Group Theory (MATH33300)
University of Bristol

Lecturer: Matthew Tointon
m.tointon@bristol.ac.uk

2025/26
Teaching block 1 (weeks 1–12)

January 29, 2025

Contents

Chapter 1

Introduction

This course follows on from *Introduction to Group Theory* (MATH10010). We will recall all of the necessary definitions from that course, but we will quote some results from it without proof.

1.1 Groups

Operations like $+$ and $\times$ that combine two numbers to produce another are called *binary operations*. More generally and formally we have the following definition.

Definition (binary operation)**.** Let X be a set. Then a *binary operation* $*$ on X is a map

$$\begin{array}{ccc} X \times X & \to & X \\ (x, y) & \mapsto & x * y. \end{array}$$

The operation $*$ is called *associative* if $x * (y * z) = (x * y) * z$ for all $x, y, z \in X$, in which case we write $x * y * z$ to mean $x * (y * z)$ (and hence $(x * y) * z$ also). An *identity* element for $*$ is an element $e \in X$ such that $e * x = x$ for every $x \in X$.

Examples.

(1) The operations of $+$, $\times$ and $-$ are all binary operations on $\mathbb{R}$, but $\div$ is not because $x \div 0$ is not defined. On the other hand, $\times$ and $\div$ are binary operations on $\mathbb{R} \setminus \{0\}$, but $+$ and $-$ are not because e.g. $1 - 1$ is not defined.

(2) Other examples of binary operations include composition on the set of functions $X \to X$; matrix multiplication on the set of $n \times n$ real matrices; and $\cap$ and $\cup$ on the set of subsets of a set. These are all associative. Which of them have identities?

(3) Associativitiy means that successive operations can be written unambiguously without brackets. For example, it is clear how to evaluate $3 \times 7 \times 6$, but should $8 \div 4 \div 2$ equal 4 or 1?

(4) The operation of $\div$ *is* associative on $\{-1, 1\}$. You can easily check this case by case, but an even quicker way is to observe that $\div$ and $\times$ coincide on $\{-1, 1\}$.

Definition (group)**.** A *group* $(G, *)$ is a set G with a binary operation $*$ such that

- $*$ is associative;
- G contains an identity element e for $*$; and
- for every element $g \in G$ there exists an element $g^{-1} \in G$ such that $g^{-1} * g = e$.

The operation $*$ is called the *group operation*. The element g^{-1} is called the *inverse* of g.

When dealing with abstract groups we usually write just G instead of $(G, *)$, and express the group operation by xy rather than $x * y$. We also write x^n to mean a product of n copies of x, and x^{-n} to mean a product of n copies of x^{-1}.

Warning. Different instances of the letter e can mean the identity in different groups, even in the same line. For example, if G and H are groups and $\varphi : G \to H$ is a map then $\varphi(e) = e$ means that φ maps the identity of G to the identity of H.

We defined identities and inverses in terms of what they do when we multiply by them on the left. The following lemma shows that they behave in a similar way when we multiply by them on the right.

Lemma 1.1. *Let G be a group. Then $ge = g$ and $gg^{-1} = e$ for all $g \in G$. Moreover, both the identity element and the inverse of any given element are unique, in the sense that if $xg = g$ or $gx = g$ for some $x, g \in G$ then $x = e$, and if $xg = e$ or $gx = e$ for some $x, g \in G$ then $x = g^{-1}$.*

Proof. First note that if $a \in G$ satisfies $a^2 = a$ then, multiplying both sides on the left by a^{-1}, we see that $a = e$. Since $(gg^{-1})^2 = g(g^{-1}g)g^{-1} = gg^{-1}$, this implies that $gg^{-1} = e$ for all $g \in G$, as required. This in turn implies that $ge = gg^{-1}g = g$ for all $g \in G$, as required.

We now prove uniqueness of inverses and the identity. If $xg = g$ then multiplying on the right by g^{-1} we obtain that $x = e$, whilst if $xg = e$ then multiplying on the right by g^{-1} we see that $x = g^{-1}$. The proofs for if $gx = g$ or $gx = e$ are similar and left as an exercise. □

We will use Lemma 1.1 throughout this course without explicit mention, and you may do the same in your homework and the exam.

Examples 1.2. Here are some examples of groups, most of which you saw in *Introduction to Group Theory*. We will use these repeatedly to illustrate various concepts throughout the course, so it is worth spending some time (re)familiarising yourself with them.

(1) The set $G = \{e\}$ with its only possible binary operation ($ee = e$) is a group, called the *trivial group*.

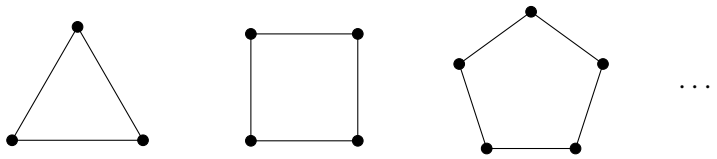
(2) The sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ all form groups under addition, and the set $\mathbb{Z}_n$ of integers mod n forms a group under addition mod n. For brevity we will generally write just $\mathbb{Z}$, $\mathbb{Z}_n$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ rather than $(\mathbb{Z}, +)$, $(\mathbb{Z}_n, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$. In groups where the operation is addition, we generally write mx, rather than x^m, to mean a sum $x + x + \cdots + x$ of m copies of x.

(3) The sets $\mathbb{R}\setminus\{0\}$, $\mathbb{R}_{>0}=(0,\infty)$ and $\mathbb{C}\setminus\{0\}$ all form groups under multiplication. We generally abbreviate

$$\begin{aligned}\mathbb{R}^* &= (\mathbb{R}\setminus\{0\},\times)\\ \mathbb{R}_{>0} &= (\mathbb{R}_{>0},\times)\\ \mathbb{C}^* &= (\mathbb{C}\setminus\{0\},\times).\end{aligned}$$

(4) Let X be a non-empty set. Then a *permutation* of X is a bijection $X\to X$. The set of all permutations of a set X forms a group under composition, called the *symmetric group on* X and denoted $\mathrm{Sym}(X)$. In the special case where $X=\{1,\dots,n\}$, we often write S_n as shorthand for $\mathrm{Sym}(\{1,\dots,n\})$.

(5) For each $n\geqslant 3$, let P_n be a regular n-gon in the Euclidean plane.



P_3 P_4 P_5

A *symmetry* of P_n is a permutation φ of the vertices of P_n such that if vertices x and y are joined by an edge then so are the vertices $\varphi(x)$ and $\varphi(y)$. The *dihedral group* D_{2n} is the group of all symmetries of P_n under composition.

Recall from *Introduction to Group Theory* that symmetries of P_n can also be represented by *rotations* and *reflections*. More concretely, if r is an anticlockwise rotation of P_n by about its centre by an angle of $2\pi/n$, and s is a reflection in a line containing the centre and one of the vertices, then elements of the form r^i are rotations, elements of the form r^is are reflections, and

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

Note that

$$r^n=e,\quad s^2=e,\quad sr=r^{-1}s, \tag{1.1}$$

and that this is enough information to allow us to evaluate any expression $r^ks^\ell r^ms^n\cdots$ in D_{2n}. **Warning.** Some authors write D_n instead of D_{2n}.

(6) Call a map $\varphi:\mathbb{Z}\to\mathbb{Z}$ a *symmetry* of $\mathbb{Z}$ if $|\varphi(m)-\varphi(n)|=|m-n|$ for all $m,n\in\mathbb{Z}$. Analogously to symmetries of polygons, symmetries of $\mathbb{Z}$ can be represented by *translations* and *reflections*. For example, for each $n\in\mathbb{Z}$ we have the *translation* of $\mathbb{Z}$ by n defined by $m\mapsto m+n$. The *infinite dihedral group* D_∞ is the group of all symmetries of $\mathbb{Z}$. More concretely, if r is the translation $n\mapsto n+1$ and s is the reflection $n\mapsto -n$ then

$$D_\infty=\left\{\begin{array}{c}\dots,r^{-2},r^{-1},e,r,r^2,\dots\\ \dots,r^{-2}s,r^{-1}s,s,rs,r^2s,\dots\end{array}\right\}.$$

As with the finite dihedral groups, $s^2 = e$ and $sr = r^{-1}s$, and this is enough to evaluate any expression $r^k s^\ell r^m s^n \cdots$.

(7) Let $n \in \mathbb{N}$. Then the set of invertible $n \times n$ real matrices form a group under matrix multiplication, denoted $\mathrm{GL}_n(\mathbb{R})$.

(8) In *Introduction to Group Theory* you saw the *direct product* of two groups. In fact this notion easily generalises to more than two groups. Indeed, if $G_1, \ldots, G_n$ are groups then the set

$$G_1 \times \cdots \times G_n = \{(a_1, \ldots, a_n) : a_i \in G_i,\ 1 \leqslant i \leqslant n\}$$

is a group with respect to the binary operation

$$(a_1, \ldots, a_n)(b_1, \ldots, b_n) := (a_1 b_1, \ldots, a_n b_n).$$

The identity element is $(e, \ldots, e)$, and $(a_1, \ldots, a_n)^{-1} = (a_1^{-1}, \ldots, a_n^{-1})$. We call $G_1 \times \cdots \times G_n$ the *direct product* of $G_1, \ldots, G_n$ (or sometimes the *outer direct product*). It is also written $\prod_{i=1}^n G_i$. We often abbreviate the direct product of n copies of the same group G by G^n.

(9) The *Klein 4-group* K_4 is defined to be the direct product $K_4 = \mathbb{Z}_2^2 = \mathbb{Z}_2 \times \mathbb{Z}_2$.

1.2 Basic terminology

In this course we will use the following notation:

- $\mathbb{N} = \{1, 2, 3, \ldots\}$
- $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$
- $[n] = \{1, \ldots, n\}$
- $[n]_0 = [n] \cup \{0\}$

Definition (commuting elements)**.** We say that elements a and b in some group *commute* if $ab = ba$.

Definition (abelian group)**.** We say G is *abelian* (or *commutative*) if $ab = ba$ for all $a, b \in G$.

I encourage you to check which of Examples 1.2 are abelian.

Definition (order of a group)**.** We write $|G|$ for the number of elements of a group G, called the *order* of G. If $|G| < \infty$ then we say that G is *finite*, whilst if $|G| = \infty$ then we say that G is *infinite*.

Definition (order of an element)**.** Let g be an element of a group. Then we define the *order* of g, written $|g|$, as follows.

- If $g^n \neq e$ for every $n \in \mathbb{N}$ then $|g| = \infty$.
- Otherwise, $|g| = \min\{n \in \mathbb{N} : g^n = e\}$.

Note that $|g| = 1$ if and only if $g = e$.

Lemma 1.3. *Let x be an element of a group and $m, n \in \mathbb{Z}$.*

(1) If $|x| < \infty$ then $x^m = e$ if and only if $|x|$ divides m.

(2) If $|x| < \infty$ then $x^m = x^n$ if and only if $|x|$ divides $m - n$.

(3) If $|x| = \infty$ then $x^m = x^n$ if and only if $m = n$.

Proof. If $|x| < \infty$ then we may write $m = q|x| + r$ with $q \in \mathbb{Z}$ and $r \in \{0, \dots, |x| - 1\}$. We then have $x^m = (x^{|x|})^q x^r = x^r$, which by definition of $|x|$ is e if and only if $r = 0$. This gives (1). For arbitrary x we have $x^m = x^n \iff e = x^{m-n}$, which implies (3) directly and implies (2) when combined with (1). □

Definition (torsion group; torsion-free group)**.** A group is said to be a *torsion* group (or *periodic* group) if every element has finite order. It is said to be *torsion-free* if no non-identity element has finite order.

Warning. A group can be neither torsion nor torsion-free, e.g. a direct product of a torsion group and a torsion-free group.

1.3 Set multiplication

We can extend the notation of the group operation to subsets as well as elements, as follows.

Definition (product set)**.** If X, Y are subsets of a group G, then we define

$$XY = \{xy : x \in X, y \in Y\} \subseteq G,$$

called the *product set* of X and Y. We also set

$$X^{-1} = \{x^{-1} : x \in X\},$$

and write

$$X^n = \overbrace{XX \cdots X}^{n \text{ copies}} = \{x_1 \cdots x_n : x_i \in X\}.$$

If $X = \{x\}$ is a singleton then we normally write xY instead of $\{x\}Y$. We sometimes adjust this notation to reflect the group operation, e.g. writing $X + Y = \{x + y : x \in X, y \in Y\}$ and $-X = \{-x : x \in X\}$ if the group operation is addition.

Examples.

(1) In any group G we have $X\varnothing = \varnothing X = \varnothing$ for all $X \subseteq G$.

(2) In any group G we have $GX = XG = G$ and $eX = Xe = X$ for all non-empty $X \subseteq G$.

(3) Let O be the set of odd numbers and E the set of even numbers in the group $\mathbb{Z}$. Then the familiar statement that the sum of every two odd numbers is even can be expressed as $O + O \subseteq E$.

(4) Let $X = \{1, \ldots, m\} \subseteq \mathbb{Z}_n$. Then $X + X = \{2, 3, \ldots, 2m\}$.

Note that

$$(XY)^{-1} = Y^{-1}X^{-1} \tag{1.2}$$

for any subsets X, Y of any group.

Definition (set multiplication)**.** Let G be a group, and write $\mathcal{P}(G)$ for the *power set* of G, that is the set of all subsets of G. Then the binary operation

$$\begin{array}{ccc} \mathcal{P}(G) \times \mathcal{P}(G) & \to & \mathcal{P}(G) \\ (X, Y) & \mapsto & XY \end{array}$$

on the subsets of G is called *set multiplication*.

Lemma 1.4. *Let G be a group. Then set multiplication is associative on $\mathcal{P}(G)$.*

Proof. This follows directly from the associativity of G: given $X, Y, Z \subseteq G$ we have

$$\begin{aligned} (XY)Z &= \{(xy)z : x \in X,\, y \in Y,\, z \in Z\} \\ &= \{x(yz) : x \in X,\, y \in Y,\, z \in Z\} \\ &= X(YZ). \end{aligned}$$

□

Note that set multiplication also has an identity: $eX = Xe = X$ for every $X \subseteq G$. However:

Warning. Subsets of G do not in general have inverses with respect to set multiplication, so set multiplication does not make $\mathcal{P}(G)$ into a group. In particular, it is worth emphasising that the set X^{-1} is *not*, in general, an inverse to X with respect to set multiplication (we always have $GG^{-1} = G$, for example).

1.4 Isomorphic groups

Definition (isomorphism; isomorphic groups)**.** Let $(G, *)$ and $(H, \bullet)$ be groups. Then an *isomorphism* from $(G, *)$ to $(H, \bullet)$ is a bijection $\varphi : G \to H$ such that $\varphi(x * y) = \varphi(x) \bullet \varphi(y)$ for all $x, y \in G$. If such a map exists, we say that G and H are *isomorphic*, denoted $G \cong H$.

Examples 1.5.

(1) For each $n \geqslant 2$ the set $Z = \{e^{2\pi ik/n} : k = 0, 1, \ldots, n-1\}$ is a group under multiplication, and the map $Z \to \mathbb{Z}_n$, $e^{2\pi ik/n} \mapsto k$ is an isomorphism.

(2) For each $n \geqslant 3$ the set $C_n = \{e, r, \ldots, r^{n-1}\} \subset D_{2n}$ of rotations of the regular n-gon form a group under composition, and the map $C_n \to \mathbb{Z}_n$, $r^k \mapsto k$ is an isomorphism.

(3) The map $\log : (\mathbb{R}_{>0}, \times) \to (\mathbb{R}, +)$ is an isomorphism.

An isomorphism from G to H is essentially a way of relabelling the elements and binary operation of G in order to obtain H.

Lemma 1.6 (from *Introduction to Group Theory*)**.** *Let G, H be groups, let $\varphi : G \to H$ be an isomorphism, and let $g \in G$. Then*

(1) $\varphi(e) = e$,

(2) $\varphi(g^{-1}) = \varphi(g)^{-1}$,

(3) $\varphi(g^i) = \varphi(g)^i$ *for all* $i \in \mathbb{Z}$, *and*

(4) $|\varphi(g)| = |g|$.

Given a group G, we write ι_G for the identity map $G \to G$. This is trivially an isomorphism, so we always have

$$G \cong G. \tag{1.3}$$

Lemma 1.7 (from *Introduction to Group Theory*)**.** *Let G, H, K be groups, and let $\varphi : G \to H$ and $\psi : H \to K$ be isomorphisms. Then*

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

(2) $\psi \circ \varphi : G \to K$ *is an isomorphism.*

Note that (1.3) and Lemma 1.7 imply that $\cong$ is an equivalence relation on any set of groups.

Lemma 1.8. *Let G be a group and let $g \in G$. Then the map $G \to G$ defined by $x \mapsto gxg^{-1}$ is an isomorphism from G to itself.*

Proof. This map is injective because $gxg^{-1} = gyg^{-1}$ implies $x = g^{-1}gyg^{-1}g = y$, and surjective because for every $x \in G$ we have $g(g^{-1}xg)g^{-1} = x$. Finally, $(gxg^{-1})(gyg^{-1}) = gx(g^{-1}g)yg^{-1} = gxyg^{-1}$. □

Definition (conjugation)**.** Given a group G and an element $g \in G$, the map $G \to G$ defined by $x \mapsto gxg^{-1}$ is called *conjugation by g*. Two elements $x, y \in G$ are said to be *conjugate* in G if there exists $g \in G$ such that $y = gxg^{-1}$. More generally, two subsets X, Y of a group G are said to be *conjugate in G* if there exists $g \in G$ such that $Y = gXg^{-1}$.

