since $G^+ \leq G$, and $g \not\in G^+$ (it’s not a rotation), the kernel of the action of $G^+$ on $D$ is trivial, so we have $G^+ \simeq S_4$.

Finally, $G \simeq G^+ \times \langle g \rangle \simeq S_4 \times C_2$. This is true because the three conditions in 3.5 are satisfied;

(i) $G^+$ has index 2 in $G$, so $G = G^+ \cup gG^+$, so anything in $G$ is a rotation or $g$ times a rotation.

(ii) $G^+ \cap \langle g \rangle = \{e\}$ is true.

(iii) every element of $G^+$ commutes with $g$; this is true by cube-waving.

Exercise. Find all elements of $S_4$ explicitly as rotations of a cube.

4.6.3 Dodecahedron / icosahedron (non-examinable)

A dodecahedron is made of pentagons and has 12 faces, 30 edges and 20 vertices. Its dual is the icosahedron.

Let $G$ and $G^+$ be as before; they act transitively on the set $F$ of faces, where $|F| = 12$. By the orbit-stabiliser theorem, $|G^+| = |F| \times |G^+_f| = 12 \times 5 = 60$ where $f$ is any face.

There are five cubes embedded into our dodecahedron; each edge of a cube appears as a diagonal of a face of our dodecahedron. Let $C$ be the set of these 5 embedded cubes. Then $G^+$ acts on $C$ faithfully (the kernel is trivial). So we obtain an injective homomorphism into $Sym(C) = S_5$.

Hence $\varphi(G^+) = A_5$, and $G^+ \simeq A_5$. Now $G$ also contains $-I$ (and note that $\pm I$ is the kernel of the action of $G$ on $C$). So $G \simeq G^+ \times \{\pm I\} \simeq A_5 \times C_2$.

5. Homomorphisms, normal subgroups and quotient groups

5.1 Homomorphisms

Recall. Let $(G,*_G)$ and $(H,*_H)$ be groups. The mapping $\varphi : G \rightarrow H$ is a homomorphism if it satisfies $\varphi(g_1 *_G g_2) = \varphi(g_1) *_H \varphi(g_2)$ for all $g_1,g_2 \in G$. If it is also bijective, then it is an isomorphism.

5.1.1 Image

The image of a homomorphism is $\varphi(G) = \{\varphi(g) : g \in G\}$. $\varphi(G) \leq H$ because; $e_H = \varphi(e_G) \in \varphi(G)$; and also, if $h_1,h_2 \in \varphi(G)$, choose $g_1,g_2 \in G$ with $\varphi(g_1^{-1}) = h_1$ and $\varphi(g_2) = h_2$. Then $h_1 h_2 = \varphi(g_1^{-1}) \varphi(g_2) = \varphi(g_1^{-1} g_2) \in \varphi(G)$.

In fact, for any $L \leq G$ it follows that $\varphi(L) \leq H$.

5.1.2 Kernel

The kernel of a homomorphism $\varphi$ is $\ker \varphi = \{g \in G : \varphi(g) = e_H\}$.

Examples. (i) $G = S_n$, $H = \{\pm 1,\infty\}$. Then $\text{sgn} : G \rightarrow H$ is a surjective homomorphism (see 1.13), with kernel the alternating group $A_n$ (since all even elements are mapped to the identity).

(ii) Let $C^* = (\mathbb{C} \setminus \{0\}, \times)$ and $H = \{z \in C^* : |z| = 1\}$. Then $\varphi : z \mapsto z/|z|$ is a homomorphism in which case the image is obviously the unit circle $H$ – in fact, $\varphi$ does nothing to the unit circle. The kernel of $\varphi$ is those elements which map to the identity 1, i.e. the positive half of the real line, i.e. $(\mathbb{R} > 0, \times)$.

Lemma 5.2. A homomorphism $\varphi : G \rightarrow H$ is injective iff $\ker \varphi = \{e_G\}$. (ker $\varphi$ is the ‘obstruction to injectivity’ of $\varphi$.)

Proof ($\Rightarrow$). If $\varphi$ is injective, assume $g \in \ker \varphi$. Then $\varphi(g) = e_H = \varphi(e_G)$, so $g = e_G$ due to $\varphi$ being injective. So the kernel is just the identity.

Proof ($\Leftarrow$). If $\ker \varphi = \{e_G\}$, let $g_1, g_2 \in G$ with $\varphi(g_1) = \varphi(g_2)$. Then $\varphi(g_1^{-1} g_2) = e_H$, so $g_1^{-1} g_2 \in \ker \varphi$. But $\ker \varphi = \{e_G\}$, so $g_1 = g_2$. Hence $\varphi$ is injective.
5.2 Normal subgroups

**Definition.** A subgroup $H \leq G$ is normal (written $H \triangleleft G$) if each left coset is equal to the corresponding right coset, i.e. $gH = Hg \forall g \in G$. Therefore for any $g \in G, k \in K \exists k' \in K$ with $gk = k'g$, i.e. $gkg^{-1} \in K \forall g \in G, k \in K$.

5.2.1 Notes

**Note.** If $G$ is abelian, then any subgroup is normal, because $gkg^{-1} = k \in K$.

**Lemma 5.4.** If $K \leq G$ of index 2, then $K \triangleleft G$ (often used in chapter 3).

**Proof.** If $g \in G$, then either $g \in K$, in which case $gK = K = Kg$, or $g \in G \setminus K$, in which case $gK = G \setminus K = Kg$. Either way, $K \triangleleft G$.

It follows from the definition that if any one member of a conjugacy class is an element of a normal subgroup, then the entire conjugacy class must be within that subgroup. Therefore any normal subgroup is a (disjoint) union of conjugacy classes.

5.2.2 Normality of the kernel

**Lemma 5.1.** If $\varphi : G \rightarrow H$ is a homomorphism, then $\ker \varphi \triangleleft G$.

**Proof.** First note $e_G \in \ker \varphi$. Next, if $g_1, g_2 \in \ker \varphi$ then so is $g_1^{-1}g_2$, since $\varphi(g_1^{-1}g_2) = \varphi(g_1^{-1})\varphi(g_2) = e_H$. Therefore $\ker \varphi \leq G$.

Now if $g \in G, k \in \ker \varphi$, then $gkg^{-1} \in \ker \varphi$, since $\varphi(gkg^{-1}) = \varphi(g)\varphi(k)\varphi(g^{-1}) = e_H$ using the fact that $\varphi(g^{-1}) = (\varphi(g))^{-1}$. Therefore the subgroup is normal. QED.

**Remark 5.3.** If $G$ acts on a set $X$, we have the homomorphism $\varphi : G \rightarrow \text{Sym}(X)$ taking $g \mapsto \varphi_g$ where $\varphi_g : x \mapsto g(x)$. The kernel of the action on $X$ is precisely $\ker \varphi$, hence it is a normal subgroup of $G$.

5.2.3 Normal factors

**Lemma 5.5 (clarification of 3.5).** If $G, G_2 \triangleleft G$ such that $G = \{G_1, G_2\}$ and $G_1 \cap G_2 = \{e\}$, then $G \cong G_1 \times G_2$.