1.5 Subgroups

Definition (closed subset; restriction of a binary operation)**.** Let X be a set with a binary operation $*$. Then a subset $A \subseteq X$ is said to be *closed* under $*$ if $a * b \in A$ whenever $a, b \in A$. The binary operation

$$\begin{array}{ccc} A \times A & \to & A \\ (a, b) & \mapsto & a * b \end{array}$$

is then called the *restriction* of $*$ to A.

Definition (subgroup)**.** A subset H of G is said to be a *subgroup* of G, denoted $H \leqslant G$, if H is closed under the operation of G, and the restriction of the operation of G makes H a group. A subgroup $H \leqslant G$ with $H \neq G$ is called a *proper subgroup*, denoted $H < G$.

Examples.

(1) In an arbitrary group G we have $\{e\} \leqslant G$ and $G \leqslant G$.

(2) The group $C_n = \{e, r, \ldots, r^{n-1}\}$ of rotations described in Example 1.5 (2) is a subgroup of the dihedral group D_{2n}.

(3) Although $(\mathbb{R}\backslash\{0\}) \subseteq \mathbb{R}$, the group $\mathbb{R}^*$ is *not* a subgroup of $(\mathbb{R}, +)$, since $\times$ is not the restriction of $+$ to $\mathbb{R} \setminus \{0\}$.

Lemma 1.9 (from *Introduction to Group Theory*)**.** *Let G be a group. Then a subset $H \subseteq G$ is a subgroup of G if and only if*

(i) $e \in H$;

(ii) H *is closed under the group operation; and*

(iii) $x^{-1} \in H$ *whenever* $x \in H$.

Example. One can check using Lemma 1.9 that

$$H = \{e, r^2, r^4, \ldots, r^{2n-2}, s, r^2s, r^4s, \ldots, r^{2n-2}s\} \leqslant D_{4n}. \tag{1.4}$$

Indeed, given $k, \ell \in [n-1]_0$ and $i, j \in \{0, 1\}$, it follows from (1.1) that

$$\begin{aligned} r^{2k}s^i r^{2\ell}s^j &= r^{2k}r^{(-1)^i 2\ell}s^i s^j \\ &= r^{2(k+(-1)^i\ell)}s^{i+j}, \end{aligned}$$

so H is closed. Furthermore, $(r^{2k})^{-1} = r^{2(n-k)} \in H$ and $(r^{2k}s)^{-1} = r^{2k}s \in H$. Note that $H \cong D_{2n}$.

It is instructive to rewrite Lemma 1.9 using set-multiplication notation, as follows.

Lemma 1.10. *Let G be a group, and let $A \subseteq G$. Then the following are equivalent:*

(i) $A \leqslant G$;

(ii) $e \in A$, $AA \subseteq A$ *and* $A^{-1} \subseteq A$;

(iii) $e \in A$, $AA = A$ *and* $A^{-1} = A$.

Proof. The equivalence (i) $\iff$ (ii) is exactly Lemma 1.9, and the implication (iii) $\implies$ (ii) is trivial. Suppose then that (ii) holds. The fact that $A^{-1} \subseteq A$ implies that $(A^{-1})^{-1} \subseteq A^{-1}$, which is to say $A \subseteq A^{-1}$, and hence $A^{-1} = A$ as required. Moreover, the fact that $e \in A$ implies that $A \subseteq AA$, so since $AA \subseteq A$ we also have $AA = A$ as required. □

We now consider a more general example.

Definition (centre)**.** Given a group G, the *centre* of G, denoted $Z(G)$, is defined via

$$Z(G) = \{z \in G : gz = zg \,\forall g \in G\}.$$

An element, subset or subgroup of G that is contained in $Z(G)$ is called a *central* element, subset or subgroup.

Lemma 1.11. *Let G be a group. Then $Z(G) \leqslant G$.*

Proof. It is clear that $e \in Z(G)$ and $Z(G)$ is closed under the group operation. Moreover,

$$\begin{aligned} z \in Z(G) &\implies zg = gz \text{ for all } g \in G \\ &\implies gz^{-1} = z^{-1}g \text{ for all } g \in G \\ &\implies z^{-1} \in Z(G), \end{aligned}$$

so the lemma follows from Lemma 1.9. □

It is often more convenient to use the following result to check whether a certain set is a subgroup, rather than Lemma 1.9.

Lemma 1.12 (subgroup test)**.** *Let A be a subset of a group G. Then the following are equivalent.*

(i) $A \leqslant G$.

(ii) $A \neq \varnothing$ *and* $x^{-1}y \in A$ *whenever* $x, y \in A$.

(iii) $A \neq \varnothing$ *and* $xy^{-1} \in A$ *whenever* $x, y \in A$.

Proof. We'll start by proving (i) $\iff$ (ii). If $A \leqslant G$, then Lemma 1.9 implies that $e \in A$, hence $A \neq \varnothing$, and also that $x^{-1}y \in A$ whenever $x, y \in A$. Conversely, suppose $A \neq \varnothing$ and $x^{-1}y \in A$ whenever $x, y \in A$. Then there exists $a \in A$, so in particular we have $e = a^{-1}a \in A$. Then, for every $x, y \in A$ we have $x^{-1} = x^{-1}e \in A$. Then, for every $x, y \in A$ we have $xy = (x^{-1})^{-1}y \in A$. Thus A is a subgroup by Lemma 1.9.

One can prove (i) $\iff$ (iii) similarly, or deduce it from (i) $\iff$ (ii) and the fact that A is a subgroup if and only if A^{-1} is a subgroup, which follows easily from Lemma 1.10. □

In set-multiplication notation, Lemma 1.12 says that $A \leqslant G$ if and only if $A \neq \varnothing$ and $A^{-1}A = A$, if and only if $A \neq \varnothing$ and $AA^{-1} = A$ (see Exercise Sheet 1 Q5(a)).

Lemma 1.13. *Let G be a group and let $\mathcal{A}$ be a set of subgroups of G. Then $\bigcap_{H \in \mathcal{A}} H$ is also a subgroup of G.*

Proof. Certainly $e \in \bigcap_{H \in \mathcal{A}} H$, so $\bigcap_{H \in \mathcal{A}} H \neq \varnothing$. Given $x, y \in \bigcap_{H \in \mathcal{A}} H$, by definition we have $x, y \in H$ for every $H \in \mathcal{A}$, and hence $x^{-1}y \in H$ for every $H \in \mathcal{A}$. By definition this means that $x^{-1}y \in \bigcap_{H \in \mathcal{A}} H$, and so $\bigcap_{H \in \mathcal{A}} H$ is indeed a subgroup by the subgroup test (Lemma 1.12). □

Lemma 1.14. *Let $H, K \leqslant G$. Then HK is a subgroup of G if and only if $HK = KH$.*

Proof. If HK is a subgroup then $HK = (HK)^{-1} = K^{-1}H^{-1} = KH$ by Lemma 1.10 and (1.2). Conversely, HK is certainly non-empty because $e = ee \in HK$, and if $HK = KH$ then

$$\begin{aligned} HK(HK)^{-1} &= HKK^{-1}H^{-1} && \text{(by (1.2))} \\ &= HKKH && \text{(by Lemma 1.10)} \\ &= HKH && \text{(by Lemma 1.10)} \\ &= HHK \\ &= HK && \text{(by Lemma 1.10),} \end{aligned}$$

so HK is a subgroup by the subgroup test (Lemma 1.12). □

1.6 Generating sets

Given a subset X of a group G, we define the *subgroup generated by* X, roughly speaking, as the 'smallest subgroup of G containing X'. We make this precise via the following proposition.

Proposition 1.15. *Let G be a group and $X \subseteq G$, and set*

$$\langle X \rangle = \{x_1^{\epsilon_1} x_2^{\epsilon_2} \cdots x_m^{\epsilon_m} : x_i \in X,\, \epsilon_i \in \{1, -1\},\, m \in \mathbb{N}_0\},$$

with the empty product corresponding to the case $m = 0$ defined to be the identity. Then $\langle X \rangle$ is a subgroup of G containing X.

Definition (subgroup generated by a set)**.** Let G be a group, and let $X \subseteq G$. We call

$$\langle X \rangle = \{x_1^{\epsilon_1} x_2^{\epsilon_2} \cdots x_m^{\epsilon_m} : x_i \in X,\, \epsilon_i \in \{1, -1\},\, m \in \mathbb{N}_0\},$$

the *subgroup generated by* X. If $\langle X \rangle = G$ then we say that X is a *generating set* for G. We say that G is *finitely generated* if there exists $X \subseteq G$ with $|X| < \infty$ such that $G = \langle X \rangle$.

Given $x_1, \ldots, x_k \in G$, we write $\langle x_1, \ldots, x_k \rangle$ to mean $\langle \{x_1, \ldots, x_k\} \rangle$.

Note in particular that, by Lemma 1.9, every subgroup containing X also contains $\langle X \rangle$, which gives a precise sense in which $\langle X \rangle$ is the 'smallest' such subgroup.

Proof of Proposition 1.15. Trivially $X \subseteq \langle X \rangle$. We have $e \in \langle X \rangle$ by definition, and given $x, y \in \langle X \rangle$, say $x = x_1^{\epsilon_1} x_2^{\epsilon_2} \cdots x_m^{\epsilon_m}$ and $y = y_1^{\delta_1} y_2^{\delta_2} \cdots y_n^{\delta_n}$ with $x_i, y_j \in X$ and $\epsilon_i, \delta_j \in \{1, -1\}$, we have

$$x^{-1}y = (x_1^{\epsilon_1} \cdots x_m^{\epsilon_m})^{-1}(y_1^{\delta_1} \cdots y_n^{\delta_n}) = x_m^{-\epsilon_m} \cdots x_1^{-\epsilon_1} y_1^{\delta_1} \cdots y_n^{\delta_n} \in \langle X \rangle,$$

and so $\langle X \rangle$ is a subgroup by the subgroup test (Lemma 1.12). □

Examples. We have $\langle \varnothing \rangle = \{e\}$ in any group. The group $\mathbb{Z}$ is generated by 1. The dihedral group D_{2n} is generated by r and s. The group $H < D_{4n}$ defined in (1.4) is generated by r^2 and s.

See Exercise Sheet 1 Q12 for an equivalent definition of $\langle X \rangle$.

1.7 Cyclic groups

Definition (cyclic group)**.** A group G is said to be *cyclic* if it is generated by a single element, that is if there exists an element $x \in G$ such that $G = \langle x \rangle$. Such an element is called a *generator* of G.

Examples. Both $\mathbb{Z}$ and $\mathbb{Z}_n$ are generated by 1, so are cyclic.

Lemma 1.16. *Let x be an element in some group. Then $\langle x \rangle = \{x^n : n \in \mathbb{Z}\}$ and $\langle x \rangle$ is abelian.*

Proof. We have $\langle x \rangle = \{x^n : n \in \mathbb{Z}\}$ by definition. Since $x^m x^n = x^{m+n} = x^n x^m$ for every $m, n \in \mathbb{Z}$, it follows also that $\langle x \rangle$ is abelian. □

Lemma 1.17. *Let x be an element in some group. If $|x| = n \in \mathbb{N}$ then $\langle x \rangle = \{e, x, \ldots, x^{n-1}\}$ with all elements distinct, whilst if $|x| = \infty$ then $\langle x \rangle = \{\ldots, x^{-2}, x^{-1}, e, x, x^2, \ldots\}$ with all elements distinct. In particular, $|\langle x \rangle| = |x|$.*

Proof. This is immediate from Lemma 1.16 and Lemma 1.3. □

Proposition 1.18 (classification of cyclic groups)**.** *Let x be an element in some group. If $|x| = \infty$ then $\langle x \rangle \cong \mathbb{Z}$, whilst if $|x| = n \in \mathbb{N}$ then $\langle x \rangle \cong \mathbb{Z}_n$.*

Proof. We will prove the case $|x| = n$; the case $|x| = \infty$ is almost identical, but simpler. Lemma 1.17 implies that we may define a bijection $\varphi : G \to \mathbb{Z}_n$ via $x^m \mapsto m \pmod n$. To see that this bijection is an isomorphism, note that $\varphi(x^\ell x^m) = \varphi(x^{\ell+m}) = \varphi(x^{\ell+m \pmod n})$ by Lemma 1.3, and hence $\varphi(x^\ell x^m) = \ell + m \pmod n = \varphi(x^\ell) + \varphi(x^m) \pmod n$, as required. □

Examples.

(1) In the dihedral group

$$D_{2n} = \{\underbrace{e, r, r^2, \ldots, r^{n-1}}_{\text{rotations}}, \underbrace{s, rs, r^2 s, \ldots, r^{n-1}s}_{\text{reflections}}\},$$

each reflection has order 2, so each $\{e, r^i s\}$ is a cyclic subgroup isomorphic to $\mathbb{Z}_2$.

(2) In the cyclic group $\mathbb{Z}_6$ we have

$$|0| = 1, \quad |1| = |5| = 6, \quad |2| = |4| = 3, \quad |3| = 2,$$

so 1 and 5 are the only generators of $\mathbb{Z}_6$.

Lemma 1.19 (from *Introduction to Group Theory*)**.** *Let G be a cyclic group and suppose $H \leqslant G$. Then H is also cyclic.*

1.8 Cosets

Definition (coset)**.** Let G be a group, let $H \leqslant G$ be a subgroup and let $x \in G$. The subset

$$xH = \{xh : h \in H\} \subseteq G$$

is called a *left coset* of H in G. We write G/H for the set of left cosets of H. The subset

$$Hx = \{hx : h \in H\} \subseteq G$$

is called a *right coset*.

Example. The left cosets of $H = \{e, s\}$ in D_6 are $eH = sH = \{e, s\}$, $rH = rsH = \{r, rs\}$, and $r^2H = r^2sH = \{r^2, r^2s\}$.

Given $h \in H$ we have $Hh = H = hH$. Indeed, hH and Hh are trivially contained in H because it is closed under the group operation, and given $x \in H$ we have $x = h(h^{-1}x) \in hH$ and $x = (xh^{-1})h \in Hh$, so that H is contained in both hH and Hh.

However, if G is non-abelian then in general we can have $Hg \neq gH$ for $g \in G \setminus H$; this is the case, for example, with the subgroup $\{e, s\} \leqslant D_6$ considered above. Nonetheless, the study of right cosets is essentially equivalent to the study of left cosets, thanks to the following result.

Lemma 1.20. *Let G be a group and $H \leqslant G$. Then the map $xH \mapsto Hx^{-1}$ is a bijection from the set of left cosets of H to the set of right cosets.*

Proof. To see that this map is well defined and injective, note that $(xH)^{-1} = Hx^{-1}$ by (1.2) and Lemma 1.10, and hence $xH = yH \iff (xH)^{-1} = (yH)^{-1} \iff Hx^{-1} = Hy^{-1}$. It is trivially surjective. □

Lemma 1.21 (basic properties of cosets)**.** *Let G be a group and $H \leqslant G$. Then the following hold.*

(1) $G = \bigcup_{x \in G} xH$.

(2) $xH = yH$ *if and only if* $x \in yH$.

(3) $xH = yH$ *if and only if* $x^{-1}y \in H$.

(4) For $x, y \in G$, *either* $xH = yH$ *or* $xH \cap yH = \varnothing$.

(5) $|xH| = |H|$ *for all* $x \in G$.

Proof.

(1) If $x \in G$ then $x = xe \in xH$.

(2) If $xH = yH$ then $x = xe \in xH = yH$. Conversely, if $x \in yH$, say $x = yh$, then $xH = yhH = yH$.

(3) We have $xH = yH \iff H = x^{-1}yH \iff x^{-1}y \in H$ by (2).

(4) If $xH \cap yH \neq \varnothing$ then there exists $g \in xH \cap yH$, and so $xH = gH = yH$ by (2).

(5) The map $H \to xH$, $h \mapsto xh$, is a bijection.

□

Remark. We can define a relation $\sim$ on G by setting $x \sim y \iff y \in xH$. It is easy to check that $\sim$ is an equivalence relation, and xH is the equivalence class containing x.

Definition (index)**.** Let G be a group and let $H \leqslant G$. Then the number of distinct left cosets of H in G is called the *index* of H in G, and denoted $[G : H]$. Note that $[G : H] = \infty$ is possible.

Example. $[D_6 : \{e, s\}] = 3$.

Remark. Lemma 1.20 implies that we could equivalently define the index as the number of distinct right cosets.

1.9 Lagrange's theorem

Lagrange's theorem is a straightforward but fundamental and extremely useful result in finite group theory. You saw it in *Introduction to Group Theory*, but it is so important that we will prove it again here.

Theorem 1.22 (Lagrange's theorem)**.** *Let G be a finite group and let $H \leqslant G$. Then*

$$|G| = [G : H]\,|H|.$$

In particular, both $[G : H]$ and $|H|$ divide $|G|$.

Proof. Lemma 1.21 implies that $G = \bigcup_{x \in G} xH$ is a disjoint union of $[G : H]$ left cosets, each of which has size $|H|$. □

You saw a number of applications of this theorem in *Introduction to Group Theory*. Here we present some simple applications for use later in the course.

Corollary 1.23. *Let G be a finite group and let $x \in G$. Then $|x|$ divides $|G|$. In particular, $x^{|G|} = e$.*

Proof. Lemma 1.17 implies that $|x| = |\langle x \rangle|$, and Lagrange's theorem implies that $|\langle x \rangle|$ divides $|G|$. □

Corollary 1.24. *Let p be a prime, and let G be a group order p. Then $G \cong \mathbb{Z}_p$, and every non-identity element of G is a generator. In particular, G has no non-trivial proper subgroups.*

Proof. Let $x \in G$ be a non-identity element. The Lagrange's theorem implies that $|\langle x \rangle|$ divides $|G|$, and the fact that $x \neq e$ implies that $|\langle x \rangle| \neq 1$, so $|\langle x \rangle| = |G|$. This forces $\langle x \rangle = G$, as claimed. The fact that $G \cong \mathbb{Z}_p$ then follows from Proposition 1.18. □

Corollary 1.25. *Let G be a group, and let p be a prime, and let $P, Q \leqslant G$ with $|P| = |Q| = p$. Then either $P = Q$ or $P \cap Q = \{e\}$.*

Proof. If $P \cap Q \neq \{e\}$ then there exists $x \in P \cap Q$ with $g \neq e$. This x generates both P and Q by Corollary 1.24, and so $P = Q$. □

Chapter 2

Homomorphisms and quotients

2.1 Homomorphisms

Definition (homomorphism)**.** Let G, H be groups. Then a map $\varphi : G \to H$ is said to be a *homomorphism* if $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in G$.