**Note.** This was used when $G$ was the group of all symmetries of a cube, $G_1$ was $G^\perp \triangleleft G$, and $G_2$ was the kernel of the action of $G$ on $D$, thus $G_2 \triangleleft G$.

**Proof.** It is enough to show, by 3.5, that $g_1g_2 = g_2g_1$ for any $g_1 \in G_1, g_2 \in G_2$. Now $g_1g_2g_1^{-1}g_2^{-1} \in G_1 \cap G_2$, because $g_1g_2g_1^{-1} \in G_1$ and $g_2g_1^{-1}g_2^{-1} \in G_2$. Also $G_1 \triangleleft G$, so $g_1g_2 = g_2g_1$.

5.3 Quotient group

**Theorem 5.6.** If $K \triangleleft G$, the set $(G : K)$ of cosets of $K$ in $G$ is a group under the operation of coset multiplication, defined by $g_1K \times g_2K = (g_1g_2)K$. This is called the quotient group $G/K$.

**Proof.** First we must check this operation is well-defined on $(G : K)$; that is, $(g_1g_2)K$ is a valid coset of $K$ in $G$ (obvious), and if $g_1K = g_1'K$ and $g_2K = g_2'K$ then $(g_1g_2)K = (g_1'g_2')K$ (less obvious).

Directly from the above expressions $g^{-1}_ig_i' \in K$ (for $i = 1,2$) so call it $k_i$. Now

$$(g_1g_2)K = (g_1'g_2')K \iff g_2^{-1}g_1^{-1}g_1'g_2' \in K$$

but

$$g_2^{-1}g_1^{-1}g_1'g_2' = g_2^{-1}g_1^{-1}g_2^{-1}g_2g_2^{-1}k_1g_2' = k_2g_2^{-1}k_1g_2' \in k_2K$$

due to the fact that $(g_2^{-1})^{-1}k_1g_2' \in K$. Finally, $k_2K = K$ so we are done.

Next we check the remaining group axioms. Associativity; $(g_1Kg_2K)g_3K = (g_1g_2)Kg_3K = (g_1g_2g_3)K = g_1K(g_2g_3)K = g_1Kg_2Kg_3K$ as required.

Identity element; $gKeK = gK$ and $eKgK = gK$ so the identity $eK = K$.

Inverses; $(gK)^{-1} = g^{-1}K$, because $gKg^{-1}K = eK = K = g^{-1}KgK$. QED.

**Remark.** Normality of $K$ is necessary.
Example. $G = D_8 = \langle a, b : a^4 = e = b^2, bab^{-1} = a^{-1} \rangle$ and $K = \langle a^2 \rangle$. Then

$$G/K = \{ K, aK, bK, abK \} \cong C_2 \times C_2 = \langle aK, bK \rangle$$

Note that $\langle b \rangle$ is not normal, but $\langle a^2 \rangle$ and all subgroups of order 4 are normal.

5.4 Isomorphism theorem

5.4.1 Statement and proof

**Theorem 5.7 (Isomorphism Theorem).** Let $\varphi : G \to H$ be a homomorphism. Then $\varphi(G) \leq H$, $\ker \varphi \trianglelefteq G$, and $G/\ker \varphi \cong \varphi(G)$.

**Note.** Any homomorphic image of $G$ is isomorphic to a quotient.

**Proof.** For the first two assertions see section 5.1.1 and Lemma 5.1.

In general, the way to show that two groups are isomorphic is to construct an isomorphism between them. Therefore we define the map $\bar{\varphi} : G/K \to \varphi(G)$ taking $gK \mapsto \varphi(g)$ (where $K = \ker \varphi$) and will show that it is an isomorphism.

- **Well-defined:** if $g_1K = g_2K$, then $g_1^{-1}g_2 \in K$, so $\varphi(g_1^{-1}g_2) = e_H = (\varphi(g_1))^{-1}\varphi(g_2)$. Therefore $\varphi(g_1) = \varphi(g_2)$, so $\bar{\varphi}(g_1K) = \bar{\varphi}(g_2K)$.
- **Injectivity:** reverse the steps above (carefully).
- **Surjectivity:** any element in $\varphi(G)$ is $\varphi(g)$ for some $g \in G$, so $\bar{\varphi}(gK) = \varphi(g)$.
- **Homomorphism:** $\bar{\varphi}(g_1Kg_2K) = \bar{\varphi}(g_1g_2K) = \varphi(g_1g_2)$ by definition of $\bar{\varphi}$. Then $\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2)$ since $\varphi$ is a homomorphism, and this $= \bar{\varphi}(g_1K)\bar{\varphi}(g_2K)$. So we are done.

5.4.2 Examples

**Example.** Let $G = (\mathbb{R}, +)$ and $H = \mathbb{C}^* = (\mathbb{C} \setminus \{0\}, \times)$, with $\varphi : G \to H$ taking $t \mapsto e^{it}$. First we check $\varphi$ is a homomorphism; $\varphi(t_1 + t_2) = e^{it_1 + it_2} = e^{it_1}e^{it_2} = \varphi(t_1)\varphi(t_2)$. Now $\varphi(G) = \mathbb{R}' = \{ z \in \mathbb{C} : |z| = 1 \}$. Then $\varphi = (2\pi) \leq \mathbb{R}$, which is normal in $(\mathbb{R}, +)$. Therefore $\mathbb{R}/(2\pi) = S'$ (the unit circle).

**Example.** Let $G = (\mathbb{Z}, +)$. Then $\langle n \rangle = \{ nk : k \in \mathbb{Z} \} \triangleleft G$. Let $H = C_n = \langle x : x^n = 1 \rangle$. Define $\varphi : G \to H$ taking $m \mapsto x^m$. This is a surjective homomorphism with kernel $\langle n \rangle$. Therefore $\mathbb{Z}/\langle n \rangle = C_n$ and $\mathbb{Z}/\langle n \rangle$ is the group of integers mod $n$, sometimes written $\mathbb{Z}/n$.

5.4.3 Applications, extensions and corollaries

**Theorem 5.8.** Let $K \trianglelefteq G$. Then $K$ is the kernel of the natural surjective homomorphism $\theta : G \to G/K$, which takes $g \mapsto gK$.

**Proof.** Homomorphism; $\theta(g_1g_2) = g_1g_2K = g_1Kg_2K = \theta(g_1)\theta(g_2)$. Surjective; just note that $gK = \theta(g)$. Now ker $\theta = \{ g \in G : \theta(g) = e_K \} = \{ g \in G : gK = K \} = K$ as required.

**Remarks.** (i) Homomorphic images of $G$ are equivalent to quotients $G/K$ with $K \trianglelefteq G$.

(ii) There are more complicated isomorphism theorems to come in IB Groups, Rings and Modules. E.g. subgroups of $G$ containing $K \leftrightarrow$ subgroups of $G/K$.