Remark. An isomorphism is the same as a bijective homomorphism.

Examples 2.1.

(1) For any groups G, H, the map $\varphi : G \to H$ with $\varphi(g) = e$ for every $g \in G$ is a homomorphism, called a *trivial* homomorphism.

(2) The identity map $\iota_G : G \to G$ is a homomorphism (indeed, an isomorphism). More generally, if $H \leqslant G$ then the *inclusion map* $H \hookrightarrow G$, defined by $h \mapsto h$, is a homomorphism.

(3) Let $G_1 \times \cdots \times G_k$ be a direct product. Then each map

$$\begin{array}{rccc} \pi_i & : & G_1 \times \cdots \times G_k & \to & G_i \\ & & (g_1, \ldots, g_k) & \mapsto & g_i \end{array}$$

is a homomorphism, called the *projection* to G_i.

(4) The determinant map $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$, $A \mapsto \det(A)$, is a homomorphism.

(5) Every linear map $\alpha : \mathbb{R}^m \to \mathbb{R}^n$ is a homomorphism.

(6) If $\varphi : G \to H$ is a homomorphism and $K \leqslant G$ then the restriction $\varphi|_K : K \to H$ is also a homomorphism.

Definition. The *image* of a homomorphism $\varphi : G \to H$ is defined to be the set

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

The *kernel* of φ is the set

$$\ker \varphi = \{g \in G : \varphi(g) = e\}.$$

It is worth knowing that a surjective homomorphism is often called an *epimorphism*, an injective homomorphism is often called a *monomorphism*, and a homomorphism $G \to G$ is often called an *endomorphism* of G, but we will not use these terms in this course or in the exam.

Lemma 2.2 (basic properties of homomorphisms)**.** *Let G, H be groups, and let $\varphi : G \to H$ be a homomorphism. Then*

(1) $\varphi(e) = e$;

(2) $\ker \varphi$ is a subgroup of G;

(3) $\varphi(G)$ is a subgroup of H;

(4) φ is injective if and only if $\ker \varphi = \{e\}$;

(5) $\varphi(x^{-1}) = \varphi(x)^{-1}$ for every $x \in G$; and

(6) $\varphi(x_1^{\epsilon_1} \cdots x_n^{\epsilon_n}) = \varphi(x_1)^{\epsilon_1} \cdots \varphi(x_n)^{\epsilon_n}$ for all $x_1, \ldots, x_n \in G$ and $\epsilon_1, \ldots, \epsilon_n \in \{1, -1\}$.

Proof. All of these statements were proved in *Introduction to Group Theory* except (6), which follows from (5) and the definition of a homomorphism by a trivial induction. □

Corollary 2.3. *Let G, H be groups, and let $\varphi : G \to H$ be a homomorphism. Suppose $K \leqslant G$ is a subgroup and $K \cap \ker \varphi = \{e\}$. Then $K \cong \varphi(K)$.*

Proof. It follows from Lemma 2.2 (4) that the restriction of φ to K is an injective homomorphism, and hence a bijective homomorphism—a.k.a. isomorphism—from K to $\varphi(K)$. □

Corollary 2.4. *Let G, H be groups, let $\varphi : G \to H$ be a homomorphism, and let $g \in G$ be an element of finite order. Then $|\varphi(g)|$ divides $|g|$.*

Proof. Lemma 2.2 (1) and (6) show that $\varphi(g)^{|g|} = e$, and so Lemma 1.3 (1) shows that $|\varphi(g)|$ divides $|g|$. □

Lemma 2.5. *Let G, H, K be groups, and suppose that $\varphi : G \to H$ and $\psi : H \to K$ are homomorphisms. Then $\psi \circ \varphi : G \to K$ is also a homomorphism. If φ and ψ are both injective, surjective or bijective then $\psi \circ \varphi$ is also injective, surjective or bijective, respectively.*

Proof. Trivial. □

Lemma 2.6. *Let G and H be groups and let $\varphi : G \to H$ be a homomorphism. Suppose $X \subseteq G$. Then $\varphi(\langle X \rangle) = \langle \varphi(X) \rangle$. In particular, the image of a cyclic group under a homomorphism is cyclic.*

Proof. We have

$$\begin{aligned}
\langle \varphi(X) \rangle &= \{y_1^{\epsilon_1} \cdots y_m^{\epsilon_m} : y_i \in \varphi(X),\, \epsilon_i \in \{1, -1\},\, m \in \mathbb{N}_0\} \\
&= \{\varphi(x_1)^{\epsilon_1} \cdots \varphi(x_m)^{\epsilon_m} : x_i \in X,\, \epsilon_i \in \{1, -1\},\, m \in \mathbb{N}_0\} \\
&= \{\varphi(x_1^{\epsilon_1} \cdots x_m^{\epsilon_m}) : x_i \in X,\, \epsilon_i \in \{1, -1\},\, m \in \mathbb{N}_0\} \qquad \text{(by Lemma 2.2 (6))} \\
&= \varphi(\langle X \rangle).
\end{aligned}$$

□

Lemma 2.7. *Let G and H be groups and let X be a generating set for G. Suppose $\varphi, \psi : G \to H$ are homomorphisms such that $\varphi(x) = \psi(x)$ for every $x \in X$. Then $\varphi = \psi$.*

Proof. This follows from Lemma 2.2 (6). □

2.2 Normal subgroups and quotient groups

It is natural to ask whether the converse to Lemma 2.2 (2) holds: is every subgroup the kernel of some homomorphism? To answer this question, let us first note that the kernel of a homomorphism from a group G has a rather special property, namely that it is closed under conjugation by any element of G, as follows.

Lemma 2.8. *Let $\varphi : G \to H$ be a homomorphism. Then $g(\ker \varphi)g^{-1} \subseteq \ker \varphi$ for every $g \in G$.*

Proof. Given $k \in \ker \varphi$ we have $\varphi(gkg^{-1}) = \varphi(g)\varphi(k)\varphi(g)^{-1} = \varphi(g)e\varphi(g)^{-1} = \varphi(g)\varphi(g)^{-1} = e$, so $gkg^{-1} \in \ker \varphi$ as required. □

Thus, to find an example of a subgroup that is not a kernel of a homomorphism, it suffices to find a subgroup that is not closed under conjugation. For example, there is no homomorphism of D_6 with kernel $\langle s \rangle$, since $rsr^{-1} = r^2 s \neq \langle s \rangle$.

It turns out that this is the only obstruction to a subgroup being the kernel of a homomorphism. In fact, we have the following result.

Theorem 2.9. *Let G be a group and let $H \leqslant G$. Then H is the kernel of some homomorphism from G to another group if and only if $gHg^{-1} \subseteq H$ for every $g \in G$.*

We will prove this theorem shortly. Motivated by it, we introduce the following definition.

Definition (normal subgroup)**.** Let G be a group. Then a subgroup $H \leqslant G$ is said to be a *normal subgroup* of G, denoted $H \trianglelefteq G$, if $gHg^{-1} \subseteq H$ for all $g \in G$.

Remark. Every subgroup of an abelian group is normal, so in the special case of abelian groups it really is the case that every subgroup is the kernel of some homomorphism.

It turns out that, given a normal subgroup, we can directly construct a homomorphism of which it is the kernel, as follows.

Theorem 2.10. *Let G be a group and $H \trianglelefteq G$. Then G/H is a group under set multiplication with identity element $H = eH$, and*

$$(aH)(bH) = (ab)H \tag{2.1}$$

for every $a, b \in G$. Moreover, the map

$$\begin{array}{rccc} \pi & : & G & \to & G/H \\ & & g & \mapsto & gH \end{array} \tag{2.2}$$

is a surjective homomorphism with kernel H.

Definition (quotient group; quotient homomorphism)**.** The group G/H with the operation of set multiplication is called the *quotient group* of G by H. The homomorphism π defined in (2.2) is called the *quotient homomorphism* from G to G/H.

In order to prove Theorem 2.10, it will be useful to note that if $H \trianglelefteqslant G$ then not only is gHg^{-1} *contained* in H for all $g \in G$, it is *equal* to H, as follows.

Lemma 2.11. *Let G be a group and let $H \trianglelefteqslant G$. Then $gHg^{-1} = H$ for all $g \in G$. In particular, $gH = Hg$ for all $g \in G$.*

Proof. For every $g \in G$ we have $(g^{-1})H(g^{-1})^{-1} \subseteq H$, i.e. $g^{-1}Hg \subseteq H$, and hence $H \subseteq gHg^{-1}$. □

Thus, if $H \trianglelefteqslant G$ then the left and right cosets of H coincide, and we may refer simply to 'cosets', without specifying left or right.

Warning. In general, if $gHg^{-1} \subseteq H$ for a given element g, this does not necessarily imply that $g^{-1}Hg = H$; see e.g. Exercise Sheet 2 Q17.

Warning. Note that $gH = Hg$ denotes an equality of *sets*; a normal subgroup H does *not* have to satisfy $gh = hg$ for all $g \in G$ and $h \in H$.

Proof of Theorem 2.10. Set multplication is associative by Lemma 1.4. For every $a, b \in G$ we have

$$\begin{aligned} aHbH &= abHH && \text{(by Lemma 2.11)} \\ &= abH && \text{(by Lemma 1.10).} \end{aligned}$$

This implies that G/H is closed under set multiplication, and that set multiplication is characterised by (2.1) as claimed. It implies that $eH = H$ is an identity element for set multiplication on G/H and that $g^{-1}H$ is an inverse for gH, so that G/H is a group under set multiplication. It also implies that π is a homomorphism. The homomorphism π is trivially surjective: for every coset gH of H we have $gH = \pi(g)$. Finally, $g \in \ker \pi \iff gH = H \iff g \in H$ by Lemma 1.21 (2). □

Theorem 2.9 is now immediate from Lemma 2.2 (2) and Lemma 2.8 and Theorem 2.10.

Lemma 2.12. *Let G be a group and suppose that $H < G$ has index 2. Then $H \trianglelefteqslant G$ and $G/H \cong \mathbb{Z}_2$.*

Proof. Let $g \in G$. If $g \in H$ then $gH = Hg = H$, so that $gHg^{-1} = H$. If $g \notin H$ then Lemma 1.21 and Exercise Sheet 1 Q6 imply that $gH = G \setminus H = Hg$, and hence that $gHg^{-1} = H$. It follows that $H \trianglelefteqslant G$ as claimed. Since $|G/H| = 2$, Corollary 1.24 then implies that $G/H \cong \mathbb{Z}_2$. □

Example. Lemma 2.12 shows that the subgroup $C_n = \langle r \rangle \leqslant D_{2n}$ is normal, and that $D_{2n}/C_n \cong \mathbb{Z}_2$. Note that the two elements of D_{2n}/C_n are the set C_n of rotations and the set sC_n of reflections. Think about how this relates to Exercise Sheet 1 Q1.

We will now see an immediate application of quotient groups, in the proof of the following special case of a famous result called Cauchy's theorem.

Theorem 2.13 (Cauchy's theorem for abelian groups)**.** *Let p be a prime, let G be a finite abelian group, and suppose that p divides $|G|$. Then G contains an element of order p.*

Proof. By induction we may assume that the theorem holds for all groups of order less than $|G|$. Let $g \neq e$ be an element of G. If p divides $|g|$ then $g^{|g|/p}$ is the required element of order p. If not then p divides $[G : \langle g \rangle]$ by Proposition 1.18 and Lagrange's theorem, and so $G/\langle g \rangle$ contains an element $h\langle g \rangle$ of order p by induction. In that case, $(h\langle g \rangle)^{|h|} = h^{|h|}\langle g \rangle = e\langle g \rangle$, so p divides $|h|$ by Lemma 1.3 and $h^{|h|/p}$ has order p. □

In fact, Cauchy's theorem also holds in non-abelian groups, but we will not prove that until we have seen *group actions* later in the course.

2.3 The homomorphism theorem

It turns out that every homomorphism is essentially a quotient homomorphism, in the following sense.

Theorem 2.14 (homomorphism theorem)**.** *Let G, H be groups. Suppose that $\varphi : G \to H$ is a homomorphism, and let $\pi : G \to G/\ker\varphi$ be the quotient homomorphism. Then there is an isomorphism $\psi : G/\ker\varphi \to \operatorname{im}\varphi$ such that $\varphi = \psi \circ \pi$. In particular, $\operatorname{im}\varphi \cong G/\ker\varphi$.*

Proof. Set $K = \ker\varphi$. Note that for $a, b \in G$ we have

$$\begin{aligned} aK = bK &\iff b^{-1}a \in K \\ &\iff \varphi(b^{-1}a) = e \\ &\iff \varphi(a) = \varphi(b), \end{aligned}$$

and so the map

$$\begin{array}{rccc} \psi & : & G/K & \to & \operatorname{im}\varphi \\ & & gK & \mapsto & \varphi(g) \end{array}$$

is well defined and injective. It is also trivially surjective, and satisfies $\varphi = \psi \circ \pi$ by definition. To see that ψ is a homomorphism, note that

$$\begin{aligned} \psi(aKbK) &= \psi(abK) \\ &= \varphi(ab) \\ &= \varphi(a)\varphi(b) \\ &= \psi(aK)\psi(bK). \end{aligned}$$

□

Examples.

(1) The trivial homomorphism $\varphi : G \to \{e\}$ has kernel G and image $\{e\}$, so $G/G \cong \{e\}$.

(2) The identity homomorphism $\iota_G : G \to G$ has kernel $\{e\}$ and image G, so $G/\{e\} \cong G$.

(3) Define $\mathrm{SL}_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) : \det(A) = 1\}$. The determinant map $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$, $A \mapsto \det(A)$, is a surjective homomorphism with kernel $\mathrm{SL}_n(\mathbb{R})$, so $\mathrm{SL}_n(\mathbb{R}) \trianglelefteq \mathrm{GL}_n(\mathbb{R})$ and $\mathrm{GL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{R}) \cong \mathbb{R}^*$.

2.4 The isomorphism theorems

In this section and the next we present three fundamental theorems detailing the relationship between various subgroups and quotients of a group. Here we call them the *first and second isomorphism theorems* and the *correspondence theorem*. **Warning.** The naming and numbering of these theorems is not consistent across the literature!

Theorem 2.15 (first isomorphism theorem)**.** *Let G be a group. Suppose $H \leqslant G$ and $N \trianglelefteq G$, and write $\pi : G \to G/N$ for the quotient homomorphism. Then $H \cap N \trianglelefteq H$ and $\pi(H) \cong H/(H \cap N)$.*

Proof. Let $\pi|_H : H \to G/N$ be restriction to H of π, recalling that this is a homomorphism. We then have $\ker \pi|_H = H \cap N$, so $H \cap N \trianglelefteq H$ by Theorem 2.9, and $\operatorname{im} \pi|_H = \pi(H)$, so $\pi(H) \cong H/(H \cap N)$ by the homomorphism theorem (Theorem 2.14). □

When interpreting the first isomorphism theorem, it is often useful to have the following alternative way of expressing $\pi(H)$.

Lemma 2.16. *Let G be a group. Suppose $H \leqslant G$ and $N \trianglelefteq G$, and write $\pi : G \to G/N$ for the quotient homomorphism. Then $HN \leqslant G$ and $\pi(H) = HN/N$.*

Proof. The normality of N implies that $hN = Nh$ for all $h \in H$, and in particular that $HN = NH$, so $HN \leqslant G$ by Lemma 1.14. The fact that $nN = N$ for every $n \in N$ implies that $HN/N = \{hnN : h \in H, n \in N\} = \{hN : h \in H\} = \pi(H)$. □

Theorem 2.17 (second isomorphism theorem)**.** *Let G be a group, and let $N \leqslant H \leqslant G$ with $N, H \trianglelefteq G$. Then $H/N \trianglelefteq G/N$ and $(G/N)/(H/N) \cong G/H$.*

Proof. Let $\varphi : G/N \to G/H$, $gN \mapsto gH$. We have

$$aN = bN \implies ab^{-1} \in N \subseteq H \implies aH = bH,$$

so φ is well defined. It is a homomorphism because

$$\varphi(aNbN) = \varphi(abN) = abH = aHbH = \varphi(aN)\varphi(bN),$$

and it is trivially surjective. Since

$$\ker \varphi = \{gN : gH = eH\} = \{gN : g \in H\} = H/N,$$

we therefore have $H/N \trianglelefteq G/N$ by Theorem 2.9 and $(G/N)/(H/N) \cong G/H$ by the homomorphism theorem. □

2.5 The correspondence theorem

In this section we investigate various relationships between the subgroups of a group and the subgroups of a quotient/homomorphic image of that group. We start by showing that the property of being a subgroup or normal subgroup is preserved under images and pullbacks of homomorphisms.

Lemma 2.18. *Let G, Γ be groups and let $\varphi : G \to \Gamma$ be a homomorphism. Then the following hold.*

(1) (a) If $H \leqslant G$ then $\varphi(H) \leqslant \Gamma$.
(b) If $H \trianglelefteq G$ then $\varphi(H) \trianglelefteq \varphi(G)$.

(2) (a) If $K \leqslant \Gamma$ then $\varphi^{-1}(K) \leqslant G$.
(b) If $K \trianglelefteq \Gamma$ then $\varphi^{-1}(K) \trianglelefteq G$.

Proof.