**Corollary 5.9.** Let $G$ act on a set $X$ by $g : x \mapsto g(x)$. Let $\varphi : G \to \text{Sym}(X)$ be a homomorphism, with $\varphi_g \in \text{Sym}(X)$, taking $g \mapsto \varphi(g)$. The kernel of the action is $\bigcap_{x \in X} G_x = \ker \varphi = K$. By the isomorphism theorem, $G/K \cong \varphi(G) \leq \text{Sym}(X)$. So it follows that $G$ has a normal subgroup of index dividing $n!$, where $|X| = n$ so that $|\text{Sym}(X)| = n!$. 
Example. Let $H \leq G$ of index $n$. Then $G$ has a normal subgroup $K$ of index dividing $n!$, because $G$ acts on the set $X$ of all left cosets of $H$ in $G$, where $|X| = n$.

### 5.5 Simple groups

A group $G$ is simple if \{e\} and $G$ are the only normal subgroups. E.g. $C_p$ (for $p$ prime) and $A_5$.

#### 5.5.1 Simplicity of $A_5$

Conjugacy classes in $S_5$;

<table>
<thead>
<tr>
<th>cycle type</th>
<th>example member</th>
<th>size</th>
<th>centraliser</th>
<th>sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1^5$</td>
<td>$\iota$</td>
<td>1</td>
<td>$S_5$</td>
<td>+</td>
</tr>
<tr>
<td>$2,1^3$</td>
<td>$(1\ 2)$</td>
<td>10</td>
<td>$(1\ 2) \times S_3$</td>
<td>$-$</td>
</tr>
<tr>
<td>$2^2, 1$</td>
<td>$(1\ 2)(3\ 4)$</td>
<td>15</td>
<td>$D_8$</td>
<td>$+$</td>
</tr>
<tr>
<td>$3,1^2$</td>
<td>$(1\ 2\ 3)$</td>
<td>10</td>
<td>$(1\ 2\ 3) \times S_2$</td>
<td>$+$</td>
</tr>
<tr>
<td>$3,2$</td>
<td>$(1\ 2\ 3)(4\ 5)$</td>
<td>20</td>
<td>$(1\ 2\ 3) \times (1\ 2\ 3)$</td>
<td>$-$</td>
</tr>
<tr>
<td>$4,1$</td>
<td>$(1\ 2\ 3\ 4)$</td>
<td>15</td>
<td>$(1\ 2\ 3\ 4)$</td>
<td>$-$</td>
</tr>
<tr>
<td>$5$</td>
<td>$(1\ 2\ 3\ 4\ 5)$</td>
<td>20</td>
<td>$\langle x \rangle$</td>
<td>$+$</td>
</tr>
</tbody>
</table>

The normal subgroups of $S_5$ are \{e\}, $A_5$, $S_5$ and no others, since any normal subgroup must be a union of conjugacy classes. The only other possibility would be, in $A_5$, to replace the conjugacy class $3,1^2$ by the conjugacy class $3,2$; but if $x$ is of cycle type $3,2$, then $x^3$ is of type $2,1^3$ which would not be present.

Conjugacy classes in $A_5$;

<table>
<thead>
<tr>
<th>cycle type</th>
<th>size</th>
<th>centraliser</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1^5$</td>
<td>1</td>
<td>$A_5$</td>
</tr>
<tr>
<td>$2^2, 1$</td>
<td>15</td>
<td>$V_4$ (Klein 4-group)</td>
</tr>
<tr>
<td>$3,1^2$</td>
<td>20</td>
<td>$\langle x \rangle$</td>
</tr>
<tr>
<td>$5$</td>
<td>12</td>
<td>$\langle x \rangle$</td>
</tr>
<tr>
<td>$5$</td>
<td>12</td>
<td>$\langle x' \rangle$</td>
</tr>
</tbody>
</table>

Note; $(1\ 2)(1\ 2\ 3\ 4\ 5)(1\ 2) = (2\ 1\ 3\ 4\ 5)$, so the two are conjugate in $S_5$ but not in $A_5$; all conjugating elements in $S_5$ are odd. Any element in $S_5$ conjugating $(1\ 2\ 3\ 4\ 5)$ to $(2\ 1\ 3\ 4\ 5)$ is in the coset $(4\ 5)C$ with $C = \langle (1\ 2\ 3\ 4\ 5) \rangle = C_n(x)$ - so in $S_5 \setminus A_5$.

Therefore $A_5$ is simple, because its only normal subgroups are \{e\} and $A_5$.

#### 5.5.2 Simple building blocks (non-examinable)

Any finite group can be broken up into simple quotients; if $H_1 \triangleleft H$ is maximal and normal, then $H/H_1$ is simple. Then take $H_2 \triangleleft H_1$ maximal and normal, and so on.

Simple groups are the ‘building blocks’ of finite groups. However, putting together simple groups can be complicated! E.g. if $|G| = 2^{10}$ then all simple factors must be $C_2$ - but there are many examples.

#### 5.5.3 Classification of finite simple groups (non-examinable)

- $C_p$, for $p$ prime
- $A_n$ for $n \geq 5$ (see IB for the proof where $n > 5$)
- $PSL_n(q)$ where $q$ is a prime power and $n \geq 2$; also $PSO_n(q)$, $PSU_n(q)$ and $PSp_n(q)$. These four are known as the ‘classical groups’

$$GL_2(5) \triangleleft \frac{SL_2(5)}{Z}$$
where \( \text{GL}_2(5) \) is integer arithmetic mod 5 and \( Z = \text{centre} = \{ \text{scalars in } \text{SL}_2(5) \} = \{ \pm I \} \)

\[ \text{PSL}_2(5) \cong A_5 \]  (see example sheet 4, question 13)

Then there are the exceptional groups; \( G_2 \) and 10 more families corresponding to Lie algebras, and the 26 sporadic groups, starting with \( M_{11} \) (which was discovered in the 1860s and has 7920 elements) and working up to \( M \), the ‘Monster’ group.

### 6. Groups of matrices I

#### 6.1 General linear group

We write \( M_n(\mathbb{R}) \) = the set of all \( n \times n \) matrices over \( \mathbb{R} \). Then we define

\[ \text{GL}_n(\mathbb{R}) = \{ A \in M_n(\mathbb{R}) : \det A \neq 0 \} \]

to be the general linear group. We will mostly deal with \( n = 2 \) or 3.

**Proposition 6.1.** \( \text{GL}_n(\mathbb{R}) \) is a group under multiplication.

**Proof.** Closure: \( \det AB = \det A \det B \), so if \( \det A \neq 0 \) and \( \det B \neq 0 \) then \( \det AB \neq 0 \). For associativity, see Vectors and Matrices. The identity is \( I_n \), and as for inverses, if \( A \in \text{GL}_n(\mathbb{R}) \), then \( \det A \neq 0 \) so \( A \) has an inverse \( A^{-1} \) with non-zero determinant.

**Example.** For \( n = 2 \),

\[ \text{GL}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : ad - bc \neq 0 \right\}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \]

**Proposition 6.2.** \( \det : \text{GL}_n(\mathbb{R}) \to \mathbb{R}^* = \{ \mathbb{R}\setminus\{0\}, \times \} \) taking \( A \mapsto \det A \) is a surjective homomorphism.