(1) (a) This is immediate from Lemma 2.2 (3).
(b) If $H \trianglelefteq G$ then for every $g \in G$ and $h \in H$ we have $\varphi(g)\varphi(h)\varphi(g)^{-1} = \varphi(ghg^{-1}) \in \varphi(H)$.

(2) (a) We have $e \in \varphi^{-1}(K)$, so $\varphi^{-1}(K) \neq \varnothing$. Moreover, if $a, b \in \varphi^{-1}(K)$ then $\varphi(a), \varphi(b) \in K$, so $\varphi(ab^{-1}) = \varphi(a)\varphi(b)^{-1} \in K$ and thus $ab^{-1} \in \varphi^{-1}(K)$, so $\varphi^{-1}(K) \leqslant G$ by the subgroup test.
(b) Given $g \in G$, if $K \trianglelefteq \Gamma$ then $\varphi(g\,\varphi^{-1}(K)\,g^{-1}) = \varphi(g)K\varphi(g)^{-1} = K$, and so $g\,\varphi^{-1}(K)\,g^{-1} \subseteq \varphi^{-1}(K)$. □

Warning. Note that $H \trianglelefteq G$ does not imply that $\varphi(H) \trianglelefteq \Gamma$; for example, let H be a non-normal subgroup of Γ, and take $G = H$ and $\varphi : G \to \Gamma$ defined by $\varphi(h) = h$.

We now prove an important correspondence between the subgroups of a quotient group G/N and the subgroups of G containing N.

Theorem 2.19 (correspondence theorem)**.** *Suppose G is a group and $N \trianglelefteq G$. Then we may define a bijection between the set of subgroups of G containing N and the set of subgroups of G/N by $H \mapsto H/N$. We may also define a bijection between the set of normal subgroups of G containing N and the set of normal subgroups of G/N by $H \mapsto H/N$.*

Proof. Write $\pi : G \to G/N$ the quotient homomorphism, and note that if $N \leqslant H \leqslant G$ then $H/N = \pi(H)$ by Lemma 2.16. It therefore follows from Lemma 2.18 (1) that $H/N \leqslant G/N$, normal if $H \trianglelefteq G$. This shows that we can indeed define maps of the claimed form. It remains to show that they are bijective.

To see that these maps are injective, suppose H and H' are two subgroups of G containing N such that $H/N = H'/N$. Then given an arbitrary $h \in H$, there exists $h' \in H'$ such that $hN = h'N$, and hence $h \in h'N \subseteq H'$; thus $H \subseteq H'$. By symmetry, $H' \subseteq H$ also, so that $H = H'$ as required.

To see that the maps are surjective, suppose that $K \leqslant G/N$ and note that $K = \pi(\pi^{-1}(K))$. Lemma 2.18 (2) implies that $\pi^{-1}(K) \leqslant G$, and that $\pi^{-1}(K) \trianglelefteq G$ if $K \trianglelefteq G/N$. Moreover, $\pi^{-1}(K)$ contains N because $eN \in K$, and Lemma 2.16 then implies that $K = \pi^{-1}(K)/N$. □

2.6 Simple groups

In any group G the subgroups G and $\{e\}$ are always normal.

Definition (simple group)**.** A non-trivial group G is said to be *simple* if G and $\{e\}$ are the *only* normal subgroups of G.

Finite simple groups play a similar role in the theory of finite groups to that played by prime numbers in number theory. Indeed, just as every natural number has a unique factorisation into prime numbers, it turns out that every finite group has an essentially unique 'factorisation' into simple groups.

The precise notion we use to obtain a 'factorisation' of a group into simple groups is the following.

Definition (composition series)**.** Let G be a group. A *composition series* of G is a sequence $\{e\} = N_0 \trianglelefteqslant N_1 \trianglelefteqslant \cdots \trianglelefteqslant N_k = G$ of subgroups with the property that each quotient N_i/N_{i-1} is simple. The simple quotients N_i/N_{i-1} are called the *composition factors* of the series. **Warning.** The subgroup N_i is only required to be normal in N_{i+1}, not necessarily in G.

It is not hard to check that every finite group admits at least one compostion series, and you are invited to do so in the exercise sheet. This is analogous to the fact that every natural number has a prime factorisation. The analogue of the uniqueness of prime factorisations is the following famous result.

Theorem 2.20 (Jordan–Hölder theorem – not examinable)**.** *Let G be a group and suppose that*

$$\{e\} = G_0 \trianglelefteqslant G_1 \trianglelefteqslant \cdots \trianglelefteqslant G_m = G \tag{2.3}$$

and

$$\{e\} = H_0 \trianglelefteqslant H_1 \trianglelefteqslant \cdots \trianglelefteqslant H_n = G \tag{2.4}$$

are two composition series for G. Then the factors of (2.3) *and* (2.4) *are the same up to isomorphism and reordering (that is to say, $m = n$ and there exists a permutation $\sigma \in S_n$ such that $G_i/G_{i-1} \cong H_{\sigma(i)}/H_{\sigma(i)-1}$ for all $i \in [n]$).*

Important though this theorem is, we mention it here only to motivate the study of simple groups, and it is not examinable. Interested readers can find a proof in two non-examinable sections at the end of this chapter.

We have seen that every finite group is 'built' form a unique collection of simple groups, namely its composition factors. This observation provides us with a two-step strategy to study all finite groups. We need to:

(1) classify the finite simple groups; and

(2) understand how groups can be 'fitted together' to build larger groups.

The classification of abelian simple groups is straightforward.

Theorem 2.21 (classification of abelian simple groups)**.** *Let G be an abelian group. Then G is simple if and only if $G \cong \mathbb{Z}_p$ for a prime p.*

Proof. If $G \cong \mathbb{Z}_p$ for a prime p then we saw in Corollary 1.24 that G has no non-trivial proper subgroups, and hence is simple.

Conversely, suppose G is simple, and let $x \in G$ with $x \neq e$. Then $\langle x \rangle \trianglelefteqslant G$ since G is abelian, so $G = \langle x \rangle$ and G is cyclic. If $|x|$ is infinite then $\{e\} < \langle x^2 \rangle < G$, contradicting simplicity, so G is finite. If $|x| = mn$ with $m, n > 1$ then $\{e\} < \langle x^m \rangle < G$, again contradicting simplicity, so $|x|$ is prime. Proposition 1.18 therefore shows that $G \cong \mathbb{Z}_p$. □

The classification of all finite simple groups is much harder, and was arguably one of the great achievements of 20th-century mathematics. The proof was completed in the 1980s (although one or two gaps have subsequently emerged and been corrected), and is spread over thousands of pages in hundreds of published papers. Towards the end of this course, we will take some small steps in the direction of this theorem.

The second step is the so-called *extension problem*. This is a very difficult problem that has been investigated since the early days of group theory in the 19th century. There has been a focus on understanding extensions that satisfy some additional conditions, but a solution to the general problem remains out of reach.

2.7 Inner direct products

In this section we will see a useful application of normal subgroups: some simple criteria for showing that a given group is isomorphic to a direct product.

Definition (inner direct product)**.** A group G is said to be an *inner direct product* of subgroups $H_1, \dots, H_n \leqslant G$, which we will write as $G = H_1 \otimes \cdots \otimes H_n$, if $H_i \trianglelefteq G$ for each i, and for each $g \in G$ there exist unique elements $h_i \in H_i$ such that $g = h_1 \cdots h_n$.

Warning. The notation $G = H_1 \otimes \cdots \otimes H_n$ is not standard, although I will – and you may – use it freely in this course.

Theorem 2.22. *Let $G, G_1, \dots, G_n$ be groups. Then $G \cong G_1 \times \cdots \times G_n$ if and only if there exist normal subgroups $H_1, \dots, H_n \trianglelefteq G$ satisfying $H_i \cong G_i$ such that $G = H_1 \otimes \cdots \otimes H_n$.*

Theorem 2.22 can be a very convenient means by which to show that a given group is isomorphic to a direct product. For example, in Exercise Sheet 1 you showed that $\mathbb{R}^* \cong \mathbb{R} \times \mathbb{Z}_2$ by defining an explicit isomorphism between the two groups. An alternative route is now open to us, namely to observe that $\mathbb{R}^* = \mathbb{R}_{>0} \otimes \{\pm 1\}$, and to recall that $\mathbb{R}_{>0} \cong \mathbb{R}$ (see Examples 1.5) and note that $(\{\pm 1\}, \times) \cong \mathbb{Z}_2$ (either directly or using Corollary 1.24).

One direction of Theorem 2.22 is easy, as follows.

Proposition 2.23. *Let $G = G_1 \times \cdots \times G_n$ be a direct product and set*

$$H_i = \{(\underbrace{e, \dots, e}_{i-1}, a_i, \underbrace{e, \dots, e}_{n-i}) \in G : a_i \in G_i\}$$

for $i = 1, \dots, n$. Then $H_i \trianglelefteq G$ and $H_i \cong G_i$ for each i, and for each $g \in G$ there exist unique elements $h_i \in H_i$ such that $g = h_1 \cdots h_n$.

Proof. For each i the map $\psi_i : G \to G$, $(a_1, \dots, a_n) \mapsto (a_1, \dots, a_{i-1}, e, a_{i+1}, \dots, a_n)$ is a homomorphism with kernel H_i, so $H_i \trianglelefteq G$, and the map $\varphi_i : H_i \to G_i$, $(e, \dots, e, a_i, e, \dots, e) \mapsto a_i$ is an isomorphism, so $H_i \cong G_i$. Finally, given elements $h_i = (e, \dots, e, a_i, e, \dots, e) \in H_i$ for each i we have $h_1 \cdots h_n = (a_1, \dots, a_n)$, so that given $g = (g_1, \dots, g_n) \in G$ we have $g = h_1 \cdots h_n$ if and only if $h_i = (e, \dots, e, g_i, e, \dots, e)$ for all i. □

We spend the rest of this section proving the harder, and more useful, converse direction of Theorem 2.22. We start by providing another condition that is equivalent to being an inner direct product; indeed, many authors use this condition as the definition of an inner direct product.

Proposition 2.24. *Let G be a group and suppose that $H_1, \dots, H_n \trianglelefteq G$. Then the following conditions are equivalent:*

(1) $G = H_1 \otimes \cdots \otimes H_n$;

(2) $G = H_1 \cdots H_n$, and $H_i \cap (H_1 \cdots H_{i-1} H_{i+1} \cdots H_n) = \{e\}$ for all $i = 1, \dots, n$.

Lemma 2.25. *Suppose that G is a group with normal subgroups $H_1, \dots, H_k \trianglelefteq G$ satisfying condition (2) of Proposition 2.24. Then whenever $i \neq j$, the elements of H_i commute with the elements of H_j.*

Proof. Let $i \neq j$. Given $h_i \in H_i$ and $h_j \in H_j$ we have $h_i h_j = h_i h_j h_i^{-1} h_j^{-1} h_j h_i$, so we need to show that $h_i h_j h_i^{-1} h_j^{-1} = e$. To see this, note that

$$\begin{aligned} h_i h_j h_i^{-1} h_j^{-1} &\in H_i (h_j H_i h_j^{-1}) = H_i, \\ h_i h_j h_i^{-1} h_j^{-1} &\in (h_i H_j h_i^{-1}) H_j = H_j, \end{aligned}$$

because $H_i, H_j \trianglelefteq G$, so that $h_i h_j h_i^{-1} h_j^{-1} \in H_i \cap H_j$. However, condition (2) of Proposition 2.24 implies in particular that $H_i \cap H_j = \{e\}$. □

Proof of Proposition 2.24. If $G = H_1 \otimes \cdots \otimes H_n$ then we certainly have $G = H_1 \cdots H_n$. Moreover, note that for each i, if $h_i = h_1 \cdots h_{i-1} h_{i+1} \cdots h_n$ with each $h_j \in H_j$ then

$$\underbrace{e \cdots e}_{i-1} h_i \underbrace{e \cdots e}_{n-i} = h_1 \cdots h_{i-1} e h_{i+1} \cdots h_n,$$

so that each $h_j = e$ by uniqueness.

Conversely, if (2) holds then by definition there exist, for each $g \in G$, elements $h_i \in H_i$ such that $g = h_1 \cdots h_n$. To see that these choices of h_i are unique, note that if in addition $g = k_1 \cdots k_n$ with each $k_i \in H_i$ then $e = k_1 \cdots k_n h_n^{-1} \cdots h_1^{-1}$. Lemma 2.25 then implies that for each i we have

$$k_i^{-1} h_i = k_1 h_1^{-1} \cdots k_{i-1} h_{i-1}^{-1} k_{i+1} h_{i+1}^{-1} \cdots k_n h_n^{-1} \in H_i \cap (H_1 \cdots H_{i-1} H_{i+1} \cdots H_n) = \{e\},$$

and hence $h_i = k_i$. □

Proof of Theorem 2.22. One direction follows from Proposition 2.23. To prove the other direction, suppose that G be a group with normal subgroups $H_1, \dots, H_k \trianglelefteq G$ such that $G = H_1 \otimes \cdots \otimes H_n$. It is easy to check using Proposition 2.24 and Lemma 2.25 that

$$\begin{array}{rccc} \varphi & : & H_1 \times \cdots \times H_n & \to & G \\ & & (h_1, \dots, h_n) & \mapsto & h_1 \cdots h_n \end{array}$$

is a homomorphism, and the definition of an inner direct product implies that it is bijective. □

2.A Proof of the Jordan–Hölder theorem (not examinable)

Definition (subnormal series; normal series)**.** Let G be a group. A *subnormal series* of G is a finite sequence

$$\{e\} = G_0 \trianglelefteqslant G_1 \trianglelefteqslant \cdots \trianglelefteqslant G_n = G \tag{2.5}$$

of subgroups of G, such that $G_i \trianglelefteqslant G_{i+1}$ for all i. The quotient groups $G_1/G_0, G_2/G_1, \ldots, G_n/G_{n-1}$ are called the *factors* of the series, and n is the *length* of the series. If each G_i is normal in G, rather than just in G_{i+1}, then the series is called a *normal series*.

Let G be a group with a subnormal series. Any subnormal series for G that results from inserting additional subgroups into the original series is called a *refinement* of that series. More formally, we make the following definition.

Definition (refinement of a subnormal series)**.** Let G be a group with a subnormal series

$$\{e\} = G_0 \trianglelefteqslant G_1 \trianglelefteqslant \cdots \trianglelefteqslant G_n = G. \tag{2.6}$$

Then we define the *refinements* of S recursively by the following rules.

- The subnormal series (2.6) is a refinement of itself.
- If $\{e\} = H_0 \trianglelefteqslant H_1 \trianglelefteqslant \cdots \trianglelefteqslant H_m = G$ is a refinement of S and $K \leqslant G$ satisfies $H_{i-1} \leqslant K \trianglelefteqslant H_i$ for some $i \in [m]$ then $\{e\} = H_0 \trianglelefteqslant \cdots \trianglelefteqslant H_{i-1} \trianglelefteqslant K \trianglelefteqslant H_i \trianglelefteqslant \cdots \trianglelefteqslant H_m = G$ is also a refinement of (2.6).

Remark. One can refine any subnormal series by repeating some of the groups. In the case of a composition series, the correspondence theorem implies that this is the only kind of refinement possible.

The Jordan–Hölder theorem is then a consequence of the following more general result.

Theorem 2.26 (refinement theorem)**.** *Let G be a group and suppose that*

$$\{e\} = G_0 \trianglelefteqslant G_1 \trianglelefteqslant \cdots \trianglelefteqslant G_m = G \tag{2.7}$$

and

$$\{e\} = H_0 \trianglelefteqslant H_1 \trianglelefteqslant \cdots \trianglelefteqslant H_n = G \tag{2.8}$$

are subnormal series for G. Then we may refine these series by inserting additional subgroups

$$G_{i-1} = G_{i,0} \trianglelefteqslant G_{i,1} \trianglelefteqslant \cdots \trianglelefteqslant G_{i,n} = G_i \qquad (i = 1, \ldots, m) \tag{2.9}$$

and

$$H_{i-1} = H_{i,0} \trianglelefteqslant H_{i,1} \trianglelefteqslant \cdots \trianglelefteqslant H_{i,m} = H_i \qquad (i = 1, \ldots, n) \tag{2.10}$$

in such a way that

$$G_{i,j}/G_{i,j-1} \cong H_{j,i}/H_{j,i-1}$$

for all $i \in [m]$, $j \in [n]$. In particular, (2.9) *and* (2.10) *have the same factors up to isomorphism and reordering.*

We prove the refinement theorem in the next section. We close this section by seeing how it implies the Jordan–Hölder theorem.

Proof of Theorem 2.20. Let $G_{i,j}$ and $H_{j,i}$ be the groups given by Theorem 2.26. Since each factor G_i/G_{i-1} is simple, the correspondence theorem implies that for each i there exists a unique j such that $G_{i,j}/G_{i,j-1} \cong G_i/G_{i-1}$, with $G_{i,j}/G_{i,j-1} \cong \{e\}$ for all other j. Similarly, for each j there exists a unique i such that $H_{j,i}/H_{j,i-1} \cong H_j/H_{j-1}$, with $H_{j,i}/H_{j,i-1} \cong \{e\}$ for all other i. Thus, up to isomorphism and reordering, the non-trivial factors of (2.9) are precisely the factors of (2.3), and the non-trivial factors of (2.10) are precisely the factors of (2.4). The desired conclusion therefore follows from the refinement theorem (Theorem 2.26). □

2.B Proof of the refinement theorem (not examinable)

We prove the refinement theorem using an argument due to Baumslag. We start with a lemma that extends the first isomorphism theorem (which is the case $L = \{e\}$).

Lemma 2.27. *Let G be a group and suppose that $H, N, L \leqslant G$ are subgroups such that*

(a) $hN = Nh$ for all $h \in H$, and

(b) $L \trianglelefteq H$.