**Proof.** \( \det AB = \det A \det B \). If \( r \in \mathbb{R}^* \), then

\[ \det \begin{pmatrix} r & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = r, \]

so it’s surjective.

#### 6.2 Special linear group

**Definition.** Let the special linear group \( \text{SL}_n(\mathbb{R}) = \ker \det \). Thus \( \text{SL}_n(\mathbb{R}) = \{ A \in M_n(\mathbb{R}) : \det A = 1 \} \lhd \text{GL}_n(\mathbb{R}) \). Also, by the isomorphism theorem,

\[ \frac{\text{GL}_n(\mathbb{R})}{\text{SL}_n(\mathbb{R})} \cong \mathbb{R}^*. \]

**Remark.** \( \mathbb{F} \) is a field if \( \mathbb{F} \) is an abelian group under addition, \( \mathbb{F}\setminus\{0\} \) is an abelian group under multiplication, and the distributive law holds. E.g. \( \mathbb{R}, \mathbb{C}, \mathbb{Q}, \mathbb{F}_p \) (the Galois field of size \( p \) = the integers mod \( p \), with \( |\mathbb{F}_p| = p \) for \( p \) prime).

We can define \( \text{GL}_n(\mathbb{F}) \) and \( \text{SL}_n(\mathbb{F}) \) for each field.

#### 6.3 Actions of these groups

An important action of \( \text{GL}_n(\mathbb{C}) \) on \( \mathbb{C}^n \), the (vector) space of columns of length \( n \) with complex entries, is \( A : v \mapsto Av \) (also written \( (A, v) \mapsto Av \)). This action is faithful, i.e. \( Av = v \forall v \in \mathbb{C}^n \Rightarrow A = I \), and has an orbit \( \mathbb{C}^n\setminus\{0\} \), as you can get from any vector to any other vector.

Another important action of \( \text{GL}_n(\mathbb{C}) \) on \( M_n(\mathbb{C}) \) is the following. Take \( P \in \text{GL}_n(\mathbb{C}) \) and \( A \in M_n(\mathbb{C}) \); then \( P : A \mapsto PAP^{-1} \). This is called the conjugation action on \( M_n(\mathbb{C}) \).
Remark. Two matrices in $M_n(\mathbb{C})$ are conjugate iff they represent the same linear transformation on the vector space $\mathbb{C}^n$ (this is proved in Linear Algebra).

A goal of Linear Algebra is to find a ‘good’ representative for each conjugacy orbit. E.g. for $n = 2$;

**Theorem 6.3.** Given $A \in M_2(\mathbb{C})$, precisely one of the following occurs;

(i) the orbit of $A$ contains the diagonal matrix \( \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \), $\lambda \neq \mu$, with eigenvalues $\lambda, \mu$

(ii) the orbit of $A$ is $\{\lambda I_2\}$ with eigenvalues $\lambda, \lambda$ (still diagonal)

(iii) the orbit of $A$ contains the matrix \( \begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix} \) with eigenvalues $\lambda, \lambda$.

**Proof.** See Vectors and Matrices.

In case (i), there is one orbit for each pair $\lambda, \mu$. In cases (ii) and (iii), there is one orbit for each $\lambda$.

7. Möbius transformations and Möbius groups

7.1 Möbius transformations

A Möbius transformation on $\mathbb{C}$ is a transformation of the form

\[ f(z) = \frac{az + b}{cz + d}, \quad a, b, c, d \in \mathbb{C}, \quad ad - bc \neq 0. \]

The restriction $ad - bc \neq 0$ is to make it injective;

\[ f(z) - f(w) = \frac{(ad - bc)(z - w)}{(cz + d)(cw + d)}. \]

But $f$ is not defined as yet at $z = -d/c$, if $c \neq 0$. So we add a new symbol $\infty$ to $\mathbb{C}$ and work over $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$, the extended complex plane.

**Definition.** $f(z)$ is defined as above, with the following extensions; if $c \neq 0$, define $f(-d/c) = \infty$ and $f(\infty) = a/c$; and if $c = 0$, define $f(\infty) = \infty$.

**Remark.** $\infty$ is just a new point in the plane, and it fits our intuition.

7.2 The Möbius group

7.2.1 Proof that it is a group

**Theorem 7.1.** The set $\mathcal{M}$ of all Möbius transformations on $\mathbb{C}_\infty$ is a group under composition. It is a subgroup of $\text{Sym}(\mathbb{C}_\infty)$.

**Proof.** We prove each group axiom in turn.

- **Closure:** if

\[ f(z) = \frac{az + b}{cz + d}, \quad g(z) = \frac{\alpha z + \beta}{\gamma z + \delta}, \quad a, b, c, d \in \mathbb{C}, \quad \alpha, \beta, \gamma, \delta \in \mathbb{C}, \quad ad - bc \neq 0, \]

then

\[ f(g(z)) = \frac{a}{c} \left( \frac{az + \beta}{\gamma z + \delta} \right) + b = \frac{\alpha z + \beta}{\gamma z + \delta}. \]

since the determinant is $(\alpha z + \beta)(\gamma z + \delta) - (\alpha \beta + \beta \delta)(\alpha z + \gamma) = (ad - bc)(\alpha \gamma - \beta \delta) \neq 0$. (We also have to treat finitely many remaining points, e.g. $c = 0$ etc.

- **Associativity:** composition of functions is associative.

- **Identity is $I : z \mapsto z$.**
• Inverses; let \( f \in \mathcal{M} \). Certainly \( f \) is a permutation of \( \mathbb{C}_\infty \). We need to find an inverse (making sure it’s in \( \mathcal{M} \)). Let

\[
f = \frac{az + b}{cz + d}, \quad g(z) = \frac{dz - b}{-cz + a}
\]

Then \( f(g(z)) = z = g(f(z)) \) \( \forall z \in \mathbb{C}_\infty \), so this works. We now consider special cases.

- Suppose \( c = 0 \), i.e.

\[
f = \frac{az + b}{d} \quad \text{for } z \in \mathbb{C}, \quad f(\infty) = \infty.
\]

Then let

\[
g(z) = \frac{dz - b}{a}, \quad g(\infty) = \infty
\]

and then \( f \circ g(z) = g \circ f(z) \) for \( z \in \mathbb{C}_\infty \) as required.

- Suppose \( c \neq 0 \). If \( z \neq -d/c \) or \( \infty \), then \( f(z) \in \mathbb{C} \) and \( g(f(z)) = z \). Similarly, unless \( z = a/c \) or \( \infty \), then \( f \circ g(z) = z \), so

\[
f : \mathbb{C}\setminus\left\{-\frac{d}{c}\right\} \to \mathbb{C}\setminus\left\{\frac{a}{c}\right\}
\]

is a bijection with inverse \( g \). Now then, if \( z = -d/c \), \( g(f(z)) = g(\infty) = -d/c = z \). And if \( z = \infty \), \( g(f(z)) = g(a/c) = \infty = z \).

It follows that \( f \) is a bijection from \( \mathbb{C}_\infty \) to \( \mathbb{C}_\infty \) with two-sided inverse \( g \).