Then

(1) $HN \leqslant G$;

(2) $LN \trianglelefteq HN$;

(3) $L(H \cap N) \trianglelefteq H$;

(4) $HN/LN \cong H/L(H \cap N)$.

Proof. Condition (a) implies in particular that $HN = NH$, so Lemma 1.14 implies that $HN \leqslant G$, giving (1). Since $L \subseteq H$ by (b), we similarly conclude that $LN \leqslant G$. Given $h \in H$ and $n \in N$, using (a) and (b) repeatedly we obtain

$$hnLN = hnNL = hNL = hLN = hLNn = LhNn = LNhn,$$

so (2) holds. We may therefore define $\pi : H \to HN/LN$ to be the restriction to H of the quotient homomorphism $HN \to HN/LN$, noting that

$$\pi(H) = HN/LN. \tag{2.11}$$

We claim that

$$H \cap LN = L(H \cap N). \tag{2.12}$$

Since $H = LH$ by (b), we conclude that $H \cap LN = LH \cap LN \supseteq L(H \cap N)$. Conversely, if $x \in H \cap LN$ then $x = \ell n$ for some $\ell \in L$ and $n \in N$, so $n = \ell^{-1}x \in H$ because $L \subseteq H$. Thus $x = \ell n \in L(H \cap N)$, and (2.12) holds as claimed. But $\ker \pi = H \cap LN$ by definition, so (2.12) implies that $\ker \pi = L(H \cap N)$. By Theorem 2.9 this gives gives (3), and by (2.11) and the homomprhism theorem it gives (4). □

Proof of Theorem 2.26. We will show that we may take $G_{i,j} = (G_i \cap H_j)G_{i-1}$ and $H_{j,i} = (H_j \cap G_i)H_{j-1}$. First, note that $(G_i \cap H_j)$ normalises G_{i-1}, so $G_{i,j}$ is indeed a group by Lemma 1.14. $H_{j,i}$ is similarly a group. Next, note that $(G_i \cap H_{j-1}) \trianglelefteqslant (G_i \cap H_j)$ because $H_{j-1} \trianglelefteqslant H_j$, so applying Lemma 2.27 (2) with $G = G_i$, $N = G_{i-1}$, $H = G_i \cap H_j$ and $L = G_i \cap H_{j-1}$ implies that $(G_i \cap H_{j-1})G_{i-1} \trianglelefteqslant (G_i \cap H_j)G_{i-1}$, which is to say $G_{i,j-1} \trianglelefteqslant G_{i,j}$, so the refinement defined by (2.9) really is a subnormal series. Lemma 2.27 (4) also implies that

$$G_{i,j}/G_{i,j-1} \cong (G_i \cap H_j)/(G_i \cap H_{j-1})(G_{i-1} \cap H_j). \tag{2.13}$$

Interchanging the roles of (2.7) and (2.8), we conclude that each $H_{j,i}$ is also a group and $H_{j,i-1} \trianglelefteqslant H_{j,i}$ with

$$H_{j,i}/H_{j,i-1} \cong (H_j \cap G_i)/(H_j \cap G_{i-1})(H_{j-1} \cap G_i).$$

This is precisely the same group as (2.13), so the proof is complete. □

Chapter 3

Symmetric groups

3.1 The basics

In this chapter we study the *symmetric groups*, which you met in *Introduction to Group Theory* and we briefly mentioned in Examples 1.2. We start by recalling their definition.

Definition (symmetric group)**.** Let X be a non-empty set. A *permutation* of X is a bijection $X \to X$. The set of all permutations of X forms a group under composition, called the *symmetric group on* X and denoted $\mathrm{Sym}(X)$. For $n \in \mathbb{N}$ we abbreviate $\mathrm{Sym}([n])$ by S_n, called the *symmetric group of degree* n. We always denote the identity permutation by e. Given two permutations $f, g \in \mathrm{Sym}(X)$, we abbreviate the composition $f \circ g$ by simply fg.

Proposition 3.1. *Let $n \in \mathbb{N}$, and suppose X and Y are two sets of size n. Then there are exactly $n!$ distinct bijections $X \to Y$. In particular, $|S_n| = n!$.*

Proof. We proceed by induction, the base case $n = 1$ being trivial. For $n \geqslant 2$, let $x \in X$, and note that by induction we have

$$\begin{aligned} \#\text{ bijections } X \to Y &= \sum_{y \in Y} \#\text{ bijections } X \to Y \text{ mapping } x \mapsto y \\ &= \sum_{y \in Y} \#\text{ bijections } X \setminus \{x\} \to Y \setminus \{y\} \\ &= \sum_{y \in Y} (n-1)! \\ &= n! \end{aligned}$$

□

One way of denoting an element $f \in S_n$ is by

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

For example, in S_6 the permutation mapping $1 \mapsto 1$, $2 \mapsto 3$, $3 \mapsto 5$, $4 \mapsto 6$, $5 \mapsto 2$, $6 \mapsto 4$ is denoted

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 5 & 6 & 2 & 4 \end{pmatrix}. \tag{3.1}$$

However, a more compact way of denoting elements of S_n is using *cycles*.

Definition (cycle; transposition)**.** Let $k \in \mathbb{N}$. A permutation $f \in S_n$ is called a *k-cycle*, or a *cycle of length k*, if there are k distinct numbers $i_1, i_2, \ldots, i_k \in [n]$ such that

$$f(i_1) = i_2, \quad f(i_2) = i_3, \quad \ldots \quad f(i_{k-1}) = f(i_k), \quad f(i_k) = i_1,$$

and such that $f(j) = j$ for every $j \in [n] \setminus \{i_1, \ldots, i_k\}$. We denote this f by $f = (i_1, i_2, \ldots, i_k)$. A 2-cycle is also called a *transposition*. Two cycles $(i_1, \ldots, i_k)$ and $(j_1, \ldots, j_\ell)$ are called *disjoint* if $\{i_1, \ldots, i_k\} \cap \{j_1, \ldots, j_\ell\} = \varnothing$.

Note that a k-cycle $(i_1, i_2, \ldots, i_k)$ has order k and satisfies

$$(i_1, i_2, \ldots, i_k) = (i_2, i_3, \ldots, i_k, i_1) = (i_3, i_4, \ldots, i_k, i_1, i_2) = \cdots,$$

$$(i_1, i_2, \ldots, i_k)^{-1} = (i_k, i_{k-1}, \ldots, i_1)$$

and

$$(i_1, \ldots, i_k) = (i_1, i_2)(i_2, i_3) \cdots (i_{k-1}, i_k). \tag{3.2}$$

Proposition 3.2 (from *Introduction to Group Theory*)**.** *Let $n \in \mathbb{N}$. Then every element of S_n can be expressed as a product of disjoint cycles whose lengths sum to n, and this expression is unique up to reordering the cycles. Every element of S_n can also be expressed as a product of transpositions.*

For example, the permutation (3.1) can be denoted more succinctly by

$$(1)(2, 3, 5)(4, 6), \tag{3.3}$$

or by

$$(3, 5)(2, 5)(4, 6).$$

Remarks.

(1) Note that disjoint cycles commute, so the caveat 'up to reordering the cycles' in Proposition 3.2 is unavoidable.

(2) Proposition 3.2 shows that S_n is generated by the set of transpositions. In fact, $S_n = \langle (1, 2), (1, 3), \ldots, (1, n) \rangle$, since $(i, j) = (1, i)(1, j)(1, i)$.

Definition (cycle type)**.** Let $f \in S_n$ be written as a product of disjoint cycles whose lengths sum to n as in Proposition 3.2, and let $\ell_1, \ldots, \ell_k$ be the lengths of these cycles, ordered so that $\ell_1 \geqslant \ell_2 \geqslant \cdots \geqslant \ell_k$. Then the k-tuple $[\ell_1, \ldots, \ell_k]$ is called the *cycle type* of f.

Thus, for example, the cycle type of the permutation (3.3) is $[3, 2, 1]$.

Lemma 3.3 (from *Introduction to Group Theory*)**.** *If $f \in S_n$ has cycle type $[\ell_1, \ldots, \ell_k]$, then $|f| = \operatorname{lcm}(\ell_1, \ldots, \ell_k)$.*

3.2 Conjugacy in S_n

Recall that two elements x, y in a group G are said to be *conjugate* if there exists $g \in G$ such that $y = gxg^{-1}$.

Lemma 3.4. *Let $n \geqslant 2$, and let $i_1, \ldots, i_k$ be distinct elements of $[n]$. Then $g(i_1, \ldots, i_k)g^{-1} = (g(i_1), \ldots, g(i_k))$ for every $g \in S_n$.*

Proof. If $m \in [n] \setminus \{g(i_1), \ldots g(i_k)\}$ then $g^{-1}(m) \notin \{i_1, \ldots, i_k\}$, and so $(i_1, \ldots, i_k)g^{-1}(m) = g^{-1}(m)$ and $g(i_1, \ldots, i_k)g^{-1}(m) = m$, as required. On the other hand, a trivial calculation shows that for each $j = 1, \ldots, k$ we have

$$g(i_1, \ldots, i_k)g^{-1}(g(i_j)) = g(i_{j+1}),$$

where the final index $i + 1$ is taken modulo k. □

Proposition 3.5. *Let $x, y \in S_n$. Then x and y are conjugate if and only if they have the same cycle type.*

Proof. First, suppose $x \in S_n$ has cycle type $[\ell_1, \ldots, \ell_k]$, and let $f_1, \ldots, f_k$ be disjoint cycles, with each f_i an ℓ_i-cycle, such that $x = f_1 \cdots f_k$. Then for every $g \in S_n$ we have $gxg^{-1} = gf_1g^{-1} \cdots gf_kg^{-1}$. Lemma 3.4 implies that the gf_ig^{-1} are disjoint cycles and that each gf_ig^{-1} is an ℓ_i-cycle, and so the conjugate gxg^{-1} also has cycle type $[\ell_1, \ldots, \ell_k]$.

Conversely, suppose x and y both have cycle type $[\ell_1, \ldots, \ell_k]$, say

$$x = (a_1^{(1)}, \ldots, a_{\ell_1}^{(1)}) \cdots (a_1^{(k)}, \ldots, a_{\ell_k}^{(k)})$$
$$y = (b_1^{(1)}, \ldots, b_{\ell_1}^{(1)}) \cdots (b_1^{(k)}, \ldots, b_{\ell_k}^{(k)}),$$

with $\{a_i^{(j)} : j \in [k], i \in [\ell_j]\} = \{b_i^{(j)} : j \in [k], i \in [\ell_j]\} = [n]$. Let $g \in S_n$ be the permutation such that $g(a_i^{(j)}) = b_i^{(j)}$ for all i, j. Then Lemma 3.4 implies that $y = gxg^{-1}$. □

3.3 Alternating groups

Proposition 3.2 says that every element of S_n can be written as a product of transpositions. This expression has no chance of being unique, since we can compose any such expression with $(1, 2)(1, 2) = e$ without changing the element it represents. Nonetheless, we do have the following result.

Lemma 3.6 (from *Introduction to Group Theory*)**.** *Let $n \in \mathbb{N}$ and $f \in S_n$. If $f = f_1 \cdots f_r = g_1 \cdots g_s$ with each f_i, g_i a transposition, then $r \equiv s \pmod 2$.*

We may therefore make the following definition.

Definition (signature)**.** Let $n \in \mathbb{N}$, and let $f \in S_n$ and write $f = f_1 \cdots f_r$ with each f_i a transposition. Then the *signature* of f, denoted $\epsilon(f)$, is defined to be

$$\epsilon(f) = (-1)^r.$$

We say that f is *even* if $\epsilon(f) = 1$, and *odd* if $\epsilon(f) = -1$.

Examples.

(1) The identity element is even, since it is a product of zero transpositions.

(2) By (3.2), a k-cycle is even if and only if k is odd.

Lemma 3.7. *Let $n \in \mathbb{N}$. Then the map $\epsilon : S_n \to (\{-1, 1\}, \times)$ is a homomorphism.*

Proof. Let $f, g \in S_n$ and write $f = f_1 \cdots f_r$ and $g = g_1 \cdots g_s$ with each f_i, g_i a transposition. Since $fg = f_1 \cdots f_r g_1 \cdots g_s$ is a product of transpositions, it follows that $\epsilon(fg) = (-1)^{r+s} = (-1)^r(-1)^s = \epsilon(f)\epsilon(g)$ and thus ϵ is a homomorphism. □

Proposition 3.8. *For each $n \in \mathbb{N}$ the set of even permutations in S_n forms a normal subgroup.*

Proof. This set is precisely $\ker \epsilon$, which is a normal subgroup by Theorem 2.9. □

Definition (alternating group)**.** Let $n \in \mathbb{N}$. Then the subgroup of even permutations in S_n is called the *alternating group of degree n*, denoted A_n.

Proposition 3.9. *Let $n \in \mathbb{N}$. Then A_n is generated by the set of 3-cycles in S_n.*

Proof. First, note that every 3-cycle is even by (3.2), so A_n certainly *contains* the subgroup of S_n generated by the 3-cycles. It remains to show that every element of A_n can be written as a product of 3-cycles.

By definition, every element of A_n is a product of an even number of transpositions. Every element of A_n is therefore a product of finitely many permutations of the form $(i, j)(k, \ell)$. It therefore suffices to show that every permutation of the form $(i, j)(k, \ell)$ can be written as a product of 3-cycles. There are three cases to consider. First, if $\{i, j\} = \{k, \ell\}$ then $(i, j)(k, \ell) = (i, j)(i, j) = e$, which is a product of zero 3-cycles. Next, if $\{i, j\}$ and $\{k, \ell\}$ have a single element in common, say $j = \ell$, then $(i, j)(k, \ell) = (i, j)(j, k) = (i, j, k)$. Finally, if i, j, k, ℓ are distinct then

$$(i, j)(k, \ell) = (i, j)(j, k)(j, k)(k, \ell) = (i, j, k)(j, k, \ell).$$ □

Proposition 3.10. *Let $n \geqslant 2$. Then A_n is the unique subgroup of index 2 in S_n.*

Proof. The homomorphism theorem implies that $S_n/A_n \cong (\{1, -1\}, \times)$, and hence $[S_n : A_n] = 2$. Conversely, suppose $H < S_n$ has index 2. We have $H \trianglelefteq S_n$ by Lemma 2.12, so we may consider the quotient homomorphism $\pi : S_n \to S_n/H$. If $x \in S_n$ is a 3-cycle then, since $|\pi(x)|$ divides both $|x| = 3$ and $|S_n/H| = 2$, it must be that $\pi(x) = eH$, and hence $x \in H$. In particular, H contains all 3-cycles, which generate A_n, so since $|H| = |A_n|$ we must have $H = A_n$. □

Remark. We have $A_1 = S_1 = \{e\}$, so the case $n = 1$ is an exception to Proposition 3.10.

3.4 Simplicity of alternating groups

In this section we will show that the alternating groups provide an infinite family of non-abelian finite simple groups, as follows.

Theorem 3.11. *The alternating group A_n is simple for $n = 3$ and every integer $n \geqslant 5$.*

Remark. Conversely, A_n is not simple for $n = 1, 2, 4$: we have $A_1 = \{e\}$ and $A_2 = \{e\}$, and you are invited to show in Exercise Sheet 3 Q2 that A_4 is not simple.

Lemma 3.12. *Let $n \geqslant 5$. All 3-cycles are conjugate in A_n.*

Proof. Since conjugacy is an equivalence relation, it is sufficient to show that an arbitrary 3-cycle (i, j, k) is conjugate to $(1, 2, 3)$. We claim that there is an element $g \in A_n$ such that $g(1) = i$, $g(2) = j$ and $g(3) = k$. Indeed, let g_0 be an arbitrary element of S_n such that $g_0(1) = i$, $g_0(2) = j$ and $g_0(3) = k$. If $g_0 \in A_n$ then the claim is satisfied; if not then we may take $g = g_0 \circ (4, 5)$. Lemma 3.4 then implies that $g(1, 2, 3)g^{-1} = (i, j, k)$. □

Proof of Theorem 3.11. It is easy to verify directly that $A_3 = \{e, (1, 2, 3), (1, 3, 2)\} = \langle (1, 2, 3) \rangle \cong \mathbb{Z}_3$, which is simple, so we may assume that $n \geqslant 5$. Let $N \neq \{e\}$ be a normal subgroup of A_n. We need to show that $N = A_n$.

Since $N \neq \{e\}$, we may pick a non-identity element $a \in N$. We consider different cases depending on which element we have picked.

Case 1: a is a 3-cycle. Since N is normal, this means that N also contains every conjugate of a in A_n, and by Lemma 3.12 this means that N contains every 3-cycle. Proposition 3.9 then implies that $N = A_n$, as required.

For the remaining cases, we will write

$$a = a_1 \cdots a_t$$

as a product of disjoint cycles written in non-increasing order of length. Note that, since $N \trianglelefteqslant A_n$, the element $aba^{-1}b^{-1} = a(ba^{-1}b^{-1})$ belongs to N for all $b \in A_n$.

Case 2: $|a_1| \geqslant 4$. Say $a_1 = (i_1, i_2, i_3, \ldots, i_r)$, and let $b = (i_1, i_2, i_3) \in A_n$. Then N contains the element

$$\begin{aligned} aba^{-1}b^{-1} &= (a(i_1), a(i_2), a(i_3))(i_3, i_2, i_1) && \text{(by Lemma 3.4)} \\ &= (i_2, i_3, i_4)(i_3, i_2, i_1) \\ &= (i_4, i_2, i_1), \end{aligned}$$

so N contains a 3-cycle and $N = A_n$ by case 1.

Case 3: $|a_1| = 3$ but a is not a 3-cycle. Say $a_1 = (i_1, i_2, i_3)$ and $a(i_4) \neq i_4$ for some $i_4 \notin \{i_1, i_2, i_3\}$, and let $b = (i_1, i_2, i_4) \in A_n$. Then N contains the element

$$\begin{aligned} aba^{-1}b^{-1} &= (a(i_1), a(i_2), a(i_4))(i_4, i_2, i_1) && \text{(by Lemma 3.4)} \\ &= (i_2, i_3, a(i_4))(i_4, i_2, i_1) \\ &= (i_1, i_4, i_3, a(i_4), i_2), \end{aligned}$$

so $N = A_n$ by case 2.

Case 4: all a_i are transpositions. Since $a \in A_n$, there must be at least two transpositions, say $a_1 = (i_1, i_2)$ and $a_2 = (i_3, i_4)$. Since $n \geqslant 5$ there exists $i_5 \in [n] \setminus \{i_1, i_2, i_3, i_4\}$, so we may set $b = (i_1, i_3, i_5)$. Then N contains the element

$$\begin{aligned} aba^{-1}b^{-1} &= (a(i_1), a(i_3), a(i_5))(i_5, i_3, i_1) && \text{(by Lemma 3.4)} \\ &= (i_2, i_4, a(i_5))(i_5, i_3, i_1). \end{aligned}$$

If $a(i_5) = i_5$ then this is the 5-cycle $(i_1, i_2, i_4, i_5, i_3)$ and we are done by case 2, whilst if $a(i_5) \neq i_5$ then $(i_2, i_4, a(i_5))$ and and (i_5, i_3, i_1) are disjoint 3-cycles and we are done by case 3. □

Chapter 4

Group actions

4.1 Group actions

In this chapter we introduce one of the most important concepts of the course: group actions. Group actions are one of the main ways in which groups interact with other mathematical objects, and are the means by which group theory arises in numerous other branches of mathematics.

We have seen various examples of groups that naturally permute the elements of certain sets. For example, the symmetric group permutes the integers $1, 2, \ldots, n$; the dihedral group D_{2n} permutes the vertices of the regular n-gon; and the linear group $\mathrm{GL}_n(\mathbb{R})$ permutes the elements of the vector space $\mathbb{R}^n$. Group actions provide a general framework for studying groups permuting the elements of sets.

Definition (group action)**.** Let G be a group and X a non-empty set. An *action of G on X*, or *G-action* on X, is a map

$$\begin{array}{ccc} G \times X & \to & X \\ (g, x) & \mapsto & g(x) \end{array}$$

that satisfies

(1) $e(x) = x$ for all $x \in X$;

(2) $(gh)(x) = g(h(x))$ for all $g, h \in G$ and $x \in X$.

We will use the notation $G \curvearrowright X$ to denote an action of G on X, or to mean 'G acts on X' or similar.

Example. As well as acting on the vertices of the regular n-gon P_n, the dihedral group D_{2n} can naturally be seen to be acting on the set of edges, or on the set of lines passing through the centre of P_n and a vertex. As well as acting on the elements of $\mathbb{R}^n$, the general linear group $\mathrm{GL}_n(\mathbb{R})$ acts on the sets of lines, planes, parallelopipeds, ellipsoids,....

Lemma 4.1. *Let G be a group and X a non-empty set. Suppose that for each $g \in G$ we assign a map $g : X \to X$. Then this assignment defines an action of G on X if and only if each map $g : X \to X$ is a bijection and the map $\varphi : G \to \mathrm{Sym}(X)$ defined by setting $\varphi(g)(x) = g(x)$ for all $g \in G$ and $x \in X$ is a homomorphism.*

Definition. We call φ the *homomorphism corresponding to the action.*

Proof. Suppose first that φ is a homomorphism. Then Lemma 2.2 implies that $\varphi(e) = e$, so that condition (1) of the definition of an action holds, and $\varphi(gh) = \varphi(g) \circ \varphi(h)$ for all $g, h \in G$ by definition of a homomorphism, so that condition (2) holds.

Conversely, suppose that the maps $g : X \to X$ define an action. Then given an arbitrary element $g \in G$, we have $g(g^{-1}(x)) = e(x) = x$ and $g^{-1}(g(x)) = e(x) = x$ for all $x \in X$, so that the map $g : X \to X$ is invertible (with inverse g^{-1}) and hence a bijection. We may therefore define a map $\varphi : G \to \mathrm{Sym}(X)$ by setting $\varphi(g)(x) = g(x)$ for each $g \in G$ and $x \in X$, and then condition (2) of the definition of an action implies that φ is a homomorphism. □

Example (the trivial action)**.** Given any group G and any non-empty set X, we can define an action via $g(x) = x$ for all $g \in G$ and $x \in X$. This is called the *trivial action* of G on X, and corresponds to the trivial homomorphism $G \to \mathrm{Sym}(X)$.

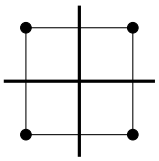
Example (the left-multiplication action)**.** Let G be a group. Then we may define an action of G on itself by setting $g(x) = gx$ for every $g, x \in G$. This is called the *left-multiplication action.* More generally, given a subgroup $H \leqslant G$, the group G acts on G/H by left multiplication: $g(xH) = gxH$.

Example (the conjugation action)**.** Let G be a group. Then we may define an action of G on itself by *conjugation*: $g(x) = gxg^{-1}$ for every $g, x \in G$.

Definition (faithful action; transitive action)**.** An action of a group G on a set X is called *transitive* if for every $x, y \in X$ there exists $g \in G$ such that $y = g(x)$. It is called *faithful* if distinct elements of G always give rise to distinct maps $X \to X$, i.e. if for all distinct $g, h \in G$ there exists $x \in X$ such that $g(x) \neq h(x)$, i.e. if the homomorphism corresponding to the action is injective.

Examples.

- A trivial action $G \curvearrowright X$ is not faithful unless $G = \{e\}$, and is not transitive unless $|X| = 1$.
- The left-multiplication action of a group on itself is both faithful and transitive.
- The group D_8 acts on the horizontal and vertical lines passing through the centre of the square, as illustrated in the following diagram.



  This action is transitive (e.g. reflection in a diagonal swaps the two lines), but not faithful (rotation by $\pi/2$ also swaps the two lines).

- Given an arbitrary group G acting on a set X, we may also define an action of G on the Cartesian product $X \times X$ via $g((x, y)) = (g(x), g(y))$. If the original action of G on X was faithful then this new action will also be faithful, but it is never transitive if $|X| > 1$, since, for example, if $x, y \in X$ with $x \neq y$ then there is no $g \in G$ such that $g((x, x)) = (x, y)$.

Definition (restrictions of actions)**.** Let G be a group acting on a non-empty set X.

- If $H \leqslant G$ is a subgroup then we may define an action of H on X by

$$\begin{array}{ccc} H \times X & \to & X \\ (g, x) & \mapsto & g(x) \end{array}$$

  This is called the *restriction* of the action to H.

- If a subset $Y \subseteq X$ satisfies $g(Y) \subseteq Y$ for all $g \in G$ then Y is said to be *invariant* under the action of G on X. In this case, we may define an action of G on Y by

$$\begin{array}{ccc} G \times Y & \to & Y \\ (g, x) & \mapsto & g(x) \end{array}$$

  This is called the *restriction* of the action to Y.

4.2 Orbits and stabilisers

We now introduce two of the most important defitions concerning group actions.

Definition (orbit; stabiliser)**.** Let G be a group acting on a set X. Then for every $x \in X$, the set

$$\mathrm{Orb}_G(x) = \{g(x) : g \in G\}$$

is called the *orbit* of x, and the set

$$\mathrm{Stab}_G(x) = \{g \in G : g(x) = x\}$$

is called the *stabiliser* of x.

Lemma 4.2 (stabilisers are subgroups)**.** *Suppose G is a group acting on a set X, and let $x \in X$. Then* $\mathrm{Stab}_G(x) \leqslant G$.

Proof. The conditions of Lemma 1.9 are trivial to verify. □

Remark. Conversely, you will see in the exercise sheet that every subgroup is a stabiliser for some action.

The most important result about orbits and stabilisers is the following famous theorem.

Theorem 4.3 (orbit–stabiliser theorem)**.** *Suppose G is a group acting on a set X, and let $x \in X$. Then*

$$|\mathrm{Orb}_G(x)| = [G : \mathrm{Stab}_G(x)].$$

In particular, if G is finite then $|G| = |\mathrm{Orb}_G(x)|\,|\mathrm{Stab}_G(x)|$ by Lagrange's theorem.

Note that $|\mathrm{Orb}_G(x)| = [G : \mathrm{Stab}_G(x)] = \infty$ is possible, e.g. for the left-multiplication action of an infinite group G on itself.

Given a group G acting on a set X, and elements $x, y \in X$, we will denote by $G_{x \to y}$ the set

$$G_{x \to y} = \{g \in G : g(x) = y\}.$$

Lemma 4.4. *Suppose G is a group acting on a set X. Let $x \in X$, $y \in \mathrm{Orb}_G(x)$, and $h \in G_{x\to y}$. Then $G_{x\to y} = h\,\mathrm{Stab}_G(x)$.*

Proof. Given $g \in G$ we have

$$\begin{aligned} g \in G_{x\to y} &\iff g(x) = h(x) \iff h^{-1}g(x) = x \\ &\iff h^{-1}g \in \mathrm{Stab}_G(x) \iff g \in h\,\mathrm{Stab}_G(x), \end{aligned}$$

as required. □

Proof of Theorem 4.3. By Lemma 4.4 we may define a map

$$\begin{array}{ccc} \mathrm{Orb}_G(x) & \to & G/\mathrm{Stab}_G(x) \\ y & \mapsto & G_{x\to y}. \end{array}$$

We claim that this map is a bijection. If y, y' are distinct elements of X then $G_{x\to y} \cap G_{x\to y'} = \varnothing$, so it is certainly injective. To see that it is surjective, note that for all $g \in G$ we have $g\,\mathrm{Stab}_G(x) = G_{x\to g(x)}$ by Lemma 4.4. □

Let us now record some other miscellaneous results about orbits.

Lemma 4.5. *Let G be a group acting on a set X. Then the relation $\sim$ defined by $x \sim y$ if $y \in \mathrm{Orb}_G(x)$ is an equivalence relation.*

Reflexivity. For every $x \in X$ we have $x = e(x)$.
Symmetry. For every $x, y \in X$, if $y = g(x)$ then $x = g^{-1}(y)$.
Transitivity. For every $x, y, z \in X$, if $y = g(x)$ and $z = h(y)$ then $z = hg(x)$. □

The orbits of a G-action on X are therefore the equivalence classes of this equivalence relation. In particular, for every $x, y \in X$ we have either $\mathrm{Orb}_G(x) = \mathrm{Orb}_G(y)$ or $\mathrm{Orb}_G(x) \cap \mathrm{Orb}_G(y) = \varnothing$, and the orbits partition X.

Definition (fixed point)**.** Let G be a group acting on a set X. Then an element $x \in X$ is called a *fixed point* for this action if $\mathrm{Orb}_G(x) = \{x\}$, or equivalently if $\mathrm{Stab}_G(x) = G$. We write

$$\mathrm{Fix}_G(X) = \{x \in X : \mathrm{Orb}_G(x) = \{x\}\}$$

for the set of fixed points.

Remark. Let G be a group acting on a set X. Then the set $\mathrm{Fix}_G(X)$ is invariant under the G-action. More generally, a subset $Y \subseteq X$ is invariant under the G-action if and only if Y is a union of orbits.

4.3 The conjugation action

In this section we will consider the conjugation action of a group on itself in a bit more detail. Let us start by noting that an arbitrary group G also acts on the set of all its subgroups by conjugation, i.e. $g(H) = gHg^{-1}$. It is a useful exercise to check this using the following lemma.

Lemma 4.6. *Let G be a group, and let $H \leqslant G$ and $g \in G$. Then $gHg^{-1} \leqslant G$ and $gHg^{-1} \cong H$.*

Proof. Lemma 1.8 shows that the conjugation map $x \mapsto gxg^{-1}$ is an isomorphism, in particular an injective homomorphism, and so its restriction to H is also an injective homomorphism. The conjugate gHg^{-1} is therefore a subgroup of G by Lemma 2.2 (3), and isomorphic to H by Corollary 2.3. □

Warning. The converse to Lemma 4.6 does not hold: there exist a group G with subgroups H, K that are isomorphic but not conjugate (see the exercise sheet).

Warning. Two subgroups of $H < G$ can be conjugate in G but not in H. Can you find an example where this occurs?

The conjugation action is so important that orbits and stabilisers have different names reserved just for this action, as follows.

Definition (conjugacy class; centraliser; normaliser)**.** Let G be a group and let $x \in G$. Then the *conjugacy class* of x , denoted x^G, is defined by

$$x^G = \{gxg^{-1} : g \in G\}.$$

The *centraliser* of x, denoted $C_G(x)$, is defined by

$$C_G(x) = \{g \in G : gxg^{-1} = x\} = \{g \in G : gx = xg\}.$$

Equivalently, x^G is the orbit of x under the conjugation action of G on itself, and $C_G(x)$ is the stabiliser of x under this action.

Given a subgroup $H \leqslant G$, the *normaliser* of H, denoted $N_G(H)$, is defined by

$$N_G(H) = \{g \in G : gHg^{-1} = H\} = \{g \in G : gH = Hg\}.$$

Equivalently, $N_G(H)$ is the stabiliser of H under the conjugation action of G on the set of subgroups of G.

We will now translate some results from the previous sections from the setting of general group actions to the setting of the conjugation action of G on itself. It is a useful exercise to prove these statements directly. The first is just Lemma 4.2 applied to the conjugation action of G on itself or on its set of subgroups.

Corollary 4.7 (centralisers and normalisers are subgroups)**.** *Let G be a group. Then $C_G(x) \leqslant G$ for all $x \in G$, and $N_G(H) \leqslant G$ for all $H \leqslant G$.*

Corollary 4.8 (orbit–stabiliser theorem for conjugation)**.** *Let G be a group and let $x \in G$. Then*

$$|x^G| = [G : C_G(x)].$$

In particular, if G is finite then $|G| = |x^G|\,|C_G(x)|$ by Lagrange's theorem.

Corollary 4.9 (orbit–stabiliser theorem for conjugation action on subgroups)**.** *Let G be a group and suppose $H \leqslant G$. Then $|\{K \leqslant G : K \text{ conjugate to } H\}| = [G : N_G(H)]$.*

Lemma 4.5 implies the following.

Corollary 4.10. *Let G be a group. Then conjugacy is an equivalence relation on the elements of G. In particular, the conjugacy classes of G partition G.*

Chapter 5

Sylow's theorems

In this chapter we will introduce Sylow's theorems, which are some of the most important theorems describing the basic structure of finite groups.

5.1 Sylow's first theorem

Sylow's first theorem is the following partial converse to Lagrange's theorem.

Theorem 5.1 (Sylow's first theorem)**.** *Let G be a finite group, and let p^k be a prime power dividing $|G|$ (so p is prime and $k \in \mathbb{N}_0$). Then G has a subgroup of order p^k.*

Remark. You have seen in Exercise Sheet 3 Q5(a)(i) that the group A_4, which has order 12, has no subgroup of order $6 = 3 \times 2$. Thus, in a sense, Sylow's first theorem can be thought of as the best possible general converse to Lagrange's theorem.

In the proof, we will use the fact that $Z(G)$ is the union of the conjugacy classes of size 1 in G (we have $z \in Z(G)$ if and only if $gzg^{-1} = z$ for all $g \in G$, if and only if $|z^G| = 1$).

Proof. The theorem is trivially true when $G = \{e\}$, so we may assume that $|G| > 1$ and, by induction, that the theorem holds for all smaller groups. Moreover, $\{e\}$ is the required subgroup when $k = 0$, so it suffices to assume that $k \geqslant 1$, and in particular that p divides $|G|$.

We divide the proof into two cases. First, suppose that p divides $|Z(G)|$. Cauchy's theorem for abelian groups (Theorem 2.13) then implies that $Z(G)$ contains an element x of order p, and then Proposition 1.18 implies that $|\langle x \rangle| = p$. The fact that x is central implies that $\langle x \rangle \trianglelefteqslant G$, and Lagrange's theorem implies that p^{k-1} divides $|G/\langle x \rangle|$. By induction, therefore, $G/\langle x \rangle$ contains a subgroup of order p^{k-1}. The correspondence theorem (Theorem 2.19) therefore implies that G contains a subgroup of order p^k, as required.

In the second case, suppose that p does not divide $|Z(G)|$. Let $\mathcal{C}(G)$ be the set of conjugacy classes in G, so that $|G| = \sum_{C \in \mathcal{C}(G)} |C|$. Since p does not divide

$$|Z(G)| = \sum_{\substack{C \in \mathcal{C}(G) \\ |C|=1}} |C|,$$

it follows that p does not divide

$$\sum_{\substack{C \in \mathcal{C}(G) \\ |C| \geqslant 2}} |C|.$$

In particular, there must exist a conjugacy class C such that $|C| \geqslant 2$ and p does not divide $|C|$. Letting $g \in C$, the orbit–stabiliser theorem implies that $|C| = [G : C_G(g)]$, so Lagrange's theorem implies that p^k divides $|C_G(g)|$ and that $|C_G(g)| < |G|$. By induction, we may therefore conclude that there is a subgroup of order p^k in $C_G(g)$, and hence in G. □

An immediate corollary of Sylow's first theorem is the following famous result.

Corollary 5.2 (Cauchy's theorem)**.** *Let p be a prime, let G be a finite group, and suppose that p divides $|G|$. Then G contains an element of order p.*

Proof. Sylow's first theorem implies that G contains a subgroup of order p, which is cyclic by Corollary 1.24, so is generated by an element of order p. □

5.2 Sylow p-subgroups

Definition (p-group)**.** Let p be a prime. Then a group G is called a *p-group* if for every $x \in G$ we have that $|x|$ is a power of p. If $H \leqslant G$ is a p-group then it is called a *p-subgroup* of G.