Therefore \( \mathcal{M} \) is a group; it is called the Möbius group.

7.2.2 Relation to general linear group

**Theorem 7.2.** There is a surjective homomorphism \( \varphi : GL_2(\mathbb{C}) \to \mathcal{M} \) with kernel

\[
Z = \left\{ \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} : \lambda \in \mathbb{C}^* \right\}
\]

taking

\[
\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto f_z \mapsto \frac{az + b}{cz + d}.
\]

**Proof.** Let \( \varphi \) act as above, and we claim that this is a homomorphism;

\[
\varphi \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \varphi \begin{pmatrix} a & b \\ c & d \end{pmatrix} \circ \varphi \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix};
\]

see the proof in 7.1. Surjectivity follows from the definition of \( \mathcal{M} \). For the kernel, we want all \( \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{C}) \) with

\[
\frac{az + b}{cz + d} = z \quad \forall z \in \mathbb{C}_\infty.
\]

Now picking \( z = \infty \) forces \( c = 0 \), picking \( z = 0 \) forces \( b/d = 0 \Rightarrow b = 0 \), and picking \( z = 1 \) forces \( a = d \). Therefore the kernel is \( \{\lambda I_2\} \) as required.

**Corollary 7.3.**

\[
\frac{GL_2(\mathbb{C})}{Z} \approx \mathcal{M}
\]

which can be proved using 7.2 and the isomorphism theorem.

We can repeat this process with the group \( SL_2(\mathbb{C}) \leq GL_2(\mathbb{C}) \), with the restriction that the \( \varphi \) above still be surjective. The kernel in this case is \( \{\pm I\} \), and hence **Corollary 7.4**;

\[
\frac{SL_2(\mathbb{C})}{\{\pm I\}} \approx \mathcal{M}
\]

7.2.3 Relation to modular group

**Exercise.** Show that the set of Möbius transformations satisfying
\[ z \mapsto \frac{az + b}{cz + d}, \quad a, b, c, d \in \mathbb{Z}, \quad ad - bc = 1 \]

is a subgroup \( \mathcal{M}_2 \) of \( \mathcal{M} \). Show also that, if \( SL_2(\mathbb{C}) \) is the group of \( 2 \times 2 \) matrices with integer entries with \( \det 1 \) (called the modular group), then

\[
\frac{SL_2(\mathbb{Z})}{\pm 1} \cong \mathcal{M}_2.
\]

### 7.3 Action of the Möbius group

#### 7.3.1 Triple transitivity

**Theorem 7.5.** The action of \( \mathcal{M} \) on \( \mathbb{C}_\infty \) is triply transitive; that is, if \( z_0, z_1 \) and \( z_\infty \) are distinct in \( \mathbb{C}_\infty \) and \( w_0, w_1, w_\infty \) are distinct in \( \mathbb{C}_\infty \), then \( \exists f \in \mathcal{M} : z_i \mapsto w_i \) for \( i = 0, 1, \infty \).

**Note.** The stabiliser of any three points in this action is trivial, hence the above transformation \( f \) is unique. (\( \mathcal{M} \) is sharply triply transitive.)

**Proof.** To take \( z_0, z_1, z_\infty \) to \( 0, 1, \infty \), we apply

\[
g : z \mapsto \left( \begin{array}{c} z_1 - z_\infty \\ z_1 - z_0 \end{array} \right) \frac{z - z_0}{z - z_\infty}, \quad \infty \mapsto \infty
\]

if all \( z_i \in \mathbb{C} \). Otherwise, we `take limits';

if \( z_\infty = \infty \) then \( g : z \mapsto \frac{z - z_0}{z_1 - z_0}, \infty \mapsto \infty \),

if \( z_1 = \infty \) then \( g : z \mapsto \frac{z - z_0}{z - z_\infty}, \infty \mapsto 1 \),

if \( z_0 = \infty \) then \( z \mapsto \frac{z_1 - z_\infty}{z - z_\infty}, \infty \mapsto 0 \).

Now if \( g \) sends \( z_0, z_1, z_\infty \) to \( 0, 1, \infty \), and \( h \) sends \( w_0, w_1, w_\infty \) to \( 0, 1, \infty \), then \( f = h^{-1}g \) sends \( z_i \) to \( w_i \) for \( i = 0, 1, \infty \). Therefore this \( f \) is the one we're after. QED.

Also, this \( f \) is unique, which can be proved as follows; if \( f' \) also takes \( z_i \) to \( w_i \) for \( i = 0, 1, \infty \), then \( f^{-1}f' \) fixes \( z_0, z_1, z_\infty \). Therefore \( f^{-1}f' = e \) (see below), so \( f' = f \).

#### 7.3.2 Stabilisers

When \( c = 0 \), the stabiliser of \( \infty \) is

\[
\mathcal{M}_\infty = \left\{ z \mapsto \frac{az + b}{d} : ad \neq 0 \right\}
\]

\[
\mathcal{M}_{0, \infty} = \left\{ z \mapsto \frac{az + b}{d} : ad \neq 0 \right\}
\]

\[
\mathcal{M}_{0,1,\infty} = \{ e \}
\]

In fact, the stabiliser of any three points \( z_0, z_1, z_\infty \) is \( \{ e \} \); taking \( g \) as above, \( g^{-1}\mathcal{M}_{0,1,\infty}g = \mathcal{M}_{z_0, z_1, z_\infty} \) (see example sheet).

### 7.4 Conjugacy classes and fixed points of Möbius transformations

**Recall.** It follows from 6.3 that any \( A \in GL_2(\mathbb{C}) \) is conjugate to one of

\[
(i) \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \quad (ii) \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix}, \quad (iii) \begin{pmatrix} 1 & \lambda \\ 0 & \lambda \end{pmatrix}.
\]

Note that if \( PAP^{-1} = B \in GL_2(\mathbb{C}) \), then \( \varphi(P)\varphi(A)(\varphi(P))^{-1} = \varphi(B) \in \mathcal{M} \).

#### 7.4.1 The three conjugacy classes

Conjugacy classes in \( \mathcal{M} \);

(i) \( f \) is conjugate in \( \mathcal{M} \) to \( z \mapsto vz \) for some \( v \in \mathbb{C}, v \neq 0, 1 \)

(ii) \( f \) is conjugate in \( \mathcal{M} \) to \( z \mapsto z \), since \( \begin{pmatrix} 1 & 0 \\ 0 & 1/\lambda \end{pmatrix} = \ker \varphi \).
(iii) $f$ is conjugate in $M$ to $z \mapsto z + \lambda^{-1}$. Now conjugate further; if $g = \varphi\left(\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}\right)$, then $g(z \mapsto z + \lambda^{-1})g^{-1} = z \mapsto z + 1$, since 
\[
\left(\begin{array}{cc} \lambda & 0 \\ 0 & 1 \end{array}\right) \left(\begin{array}{cc} 1 & \lambda - 1 \\ 0 & 1 \end{array}\right) \left(\begin{array}{cc} \lambda & 0 \\ 0 & 1 \end{array}\right) = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right).
\]
Therefore, **Theorem 7.7.** Any non-identity Möbius transformation is conjugate to one of
(i) $z \mapsto vz$ for $v \in \mathbb{C}$, $v \neq 0,1$
(ii) $z \mapsto z + 1$.