Corollary 5.3. *A finite group G is a p-group if and only if $|G| = p^m$ for some $m \in \mathbb{N}_0$.*

Proof. If $|G| = p^m$ then Lagrange's theorem implies that every element has order dividing p^m, and hence a power of p. Conversely, if $|G|$ is divisible by a prime $q \neq p$ then Cauchy's theorem implies that G has an element of order q, which is not a power of p. □

Definition (Sylow p-subgroup)**.** Let G be a group and let p be a prime. A p-subgroup H of G is a *Sylow p-subgroup* if it is a maximal p-subgroup of G, in the sense that it is not a proper subgroup of any other p-subgroup of G.

Given a group G, we write $\mathrm{Syl}_p(G)$ for the set of Sylow p-subgroups of G, and $n_p(G)$ for the number of Sylow p-subgroups, i.e. $n_p(G) = |\mathrm{Syl}_p(G)|$.

Given a prime p, every finite group G has at least one Sylow p-subgroup. Indeed, $\{e\}$ is a p-subgroup, so G has a at least one p-subgroup, and so since it is finite it has a p-subgroup of maximum size. This is then automatically a Sylow p-subgroup. Sylow's first theorem implies the stronger statement that if G has order $p^r m$, with m not divisible by p, then G at least one Sylow p-subgroup of order p^r (a subgroup of order p^r must be a Sylow p-subgroup by Lagrange's theorem and Corollary 5.3).

5.3 Sylow's second theorem

It is easy to check that a group G acts by conjugation on its Sylow p-subgroups, as follows.

Lemma 5.4. *Let G be a group, let p be a prime, and let $H \leqslant G$ be a p-subgroup. Then for every $g \in G$ the conjugate gHg^{-1} is also a p-group. Moreover, if H is a Sylow p-subgroup then so is gHg^{-1}.*

Proof. Lemma 4.6 implies that gHg^{-1} is a subgroup isomorphic to H, and hence a p-group by Lemma 1.6 (4). If gHg^{-1} were not a Sylow p-subgroup then it would be a proper subgroup of some p-group $K \leqslant G$, but then H would be a proper subgroup of the p-group $g^{-1}Kg$, and so H would not be a Sylow p-subgroup either. □

Sylow's second theorem asserts that this action is transitive for a finite group.

Theorem 5.5 (Sylow's second theorem)**.** *Let G be a finite group and let p be a prime. Then the Sylow p-subgroups of G are all conjugate to one another.*

Proof. Expressing $|G| = p^r m$ with $r \in \mathbb{N}_0$, $m \in \mathbb{N}$ and $p \nmid m$, Theorem 5.1 implies that there exists a Sylow p-subgroup $P \leqslant G$ of order p^r; it suffices to show that an arbitrary Sylow p-subgroup of G is conjugate to P. Let H be a Sylow p-subgroup of G, and let H act on G/P via $h(gP) = hgP$. Let $\mathcal{O}$ be the set of orbits of this action, and recall that the orbits partition G/P (see Lemma 4.5) so that

$$m = [G : P] = \sum_{O \in \mathcal{O}} |O|.$$

In particular, there exists at least one orbit O with $|O|$ not divisible by p. However, $|O|$ divides $|H|$ by the orbit-stabiliser theorem (Theorem 4.3), and $|H|$ is a power of p by Corollary 5.3, so this forces $|O| = 1$. This means that the action of H on G/P has a fixed point, i.e. that there exists $g \in G$ such that $hgP = gP$ for all $h \in H$. However, this implies in particular that $g^{-1}hgP = P$ for all $h \in H$, hence that $g^{-1}hg \in P$ for all $h \in H$ by Lemma 1.21, and hence that $g^{-1}Hg \leqslant P$. Since $g^{-1}Hg$ is a Sylow p-subgroup by Lemma 5.4, by maximality we must have $g^{-1}Hg = P$, as required. □

Corollary 5.6. *Let G be a finite group, let p be a prime, and let P be a Sylow p-subgroup of G. Then $n_p(G) = [G : N_G(P)]$. In particular, $P \trianglelefteq G$ if and only if P is the unique Sylow p-subgroup of G.*

Proof. Lemma 5.4 implies that G acts on $\mathrm{Syl}_p(G)$ by conjugation; Sylow's second theorem implies that $\mathrm{Orb}_G(P) = \mathrm{Syl}_p(G)$; and then the orbit–stabiliser theorem implies that $n_p(G) = [G : N_G(P)]$. □

5.4 Sylow's third theorem

Theorem 5.7 (Sylow's third theorem)**.** *Let p be a prime, and let G be a finite group. Then*

(1) $n_p(G) \equiv 1 \pmod{p}$, and

(2) $n_p(G)$ divides $|G|$.

The main additional ingredient we need for Theorem 5.7 is the following lemma.

Lemma 5.8. *Let G be a finite group, let p be a prime, and let P be a Sylow p-subgroup of G. Let P act on $\mathrm{Syl}_p(G)$ by conjugation (which we may do by Lemma 5.4). Then $\mathrm{Fix}_P(\mathrm{Syl}_p(G)) = \{P\}$.*

Proof. Certainly $P \in \mathrm{Fix}_P(\mathrm{Syl}_p(G))$, so suppose $Q \in \mathrm{Fix}_P(\mathrm{Syl}_p(G))$. We claim that PQ is a p-subgroup of G; since P and Q are both Sylow p-subgroups contained in PQ, this will imply that $Q = PQ = P$, as required. By definition, $P \leqslant N_G(Q)$. Since $Q \trianglelefteq N_G(Q)$, Lemma 2.16 therefore implies that $PQ \leqslant N_G(Q)$ and that, writing $\pi : N_G(Q) \to N_G(Q)/Q$ for the quotient homomorphism, $\pi(P) = PQ/Q$. The first isomorphism theorem therefore implies that $PQ/Q \cong P/(P \cap Q)$. Lagrange's theorem therefore implies that $[PQ : Q]$ divides $|P|$, and hence is a power of p by Corollary 5.3. Since $|Q|$ is also a power of p by Corollary 5.3, Lagrange's theorem implies that $|PQ|$ is a power of p, and hence, by Corollary 5.3, that PQ is a p-group, as claimed. □

Proof of Theorem 5.7.

(1) Let P be a Sylow p-subgroup of G, and note that P acts on $\mathrm{Syl}_p(G)$ by conjugation by Lemma 5.4. Partition $\mathrm{Syl}_p(G)$ into the orbits of this action, and note that the orbit–stabiliser theorem implies that those orbits with size greater than 1 have size divisible by p. Lemma 5.8 implies that the unique orbit with size 1 is $\{P\}$, so $\mathrm{Syl}_p(G) \equiv 1 \pmod{p}$.

(2) Corollary 5.6 and Lagrange's theorem show that $n_p(G)$ divides $|G|$.

□

In summary, Sylow's theorems tell us the following.

Theorem 5.9 (summary of Sylow's theorems)**.** *Let G be a finite group, and suppose $|G| = p^r m$ with p prime, $r \in \mathbb{N}_0$, $m \in \mathbb{N}$, and $p \nmid m$. Then the Sylow p-subgroups of G are exactly the subgroups of order p^r, and they are all conjugate to one another. Moreover, $n_p(G) \equiv 1 \pmod{p}$ and $n_p(G) \mid m$.*

Proof. Lagrange's theorem and Corollary 5.3 imply that every subgroup of order p^r is a Sylow p-subgroup. Sylow's first theorem implies that there exists some such group, and Sylow's second theorem implies that every other Sylow p-subgroup is conjugate to it, and hence, in particular, of the same cardinality by Lemma 4.6. Sylow's third theorem implies that $n_p(G) \equiv 1 \pmod{p}$, which in particular means that $n_p(G)$ is not divisible by p. Sylow's third theorem also implies that $n_p(G)$ divides $|G|$, which combined with the fact that $n_p(G)$ is not divisible by p implies that $n_p(G) \mid m$. □

Chapter 6

The smallest non-abelian finite simple group

The main aim of this chapter is to show that A_5 is the smallest non-abelian finite simple group, as follows.

Theorem 6.1. *Let G be a non-abelian finite simple group. Then $|G| \geqslant 60$.*

Theorem 6.2. *Let G be a simple group of order 60. Then $G \cong A_5$.*

This is one small step on the great journey that is the classification of finite simple groups.

From now on, given groups H and G, we will write $H \lesssim G$ to mean that H is isomorphic to a subgroup of G, i.e. that there exists $K \leqslant G$ such that $H \cong K$.

6.1 Ruling out the existence of simple groups with given order

The ostensible aim of this section is to prove Theorem 6.1 by ruling out the existence of non-abelian finite simple groups with order less than 60. However, it also has the broader aim of developing general techniques for ruling out the existence of simple groups with given order.

We start by classifying the finite simple p-groups.

Proposition 6.3. *Let p be a prime and let G be a non-trivial finite p-group. Then $|Z(G)|$ is divisible by p. In particular, since $e \in Z(G)$, we have $|Z(G)| \geqslant p$.*

Proof. Write $\mathcal{C}(G)$ for the set of conjugacy classes in G. As in the proof of Sylow's first theorem, we have

$$|G| = |Z(G)| + \sum_{\substack{C \in \mathcal{C}(G) \\ |C|>1}} |C|.$$

Corollary 5.3 implies that $|G| = p^m$ for some $m \in \mathbb{N}$. This in turn implies that p divides $|G|$ and, by the orbit–stabiliser theorem, that p divides $|C|$ for all $C \in \mathcal{C}(G)$ with $|C| > 1$, so that p divides $|Z(G)|$ as required. □

Corollary 6.4. *Let p be a prime and let G be a finite p-group. Then G is simple if and only if $G \cong \mathbb{Z}_p$.*

Proof. Suppose G is a finite simple p-group. Proposition 6.3 implies that $Z(G)$ is a non-trivial subgroup of G, and it is trivially normal, so we must have $Z(G) = G$, so that G is abelian. The desired result then follows from Theorem 2.21. □

Lemma 6.5. *Let G be a finite simple group, and let p be a prime dividing $|G|$. Then either $G \cong \mathbb{Z}_p$ or $n_p(G) > 1$.*

Proof. Sylow's first theorem implies that G has at least one non-trivial Sylow p-subgroup P. Corollary 6.4 implies that either $G \cong \mathbb{Z}_p$ or $G \neq P$, and in the latter case simplicity of G and Corollary 5.6 imply that $n_p(G) > 1$, as required. □

Proposition 6.6. *There are no simple groups of orders 40 or 54.*

Proof. Let G be group. If $|G| = 40$ then Sylow's third theorem implies that $n_5(G) = 1$, so G is not simple by Lemma 6.5. If $|G| = 54$ then Sylow's third theorem implies that $n_3(G) = 1$, so G is not simple by Lemma 6.5. □

Proposition 6.7. *There are no simple groups of order 56.*

Proof. Let G be a group of order 56. If $n_7(G) = 1$ then G is not simple by Lemma 6.5. If $n_7(G) > 1$ then Sylow's third theorem implies that $n_7(G) = 8$. Sylow's theorems imply that each Sylow 7-subgroup has order 7, and Corollary 1.25 implies that each distinct pair of these subgroups them intersects only at the identity, so G has $8 \times 6 = 48$ distinct elements of order 7. This leaves only enough room for a unique Sylow 2-subgroup, since these have order 8 by Sylow's theorems, so Lemma 6.5 implies that G is not simple. □

Proposition 6.8. *Let p, q, r be primes, not necessarily distinct. Then there are no simple groups of order pq or pqr.*

Remark. The simple group A_5 has order $60 = 2^2 \times 3 \times 5$, so there is no equivalent of Proposition 6.8 for products of four primes.

Proof. Let G be a finite simple group. Suppose first that $|G| = pq$. Corollary 6.4 implies that $p \neq q$, and so Sylow's third theorem and Lemma 6.5 imply that $n_p(G) = q$ and $n_q(G) = p$, and then Sylow's third theorem implies that $p \equiv 1 \pmod{q}$ and $q \equiv 1 \pmod{p}$. This gives a contradiction, since the smaller of p and q cannot be 1 mod the larger.

Suppose next that $|G| = pq^2$. Corollary 6.4 again implies that $p \neq q$, and then Sylow's third theorem and Lemma 6.5 imply that $n_p(G) \in \{q, q^2\}$ and $n_q(G) = p$. If $n_p(G) = q$ then we have the same contradiction as before, so we must have $n_p(G) = q^2$. Since a Sylow p-subgroup of G has order p, Corollary 1.25 then implies that there are $q^2(p-1)$ elements of order p in G. However, $q^2(p-1) = |G| - q^2$, so there is only enough room left for a unique Sylow q-subgroup, contradicting Lemma 6.5.

Finally, suppose that $|G| = pqr$. Corollary 6.4 and the previous paragraph imply that p, q and r are all distinct, so without loss of generality we may assume that $p < q < r$. Sylow's third theorem

implies that $n_r(G)$ divides pq and is $1 \pmod r$, so Lemma 6.5 and the fact that $p < q < r$ forces $n_r(G) = pq$. Sylow's third theorem also implies that $n_q(G)$ divides pr and is $1 \pmod q$, so Lemma 6.5 and the fact that $p < q$ forces $n_q(G) \geqslant r$. Finally, Sylow's third theorem implies that $n_p(G)$ divides qr, so Lemma 6.5 and the fact that $q < r$ forces $n_p(G) \geqslant q$. Corollary 1.25 therefore implies that G contains $pq(r-1)$ distinct elements of order r, at least $r(q-1)$ distinct elements of order q, and at least $q(p-1)$ distinct elements of order p. The total number of distinct elements in G is therefore at least

$$\begin{aligned} pq(r-1) + r(q-1) + q(p-1) &= |G| - pq + rq - r + qp - q \\ &= |G| + rq - r - q \\ &> |G| \end{aligned}$$

which is again a contradiction. □

To finish the proof of Theorem 6.1 we use the following straightforward but very useful lemma.

Lemma 6.9. *Suppose G is a simple group acting non-trivially on a set X. Then then the action is faithful and $G \lesssim \mathrm{Sym}(X)$.*

Proof. Let $\varphi : G \to \mathrm{Sym}(X)$ be the homomorphism corresponding to the action. We have $\ker \varphi \neq G$ because the action is not trivial, so simplicity implies that $\ker \varphi = \{e\}$, so that the action is indeed faithful and $\varphi : G \to \varphi(G)$ is an isomorphism. □

Corollary 6.10. *Let G be a non-abelian finite simple group and let p be a prime dividing $|G|$. Then $G \lesssim S_{n_p(G)}$. In particular, $|G|$ divides $n_p(G)!$ by Lagrange's theorem.*

Proof. Sylow's second theorem and Lemma 6.9 imply that $G \lesssim \mathrm{Sym}(\mathrm{Syl}_p(G))$. □

Proof of Theorem 6.1. Suppose $|G| \leqslant 60$. Corollary 6.4 implies that $|G|$ is not a prime power, Proposition 6.8 rules out all orders that are products of two or three primes, and Propositions 6.6 and 6.7 rule out $|G| \in \{40, 54, 56\}$, leaving $|G| \in \{24, 36, 48, 60\}$ as the only possibilities. If $|G| = 24$ or 48 then $n_2(G) = 3$ by Sylow's third theorem and Lemma 6.5, whilst if $|G| = 36$ then $n_3(G) = 4$ by Sylow's third theorem and Lemma 6.5. In each case this contradicts Corollary 6.10, so it must be that $|G| = 60$ as required. □

6.2 Simple groups of order 60

The proof of Theorem 6.2 uses a further result of independent interest.

Proposition 6.11. *Let $n \geqslant 5$ be an integer, and suppose $H \leqslant A_n$ has index n. Then $H \cong A_{n-1}$.*

Remark. In Exercise Sheet 6 Q5(a) and Q5(b) you are invited to show that A_n has no proper subgroups of index less than n except when $n = 4$. Combined with Proposition 6.11, this classifies all the subgroups of A_n of index at most n up to isomorphism.

Exercise Sheet 6 Q5(c) asks whether it is true that every subgroup of A_n of index n is the stabiliser of some $k \in [n]$ (Proposition 6.11 only shows that every such subgroup is *isomorphic* to the stabiliser of some $k \in [n]$).

It will be convenient to separate out the following general lemma.

Lemma 6.12. *Suppose G is a group acting on a set X, and $F \subseteq \mathrm{Fix}_G(X)$. Then $X \setminus F$ is invariant under the action of G, and the action of G on X is faithful if and only if its restriction to $X \setminus F$ is faithful.*

Proof. Since F consists of single-element orbits and the orbits partition X, the set $X \setminus F$ is a union of orbits and hence invariant as claimed.

If the restriction to $X \setminus F$ then certainly the action on the whole of X is faithful. Conversely, suppose that the action of G on X is faithful. This means that for every non-identiy $g \in G$ there exists $x \in X$ such that $g(x) \neq x$; since $F \subseteq \mathrm{Fix}_G(X)$, it must be that these elements x belong to $X \setminus F$, so that the restricted action is indeed faithful. □

Proof of Proposition 6.11. Let A_n act on A_n/H by left multiplication, i.e. $g(xH) = gxH$. Since this action is transitive, it is certainly not trivial, so it is faithful by Theorem 3.11 and Lemma 6.9. The restriction of this action to H is therefore a faithful action of H on A_n/H.

Now Lemma 1.21 (2) implies that H (viewed as an element of A_n/H) is fixed by every element of H (viewed as the group acting on A_n/H). Lemma 6.12 therefore implies that left multiplication (i.e. $g(xH) = gxH$) also defines a faithful action of H on $A_n/H \setminus \{H\}$. The homomorphism corresponding to this action is therefore injective, and Corollary 2.3 therefore implies that $H \lesssim \mathrm{Sym}(A_n/H \setminus \{H\}) \cong S_{n-1}$, say $H \cong K \leqslant S_{n-1}$. Proposition 3.10 then implies that $K = A_{n-1}$. □

Proof of Theorem 6.2. Sylow's third theorem and Lemma 6.5 imply that $n_5(G) = 6$. Corollary 6.10 therefore implies that $G \lesssim S_6$, say $G \cong G' \leqslant S_6$.