7.4.2 Fixed points

**Corollary 7.8.** Fixed points of non-identity Möbius transformations are, correspondingly;
(i) precisely two fixed points, 0 and $\infty$
(ii) precisely one fixed point, $\infty$
If 3 points are fixed, we have the identity, but we already knew that.
**Note.** If $f$ fixes the point $\alpha$, then $gf^{-1}g$ fixes $g(\alpha)$.

7.4.3 Alternative direct approach

Let $f(z) = \frac{az + b}{cz + d}$, $f \neq e$.

Consider the fixed points $z_i$. Then $f(z_i) = z_i$, so
\[
\frac{az_i + b}{cz_i + d} = z_i
\]
so the $z_i$s are roots of the quadratic (or linear) equation $cz^2 + (d - a)z - b = 0$. If this is non-trivial, we will have 1 or 2 complex roots.

If there are two fixed points $z_1$ and $z_2$, let $h \in M$ with $h(z_1) = 0$, $h(z_2) = \infty$. Then $hf h^{-1}$ fixes 0, $\infty$, so $hf h^{-1}(z) = (\alpha/\delta)z$ for $\alpha/\delta \neq 0,1$.

If there is one fixed point $z$, let $h(z_1) = \infty$. Then $hf h^{-1}$ fixes precisely the point $\infty$. Thus $hf h^{-1}(z) = az + \beta$. But $hf h^{-1}$ has precisely one fixed point, $(\alpha^{-1})z + \beta = 0$, so $\alpha - 1 = 0$ and $hf h^{-1}(z) = z + \beta$, and conjugating as in (iii) of 7.7 we get $f$ conjugate to $z \mapsto z + 1$.

7.4.4 Iteration of Möbius functions

**Example 7.9.** This can be used to investigate iteration of Möbius functions. E.g. $f$ fixes a unique point $z_0 \in \mathbb{C}_\infty$. What happens to $f^n(z)$ as $n$ increases?

For some $h \in M$, $t = hf h^{-1} : z \mapsto z + 1$. Now $t^n(z) = z + n$, which goes to $\infty$ as $n$ increases, for any $z \in \mathbb{C}_\infty$. $t^n(h(z))$ also goes to $\infty$. Now, for any $z$, $f^n(z) = h^{-1} t^n(h(z)) \rightarrow h^{-1}(\infty) = z_0$ as $n$ increases.

7.5 Composition of Möbius transformations

**Proposition 7.10.** Any Möbius transformation can be written as a composition of the Möbius transformations of just three types;
(i) $z \mapsto az$ with $a \neq 0$, representing dilation or rotation
(ii) $z \mapsto z + b$, representing translation
(iii) $z \mapsto 1/z$, representing inversion.

**Proof.** Let $f(z) = \frac{az + b}{cz + d}$.
If \( c = 0 \) then \( f = f_2 \circ f_1 \), with \( f_1(z) = (a/d)z \) and \( f_2(z) = z + b/d \).

If \( c \neq 0 \),

\[
f(z) = \frac{az + b}{cz + d} = \frac{a}{c} \frac{z + \frac{b}{c}}{z + \frac{d}{c}} = A + \frac{B}{z + \frac{d}{c}}
\]

with

\[
A = \frac{a}{c}, \quad B = -\frac{ad - bc}{c^2}.
\]

So \( f = f_4 \circ f_3 \circ f_2 \circ f_1 \), with \( f_1(z) = z + d/c, f_2(z) = 1/z, f_3(z) = Bz \) and \( f_4(z) = z + A \).

### 7.6 Circles and straight lines

#### 7.6.1 General equation and relation

**Remark 7.11.** The general equation of a circle or straight line in \( \mathbb{C} \) is

\[
Azz + B\bar{z} + \bar{B}z + C = 0, \quad A, C \in \mathbb{R}, \quad |B|^2 > AC.
\]

For \( z = x + iy \) and \( B = b_1 + ib_2 \), we get \( A(x^2 + y^2) + b_1x + b_2y + C = 0 \) which is the equation of a circle or straight line. (It is a straight line iff \( A = 0 \) and a circle through the origin iff \( C = 0 \).)

Note that a straight line is a ‘circle’ in \( \mathbb{C}_\infty \); it is \( L \cup \{\infty\} \) with \( L \) a line in \( \mathbb{C} \), and called the ‘circle through infinity’.

#### 7.6.2 Effect of Möbius transformations

**Theorem 7.12.** Möbius transformations take circles or straight lines onto circles or straight lines (not respectively).

**Proof.** We need to show that if \( f \in \mathcal{M} \), then \( f \) (circle or straight line) = \( f \) (circle or straight line). By 7.10, it is enough to check this assertion for an inversion, as it is geometrically clear for dilations, rotations and translations.

Let \( f(z) = 1/z \), then \( Azz + B\bar{z} + \bar{B}z + C = 0, \quad A, C \in \mathbb{R} \). Then putting \( w = 1/z \) we have \( w \) satisfying the equation \( Cw\bar{w} + Bw + \bar{B}w + A = 0 \).

**Example.** Find the image of the real axis under

\[
f(z) = \frac{z - i}{z + i}
\]

It is a circle or straight line containing \( f(\infty) = 1, f(0) = -1 \) and \( f(1) = -i \). Therefore we get the unit circle in \( \mathbb{C} \). Further, the image of the half-plane \( \text{Im}(z) \geq 0 \) is the inside of the unit circle, since \( f(i) = 0 \) which is inside the circle.

**Example.**
7.7 Cross-ratios

7.7.1 Definition and formula

Definition. Let \( z_i \in \mathbb{C}_\infty \) be distinct for \( i = 1, 2, 3, 4 \). Then the cross-ratio \( [z_1, z_2, z_3, z_4] \) is the element \( x \) of \( \mathbb{C}_\infty \) such that, if \( f \in \mathcal{M} \) with \( f : z_1 \mapsto 0, z_2 \mapsto 1, z_3 \mapsto \infty \), then \( f : z_4 \mapsto x \). In other words, \( [z_1, z_2, z_3, z_4] = f(z_4) \) for that \( f \in \mathcal{M} \).

Thus, if \( z_j \neq \infty \) then

\[
[z_1, z_2, z_3, z_4] = \frac{z_2 - z_3}{z_2 - z_1} \times \frac{z_4 - z_1}{z_4 - z_3}
\]

(see 7.5). And if some \( z_j = \infty \), then ‘take limits’; e.g.

\[
[\infty, z_2, z_3, z_4] = \frac{z_2 - z_3}{z_4 - z_3}
\]

Remark. Different permutations of \( z_1, z_2, z_3, z_4 \) in different places in literature.

7.7.2 Transformations between sets of four points

Theorem 7.13. Given \( [z_1, z_2, z_3, z_4] \) distinct \( \in \mathbb{C}_\infty \) and \( [w_1, w_2, w_3, w_4] \) distinct \( \in \mathbb{C}_\infty \), there exists \( f \in \mathcal{M} \) with \( f(z_i) = w_i \) for \( 1 \leq i \leq 4 \) iff \( [z_1, z_2, z_3, z_4] = [w_1, w_2, w_3, w_4] \).

Example.

Let \( f \in \mathcal{M} \) take \( z_1 \mapsto 0, z_2 \mapsto 1 \) and \( z_3 \mapsto \infty \), so that \( z_4 \mapsto \) the cross-ratio. Then the circle/straight line on \( 0,1,\infty \) is the real axis. Therefore \( z_1, z_2, z_3, z_4 \) lie on a circle or straight line iff \( [z_1, z_2, z_3, z_4] \in \mathbb{R} \).

Proof (\( \Rightarrow \)). We need to show that \( f \in \mathcal{M} \) preserves cross-ratios, i.e.

\[
[z_1, z_2, z_3, z_4] = [f(z_1), f(z_2), f(z_3), f(z_4)].
\]

Let \( g \in \mathcal{M} \) take \( f(z_1) \mapsto 0, f(z_2) \mapsto 1, f(z_3) \mapsto \infty \) and \( f(z_4) \mapsto [f(z_1), f(z_2), f(z_3), f(z_4)] \). Then \( g f \) takes \( z_1 \mapsto 0, z_2 \mapsto 1, z_3 \mapsto \infty \) and \( z_4 \mapsto [f(z_1), f(z_2), f(z_3), f(z_4)] \). But \( g f(z_4) = [z_1, z_2, z_3, z_4] \) by definition. So the identity is proved.

Proof (\( \Leftarrow \)). Assume \( [z_1, z_2, z_3, z_4] = [w_1, w_2, w_3, w_4] = x \). Let \( h \in \mathcal{M} \) take \( z_1 \mapsto 0, z_2 \mapsto 1, z_3 \mapsto \infty \) and \( z_4 \mapsto x \), and let \( k \in \mathcal{M} \) take \( w_1 \mapsto 0, w_2 \mapsto 1, w_3 \mapsto \infty \) and \( w_4 \mapsto x \). Then \( k^{-1} h \in \mathcal{M} \) with \( z_i \mapsto w_i \), for \( 1 \leq i \leq 4 \).

7.8 Other views of \( \mathbb{C}_\infty \) (non-examinable)

7.8.1 Action of \( GL_2(\mathbb{C}) \) on \( \mathbb{P}'(\mathbb{C}) \)

\( GL_2(\mathbb{C}) \) acts on \( \mathbb{C}^2 \), and so it acts on \( \mathbb{P}'(\mathbb{C}) \) which is the set of 1-dimensional subspaces of \( \mathbb{C}^2 \). The kernel of the action on \( \mathbb{P}'(\mathbb{C}) \) is \( \mathbb{Z} = \{ \lambda I_2 : \lambda \in \mathbb{C}^* \} \).

Look at \( \left( \begin{array}{c} x \\ y \end{array} \right) \), called \( x/y \), which is an element of \( \mathbb{C}^2 \) if \( y \neq 0 \), and \( \infty \) if \( y = 0 \).

If \( y \neq 0 \):

\[
\left( \begin{array}{cc} a & b \\ c & d \end{array} \right) \left( \begin{array}{c} x \\ y \end{array} \right) = \left( \begin{array}{c} ax + b \\ cx + d \end{array} \right), \quad x \mapsto \frac{ax + b}{cx + d}
\]

If \( y = 0 \):

\[
\left( \begin{array}{cc} a & b \\ c & d \end{array} \right) \left( \begin{array}{c} 1 \\ 0 \end{array} \right) = \left( \begin{array}{c} a \\ c \end{array} \right), \quad \infty \mapsto \frac{a}{c}
\]
7.8.2 The Riemann sphere – stereographic projection

Let $S$ be the unit sphere with centre $0$, let $C$ be the equatorial plane $\{(x, y, 0)\}$ and let $\zeta$ be the north pole $(0,0,1)$.

Then the stereographic projection $\varphi : C \to S \setminus \{\zeta\}$ takes $z \mapsto z'$, the point on $S$ that lies on the straight line from $z$ to $\zeta$. As $|z|$ gets large, $|\varphi(z)|$ gets near to $\zeta$.

We can extend the definition of $\varphi$ by letting $z \mapsto z'$ and $\infty \mapsto \zeta$, so that $\varphi : C_\infty \to S$. We can now study any $f : C_\infty \to C_\infty$ as $\varphi f \varphi^{-1} : S \to S$.

8. Matrix groups II – orthogonal groups

8.1 Introduction

8.1.1 Orthogonal group

Definition. The set $O(n) = \{A \in GL_n(\mathbb{R}) : AA^T = I_n\}$.

Note. $A \in O_n \Rightarrow \det A = \pm 1$, because $\det A = \det A^T$ and $AA^T = I$.

Lemma 8.1. $O(n) \leq GL_n(\mathbb{R})$, called the orthogonal group.

Proof. $O(n)$ is a subset, since $A \in O_n \Rightarrow A \in GL_n(\mathbb{R})$. Also, $I \in O(n)$. It is sufficient to prove that if $A, B \in O(n)$ then $A^{-1}B \in O(n)$;

$$A^{-1}B(A^{-1}B)^T = A^{-1}BB^T(A^{-1})^T = A^{-1}A^T = I.$$

8.1.2 Special orthogonal group

Definition. The group $SO(n) = \{A \in O(n) : \det A = 1\}$. We use the isomorphism theorem to show this is a subgroup of $O(n)$; the homomorphism $\det O(n) \to \{\pm 1\} < \mathbb{R}^*$ is surjective, so

$$SO(n) = \ker \det \frac{O(n)}{SO(n)} \simeq C_2.$$

8.1.3 Isometry

Note. $AA^T = I \iff$ the rows of $A$ form an orthonormal basis of $\mathbb{R}^n$, and $A^T A = I \iff$ the columns of $A$ form an orthonormal basis of $\mathbb{R}^n$.

Lemma 8.2. Assume $A \in O(n)$. Then (i) $Ax \cdot Ay = x \cdot y$ and (ii) $|Ax| = |x|$, so $A$ is an isometry on the Euclidean space $\mathbb{R}^n$.

Proof. (i)

$$Ax \cdot Ay = \sum_i \left( \sum_k a_{ik}x_k \right) \left( \sum_j a_{ij}y_j \right) = \sum_{i,k,j} a_{ik}a_{ij}x_ky_j = \sum_{i,k} \delta_{k-j}y_j = \sum_j x_jy_j = x \cdot y$$

(ii) $|Ax|^2 = Ax \cdot Ax = x \cdot x = |x|^2$.

Remark. It is not hard to show that any isometry of the Euclidean space $\mathbb{R}^n$ that fixes $0$ is linear, so $x \mapsto Ax$ with $A \in O(n)$.