We claim that $G' \leqslant A_6$. To see this, let $\pi : S_6 \to S_6/A_6$ be the quotient homomorphism, and consider its restriction $\pi|_{G'} : G' \to S_6/A_6$. Since $|G'| = 60 > 2 = |S_6/A_6|$, this restriction homomorphism is not injective, so $\ker \pi|_{G'} \neq \{e\}$. Simplicity therefore implies that $\ker \pi|_{G'} = G'$, hence $G' \leqslant A_6$ as claimed. Proposition 6.11 then implies that $G' \cong A_5$. □

Chapter 7

Free groups and presentations

7.1 Free groups

Suppose you are told that G is a group generated by elements x and y, but are given no further information. Given two arbitrary expressions in the generators, say $x^2y^5x^{-3}$ and y^2x^2, there is not much chance of being able to tell whether they are equal in G. For certain specific expressions, however, this is possible. For example, we do not need any further information about G to know that $xyy^{-1} = x$.

Roughly speaking, the group G is called *free* on x and y if the only relations between x and y that hold in G are those, like $xyy^{-1} = x$, that are forced by the definition of a group. The aim of this section is to define this notion more precisely.

Definition (word)**.** Let X be a set. A *word* w on X is a finite sequence of symbols of the form $w = x_1^{\epsilon_1} \cdots x_k^{\epsilon_k}$, with each $x_i \in X$ and each $\epsilon_i \in \{1, -1\}$. We define k to be the *length* of w, written $|w|$. If $k = 0$ then w is called the *empty word*, denoted e.

Given $x \in X$, we normally abbreviate x^1 by x. Given in addition $n \in \mathbb{N}$, we write x^n to mean the word $xx \cdots x$ consisting of n copies of x, and x^{-n} to mean the word $x^{-1}x^{-1} \cdots x^{-1}$ consisting of n copies of x^{-n}.

If $w = x_1^{\epsilon_1} \cdots x_k^{\epsilon_k}$ and $v = y_1^{\delta_1} \cdots y_\ell^{\delta_\ell}$ are words on X then we define the *product word* $wv = x_1^{\epsilon_1} \cdots x_k^{\epsilon_k} y_1^{\delta_1} \cdots y_\ell^{\delta_\ell}$, and the *inverse word* $w^{-1} = x_k^{-\epsilon_k} \cdots x_1^{-\epsilon_1}$.

If $w = x_1^{\epsilon_1} \cdots x_k^{\epsilon_k}$ is a word on X then a sequence $x_i^{\epsilon_i} x_{i+1}^{\epsilon_{i+1}} \cdots x_j^{\epsilon_j}$ with $i \leqslant j$ is called a *subword* of w. The empty word is also considered a subword of every word.

Definition (equivalent words)**.** Two words w, w' on a set X are called *adjacent* if there exist $x \in X$, $\epsilon \in \{1, -1\}$, and words u, v such that either $w = uv$ and $w' = ux^\epsilon x^{-\epsilon} v$, or $w = ux^\epsilon x^{-\epsilon} v$ and $w' = uv$. Two words w, w' are then called *equivalent*, written $w \sim w'$, if there exist finitely many words $w_1, \ldots, w_m$ such that $w = w_1$, $w' = w_m$, and each pair w_i, w_{i+1} is adjacent. The relation $\sim$ is clearly an equivalence relation, and we denote the equivalence class of a word w by $[w]$.

Note that if X is a subset of a group G then equivalent words in X always represent the same element of G.

Proposition 7.1. *Let X be a set, and let F be the set of equivalence classes of words on X. Then we may define a product on F via $[v][w] = [vw]$, and this product makes F into a group with generating set $\{[x] : x \in X\}$.*

Proof. First note that if $v \sim v'$ and $w \sim w'$ then $vw \sim v'w \sim v'w'$, so the product is well defined. It is also associative, since $([u][v])[w] = [uvw] = [u]([v][w])$. Furthermore, for every $[w] \in F$ we have $[e][w] = [w]$, so $[e]$ is an identity, and $[w^{-1}][w] = [e]$, so $[w]$ has an inverse, so F is a group. The fact that $\{[x] : x \in X\}$ generates F is immediate by definition. □

Definition (free group)**.** Let X be a set. Then the set of equivalence classes of words on X under the product defined in Proposition 7.1 is called the *free group* on X, denoted F_X. Given $r \in \mathbb{N}$, we write F_r for the free group on r symbols $x_1, \dots, x_r$.

You are invited to show in Exercise Sheet 7 Q2 that if X and Y are two sets of the same cardinality then $F_X \cong F_Y$, so F_r is well defined up to isomorphism.

We generally abuse notation slightly and drop the square brackets from elements of F_X, writing simply w instead of $[w]$. Thus, X may be viewed as a subset of F_X.

The key property of free groups is the following.

Theorem 7.2 (the universal property of free groups)**.** *Let X be a set, let G be a group, and let $\varphi : X \to G$. Then there is a unique homomorphism $\psi : F_X \to G$ such that $\psi(x) = \varphi(x)$ for every $x \in X$.*

You saw in Exercise Sheet 2 Q4 that, if X is a generating set of an arbitrary group Γ and G is some other group, then in general not every map $X \to G$ extends to a homomorphism $\Gamma \to G$, so the universal property is therefore a very special property of free groups. In fact, as you are invited to show in Exercise Sheet 7 Q1, free groups are the only groups to have this property, and indeed some authors take having the universal property as the definition of a free group.

Proof. Write W for the set of words on X, and define

$$\begin{array}{rccc} \psi_0 & : & W & \to & G \\ & & x_1^{\epsilon_1} \cdots x_k^{\epsilon_k} & \mapsto & \varphi(x_1)^{\epsilon_1} \cdots \varphi(x_k)^{\epsilon_k}, \end{array}$$

where each $x_i \in X$ and each $\epsilon_i \in \{1, -1\}$. If $v, w \in W$ are adjacent, say differing by a subword $x^\epsilon x^{-\epsilon}$, then $\psi_0(v) = \psi_0(w)$ since $\varphi(x)^\epsilon \varphi(x)^{-\epsilon} = e$ in G. It follows that ψ_0 is constant on equivalence classes, so we may define

$$\begin{array}{rccc} \psi & : & F_X & \to & G \\ & & [w] & \mapsto & \psi_0(w). \end{array}$$

The map ψ is a homomorphism since $\psi(vw) = \psi_0(vw) = \psi_0(v)\psi_0(w) = \psi(v)\psi(w)$ by definition of ψ_0. Since X generates F_X by Proposition 7.1, Lemma 2.7 implies that ψ is unique. □

Definition (free group… again)**.** If G is a group and $X \subseteq G$ then Theorem 7.2 shows that there is a unique homomorphism $\psi : F_X \to G$ such that $\psi(x) = x$ for every $x \in X$. If this is an isomorphism then G is said to be *free on* X.

We give a more concrete formulation of this definition in the next section.

Corollary 7.3. *Let G be a group with a generating set X. There is a unique surjective homomorphism $\psi : F_X \to G$ satisfying $\psi(x) = x$ for every $x \in X$. In particular, G is isomorphic to a quotient of F_X.*

Proof. The existence of ψ follows from applying Theorem 7.2 to the inclusion map $X \hookrightarrow G$, and then its surjectivity follows from Lemma 2.6. □

Note in particular that every group G admits at least one generating set, namely G itself, so Corollary 7.3 and the homomorphism theorem show that every group is isomorphic to a quotient of some free group.

7.2 Reduced words

Definition (reduced word)**.** A word on a set X is called *reduced* if it does not contain any subwords of the form xx^{-1} or $x^{-1}x$ with $x \in X$.

Proposition 7.4. *Let X be a set. Then each equivalence class of words on X contains precisely one reduced word.*

Proof. We first give a straightforward algorithm that, given an arbitrary word w, produces an equivalent reduced word. If w is not already reduced then it contains a subword of the form xx^{-1} or $x^{-1}x$, and on deleting that subword we obtain an equivalent word w_1 of shorter length. We apply the same procedure to w_1 to obtain a shorter equivalent word w_2, and so on. Since the length reduces at each step, this process must terminate, at which point we have a reduce word equivalent to w.

To see that no equivalence class contains more than one reduced word, write R for the set of reduced words and for each $x \in X$ and $\epsilon \in \{1, -1\}$, define a map $\sigma_{x,\epsilon} : R \to R$ by setting

$$\sigma_{x,\epsilon}(x_1^{\epsilon_1} \cdots x_k^{\epsilon_k}) = \begin{cases} x^{\epsilon} x_1^{\epsilon_1} \cdots x_k^{\epsilon_k} & \text{if } x_1^{\epsilon_1} \neq x^{-\epsilon} \\ x_2^{\epsilon_2} \cdots x_k^{\epsilon_k} & \text{if } x_1^{\epsilon_1} = x^{-\epsilon}. \end{cases}$$

Since $\sigma_{x,\epsilon}$ is invertible via $\sigma_{x,-\epsilon}$, it is bijective, which is to say that $\sigma_{x,\epsilon} \in \mathrm{Sym}(R)$. Theorem 7.2 therefore implies that there is a homomorphism $\psi : F_X \to \mathrm{Sym}(R)$ satisfying $\psi([x]) = \sigma_{x,1}$ for each $x \in X$. This implies in particular that

$$\psi([w])(e) = w \qquad \forall w \in R. \tag{7.1}$$

Now suppose that u, v are two reduced words lying in the same equivalence class. Then

$$\begin{aligned} u &= \psi([u])(e) && \text{(by (7.1))} \\ &= \psi([v])(e) && \text{(since } u \sim v\text{)} \\ &= v && \text{(by (7.1)),} \end{aligned}$$

as required. □

The reduced form of a word gives a canonical form in which we can express an arbitrary element of F_X. Moreover, given any word on X, the proof of Proposition 7.4 gives a straightforward algorithm to find this canonical form.

This also leads to a more technically convenient formulation of a group being free on a generating set, as follows.

Theorem 7.5. *Let G be a group generated by a set X. Then G is free on X if and only if no non-empty reduced word on X equals the identity in G.*

Proof. Let $\psi : F_X \to G$ be the unique surjective homomorphism such that $\psi([x]) = x$ for every $x \in X$ given by Corollary 7.3, and note that G is free on X if and only if ψ is injective. Note also that $\psi([x_1^{\epsilon_1} \cdots x_k^{\epsilon_k}]) = x_1^{\epsilon_1} \cdots x_k^{\epsilon_k}$ for all $x_i \in X$ and $\epsilon_i \in \{1, -1\}$ by the definition of the product on F_X and the definition of a homomorphism. Then

$$\begin{aligned} G \text{ is free on } X &\iff \psi \text{ is injective} \\ &\iff \psi([w]) \neq e \,\forall\, w \not\sim e \\ &\iff \psi([w]) \neq e \,\forall \text{ reduced } w \not\sim e \\ &\iff w \neq e \,\forall \text{ reduced } w \not\sim e. \end{aligned}$$

□

7.3 Group presentations

The aim of this section is to be able to understand the notation

$$D_{2n} = \langle r, s \mid r^n = s^2 = e,\ srs = r^{-1} \rangle, \tag{7.2}$$

which is called a *presentation* of D_{2n}. Intuitively, this captures the fact that D_{2n} is generated by the elements r, s, and that the equations in r and s that hold in D_{2n} are precisely those that are consequences of the equations $r^n = s^2 = e$ and $srs = r^{-1}$.

Although it is notationally convenient to write expressions like $r^n = s^2 = e$ on the right-hand side of a presentation as in (7.2), our upcoming definition of a presentation is cleaner if the right-hand side consists instead of words that are declared to equal the identity. Thus, we will sometimes rewrite (7.2) as

$$D_{2n} = \langle r, s \mid r^n, s^2, rsrs \rangle. \tag{7.3}$$

In practice we will use these two different versions of the notation interchangeably, which is a common abuse of notation in the literature.

In general, declaring that a certain word equals the identity also forces lots of other words to equal the identity. For example, stipulating that $r^n = e$ and $s^2 = e$ in (7.3) also forces $r^n s^2 = e$ and $g r^n g^{-1} = e$ for every group element g. This reflects the fact that the words in r, s equal to the identity in D_{2n} make up the kernel of the surjective homomorphism $F_{r,s} \to D_{2n}$ satisfying $r \mapsto r$ and $s \mapsto s$ given by Corollary 7.3, which is a normal subgroup. The idea, then, is that $\langle r, s \mid r^n, s^2, rsrs \rangle$ should be the quotient of $F_{r,s}$ by the 'smallest normal subgroup' containing $r^n, s^2, rsrs$.

Definition (normal subgroup generated by a set)**.** Let G be a group. Then given $R \subseteq G$, we define the *normal subgroup of G generated by R*, denoted $\langle R \rangle^G$, via

$$\langle R \rangle^G = \bigcap_{\substack{N \trianglelefteqslant G \\ R \subseteq N}} N.$$

Note that $\langle R \rangle^G \trianglelefteqslant G$ by Exercise Sheet 2 Q1(b), and that trivially $R \subseteq \langle R \rangle^G$.

Definition (group presentation)**.** Given a set X and a set R of words on X, we define $\langle X \mid R \rangle$ to be the quotient $F_X / \langle R \rangle^{F_X}$. We call $\langle X \mid R \rangle$ the *presentation* with *generators* X and *relators* R. Given a word w on X, we generally abuse notation and write w to mean the element $w\langle R \rangle^{F_X} \in \langle X \mid R \rangle$.

We also abuse notation and write $G = \langle X \mid R \rangle$ to mean that $G \cong \langle X \mid R \rangle$, or equivalently that there is a surjective homomorphism $\psi : F_X \to G$ with kernel $\langle R \rangle^{F_X}$. In this case, we say that $\langle X \mid R \rangle$ is a *presentation* of G, and call ψ the *homomorphism realising the presentation*. Again, we abuse notation in this case and write w to mean $\psi([w])$.

Theorem 7.6 (von Dyck's theorem)**.** *Let X be a set, and let R_1 and R_2 be sets of words on X such that $R_1 \subseteq R_2$. Suppose G_1 is a group with presentation $\langle X \mid R_1 \rangle$ realised by the homomorhism $\psi_1 : F_X \to G_1$, and G_2 is a group with presentation $\langle X \mid R_2 \rangle$ realised by the homomorhism $\psi_2 : F_X \to G_2$. Then we may define a surjective homomorphism $G_1 \to G_2$ by $\psi_1(w) \mapsto \psi_2(w)$.*

Proof. The fact that $R_1 \subseteq R_2$ implies that $\ker \psi_1 \leqslant \ker \psi_2$. The map in question is well defined because given $w, v \in F_X$ we have $\psi_1(w) = \psi_1(v) \implies w^{-1}v \in \ker \psi_1 = \langle R_1 \rangle^{F_X} \leqslant \langle R_2 \rangle^{F_X} = \ker \psi_2 \implies \psi_2(w) = \psi_2(v)$. It is trivially a homomorphism, and surjective because ψ_2 is surjective. □

Example. In Exercise Sheet 7 Q8 you are invited to verify that (7.2) really is a presentation of D_{2n}. To set you on the way, let us see that $D_\infty = \langle r, s \mid s^2 = e,\ srs = r^{-1} \rangle$. Indeed, let

$$G = \langle r, s \mid s^2 = e,\ srs = r^{-1} \rangle, \tag{7.4}$$

and note that

$$D_\infty = \langle r, s \mid R \rangle,$$

where R is the set of *all* relations in r, s that hold in D_∞. Since the relations of (7.4) hold in D_∞, and so are a subset of R, von Dyck's theorem implies that there is a surjective homomorphism $\varphi : G \to D_\infty$ such that $\varphi(r) = r$ and $\varphi(s) = s$. Using the relations of (7.4), we can write an arbitrary element $g \in G$ in the form $g = r^m s^i$ with $m \in \mathbb{Z}$ and $i \in \{0, 1\}$, and we then have $\varphi(g) = r^m s^i \in D_\infty$. In particular, we have $\varphi(g) = e \implies m = i = 0 \implies g = e$, so that φ is injective, hence bijective, and $G \cong D_\infty$.

Example. Let us now see that $\langle a, b \mid a^2 = b^2 = e \rangle$ is also a presentation of D_∞. Indeed, let G be as above and set

$$H = \langle a, b \mid a^2 = b^2 = e \rangle,$$

and note that by von Dyck's theorem there exist surjective homomorphisms $\varphi_1 : G \to H$ such that $\varphi_1(r) = ab$ and $\varphi_1(s) = b$, and $\varphi_2 : H \to G$ such that $\varphi_2(a) = rs$ and $\varphi_2(b) = s$. These are mutually inverse (by Lemma 2.7 it is enough to check that $\varphi_1\varphi_2(a) = a$, $\varphi_1\varphi_2(b) = b$, $\varphi_2\varphi_1(r) = r$, $\varphi_2\varphi_1(s) = s$), hence invertible, hence isomorphisms, and so $G \cong H$.

Warning. Although the presentations of D_{2n} and D_∞ allow us to 'evaluate' every expression in the generators, in the sense that we can write every word in the standard form $r^n s^i$ with $n \in \mathbb{Z}$ and $i \in \{1, -1\}$, in general this is not always possible. Indeed, the famous *Novikov–Boone theorem* states that there exists a presentation $\langle X \mid R \rangle$ with X and R both finite such that there is no algorithm to determine whether a given word on X evaluates to the identity. There is also no algorithm to determine whether an arbitrary presentation $\langle X \mid R \rangle$ results in the trivial group.