8.2 In two dimensions

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

$$AA^T = A^TA = I$$
\[ a^2 + b^2 = 1 = c^2 + d^2, \quad a^2 + c^2 = 1 = b^2 + d^2 \]
\[ ac + bd = 0 = ab + cd \]

For \( 0 \leq \theta < 2\pi \), let
\[
\begin{pmatrix}
 a \\
 b \\
 c \\
 d
\end{pmatrix} = \begin{pmatrix}
 \cos \theta \\
 -\sin \theta \\
 \sin \theta \\
 \cos \theta
\end{pmatrix} \quad \text{so} \quad \begin{pmatrix}
 b \\
 d
\end{pmatrix} = \begin{pmatrix}
 -\sin \theta \\
 \cos \theta
\end{pmatrix} \quad \text{or} \quad \begin{pmatrix}
 \sin \theta \\
 -\cos \theta
\end{pmatrix}.
\]

Therefore \( A \) is of one of the forms
\[
\begin{pmatrix}
 \cos \theta & -\sin \theta \\
 \sin \theta & \cos \theta
\end{pmatrix}, \quad \begin{pmatrix}
 \cos \theta & \sin \theta \\
 \sin \theta & -\cos \theta
\end{pmatrix}
\]

The first of these has determinant 1 and is a rotation by \( \theta \). This can be expressed as \( z \mapsto e^{i\theta}z \), since if \( z = x + iy \), then \( e^{i\theta}z = (\cos \theta + i \sin \theta)(x + iy) = (x \cos \theta - y \sin \theta) + (x \sin \theta + y \cos \theta) \). Matrices of this form make up \( O(2) \).

The second has determinant \(-1\) and is a reflection in the line \( \frac{y}{x} = \frac{\sin \theta/2}{\cos \theta/2} \).

This can be expressed as \( z \mapsto e^{i\theta} \overline{z} \). This can be checked by checking fixed points;
\[
e^{i\theta} \overline{z} = z \iff e^{i\theta} \overline{z} = e^{-i\theta}z \iff ze^{-i\theta} = t \in \mathbb{R} \iff z = te^{-i\theta},
\]
which implies that \( z \mapsto e^{i\theta} \overline{z} \) is a reflection in the line \( te^{i\theta/2} \). Matrices of this form make up \( O(2) \setminus SO(2) \).

### 8.3 In three dimensions

#### 8.3.1 Rotations in \( SO(3) \)

**Theorem 8.4.** Every matrix \( A \in SO(3) \) has an eigenvector with eigenvalue 1 (the axis of a rotation).

**Proof.** Let \( \lambda \) be a real eigenvalue of \( A \) (at least one must exist, since the characteristic polynomial is a cubic in \( \mathbb{R} \)). Let \( \mathbf{v} \) be an eigenvector corresponding to \( \lambda \). Then \( |A\mathbf{v}| = |\lambda| \mathbf{v} \) so \( \lambda = \pm 1 \). The other eigenvalues of \( A \) must be

(i) \( \alpha, \bar{\alpha} \) with \( \alpha \in \mathbb{C} \setminus \mathbb{R} \), in which case \( \det A = \alpha\bar{\alpha} \lambda = |\alpha| \lambda \Rightarrow \lambda = \pm 1 \)

(ii) all real and \( \pm 1 \) (or the eigenvalues are \( +1, -1 \) and \( -1 \)) so we must have \( A = \pm I \).

So either way, \( A \) has +1 as an eigenvalue.

**Exercise.** When do two rotations in \( SO(3) \) commute?

**Theorem 8.5.** Any matrix in \( SO(3) \) is conjugate to one of the form
\[
\begin{pmatrix}
 \cos \theta & -\sin \theta & 0 \\
 \sin \theta & \cos \theta & 0 \\
 0 & 0 & 1
\end{pmatrix}
\]

for some \( \theta \in (0,2\pi) \).

**Proof.** Let \( A \in SO(3) \) and let \( \mathbf{v} \in \mathbb{R}^3 \) with \( A\mathbf{v} = \mathbf{v} \) and \( |\mathbf{v}| = 1 \). Let \( \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \) be the standard orthonormal basis.

Let \( P \in SO(3) \) rotate \( \mathbf{v} \) to \( \mathbf{e}_3 \), i.e. \( P\mathbf{v} = \mathbf{e}_3 \). Then \( PAP^{-1} \) fixes \( \mathbf{e}_3 \) and maps the plane \( \Pi \) (consisting all vectors perpendicular to \( \mathbf{e}_3 \)) to itself. So \( PAP^{-1} \) is a rotation, on \( \Pi \), by some angle \( (0,2\pi) \). Hence the assertion.

#### 8.3.2 Reflections in \( O(3) \setminus SO(3) \)

Now consider the elements of \( O(3) \setminus SO(3) \). E.g. reflections \( r \) in \( \Pi \), a plane through \( \mathbf{0} \), \( \mathbf{n} \) perpendicular to \( \Pi \), \( |\mathbf{n}| = 1 \).

Consider the map \( r_{\mathbf{n}}(\mathbf{x}) = \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{n})\mathbf{n} \) for each \( \mathbf{x} \in \mathbb{R}^3 \). Note that \( \mathbf{n} \mapsto \mathbf{n} \) and also that if \( \mathbf{x} \in \Pi \) then \( \mathbf{x} \mapsto \mathbf{x} \) with respect to a basis \( \mathbf{n}, \mathbf{e}_2, \mathbf{e}_3 \) with \( \mathbf{e}_2, \mathbf{e}_3 \in \Pi, r^2 = 1, \det r = -1 \).
Note 8.7.

\[ O(3) = SO(3) \bigcup \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} SO(3) \]

where the symbol represents a disjoint union. In fact, \( O(3) \approx SO(3) \times \{ \pm I_3 \} \) - see example sheet 4.

Note 8.8. If \( a \) and \( b \) are as shown in the diagram with \( |a| = |b| = 1 \), i.e.

\[ n = \frac{a - b}{|a - b|} \]

then \( r_n : a \leftrightarrow b \).

8.3.3 Product of reflections

Theorem 8.9. Any element of \( O(3) \) is the product of at most three reflections.

Proof. Let \( f(x) = Ax \) with \( A \in O(3) \). Let \( e_1, e_2, e_3 \) be the standard orthonormal basis of \( \mathbb{R}^3 \).

Now \( |f(e_3)| = |e_3| = 1 \), since \( A \) is an isometry. So there is a reflection \( r_1 \) in a plane through \( 0 \) with \( r_1 f(e_3) = e_3 \). Then \( r_1 f \) fixes \( e_3 \), so maps the plane \( \Pi \) generated by \( e_1, e_2 \) onto itself (\( \Pi \) is \( e_3^\perp \)).

So there is a reflection \( r_2 \) with \( r_2(e_3) = e_3 \) and \( r_2(f(e_2)) = e_2 \). So \( r_2 r_1 f : e_3 \mapsto e_3 \) and \( e_2 \mapsto e_2 \), which means \( e_1 \mapsto \pm e_1 \).

Now take \( r_3 \) to be a reflection with \( n = e_1 \) if \( r_2 r_1 f(e_1) = -e_1 \) and \( r_3 = I \) otherwise. So \( r_3 r_2 r_1 f = I \), so \( f = r_1 r_2 r_3 \).