Section 6

Soluble Groups

We have already met the concept of a composition series for a group. In the next section we shall consider groups whose composition factors are all abelian. We can think of this as the class of groups we can build using only abelian groups.

To give a general description of these groups we need the following concept.

Definition 6.1 Let G be a group and $x, y \in G$. The *commutator* of x and y is the element
$$[x, y] = x^{-1}y^{-1}xy.$$

Note that the following equations hold immediately:

$$[x, y] = x^{-1}x^{y}$$
$$[x, y] = (y^{-1})^{x}y$$

and

$$xy = yx[x, y]. \tag{6.1}$$

The latter tells us that the commutator essentially measures by how much x and y fail to commute.

Lemma 6.2 *Let G and H be groups, let $\phi \colon G \to H$ be a homomorphism and let $x, y, z \in G$. Then*

(i) $[x, y]^{-1} = [y, x]$;

(ii) $[x, y]\phi = [x\phi, y\phi]$;

(iii) $[x, yz] = [x, z]\,[x, y]^{z}$;

(iv) $[xy, z] = [x, z]^{y}\,[y, z]$.

PROOF: (i) $[x, y]^{-1} = (x^{-1}y^{-1}xy)^{-1} = y^{-1}x^{-1}yx = [y, x]$.

(ii) $[x, y]\phi = (x^{-1}y^{-1}xy)\phi = (x\phi)^{-1}(y\phi)^{-1}(x\phi)(y\phi) = [x\phi, y\phi]$.

(iii) For this and part (iv), we shall rely on Equation (6.1) and view it as telling us how to exchange group elements at the expense of introducing commutators. (This is known as 'collection'.) So

$$xyz = yzx[x, yz]$$

but if we collect one term at a time we obtain

$$\begin{aligned} xyz &= yx[x, y]z \\ &= yxz[x, y]^z \\ &= yzx[x, z]\,[x, y]^z. \end{aligned}$$

Hence

$$yzx[x, yz] = yzx[x, z]\,[x, y]^z,$$

so

$$[x, yz] = [x, z]\,[x, y]^z.$$

(iv)

$$xyz = zxy[xy, z]$$

and

$$\begin{aligned} xyz &= xzy[y, z] \\ &= zx[x, z]y[y, z] \\ &= zxy[x, z]^y[y, z]. \end{aligned}$$

Comparing we deduce

$$[xy, z] = [x, z]^y[y, z].$$

□

Both parts (iii) and (iv) can be proved by a more simple-minded expansion of the terms on both sides, but I believe more can be learnt and understood via the collection process.

Definition 6.3 Let G be a group. The *derived subgroup* (or *commutator subgroup*) G' of G is the subgroup generated by all commutators of elements from G:

$$G' = \langle\, [x, y] \mid x, y \in G \,\rangle.$$

Part (i) of Lemma 6.2 tells us that the inverse of a commutator is again a commutator, but we have no information about products of commutators. Consequently, a typical element of G' has the form

$$[x_1, y_1]\,[x_2, y_2] \dots [x_n, y_n]$$

where $x_i, y_i \in G$ for each i.

Iterating this construction yields the derived series:

Definition 6.4 The *derived series* $(G^{(i)})$ (for $i \geqslant 0$) is the chain of subgroups of the group G defined by

$$G^{(0)} = G$$

and

$$G^{(i+1)} = (G^{(i)})' \qquad \text{for } i \geqslant 0.$$

So $G^{(1)} = G'$, $G^{(2)} = (G')' = G''$, etc. We then have a chain of subgroups

$$G = G^{(0)} \geqslant G^{(1)} \geqslant G^{(2)} \geqslant \cdots .$$

We shall see later that this is indeed a series in the sense of Definition 4.1 (in that each term is normal in the previous). Indeed far more is true as we shall see.

Definition 6.5 A group G is called *soluble* (*solvable* in the U.S.) if $G^{(d)} = \mathbf{1}$ for some d. The least such d is called the *derived length* of G.

Since when forming the derived series, we take the derived subgroup of the previous term at each stage, once we have a repetition then the series becomes constant. Thus if G is a soluble group of derived length d, its derived series has the form

$$G = G^{(0)} > G^{(1)} > G^{(2)} > \cdots > G^{(d)} = \mathbf{1}.$$

We seek to understand the properties a soluble group really has and to produce equivalent formulations so that examples can be more easily described. Accordingly, we being by establishing basic properties of the derived subgroup and the derived series.

Lemma 6.6 (i) *If H is a subgroup of G, then $H' \leqslant G'$.*

(ii) *If $\phi \colon G \to K$ is a homomorphism, then $G'\phi \leqslant K'$.*

(iii) *If $\phi \colon G \to K$ is a surjective homomorphism, then $G'\phi = K'$.*

Proof: (i) If $x, y \in H$, then $[x, y]$ is a commutator of elements of G so belongs to the derived subgroup of G:

$$[x, y] \in G' \qquad \text{for all } x, y \in H.$$

Therefore

$$\langle\, [x, y] \mid x, y \in H \,\rangle \leqslant G',$$

so $H' \leqslant G'$.

(ii) If $x, y \in G$, then $[x, y]\phi = [x\phi, y\phi] \in K'$. Since K' is closed under products, it follows that any product of commutators in G is mapped into K' by ϕ. Thus $G'\phi \leqslant K'$.

(iii) Let $a, b \in K$. Since ϕ is surjective, there exists $x, y \in G$ such that $a = x\phi$ and $b = y\phi$. Thus

$$[a, b] = [x\phi, y\phi] = [x, y]\phi \in G'\phi.$$

Thus

$$[a, b] \in G'\phi \qquad \text{for all } a, b \in K.$$

This forces $K' \leqslant G'\phi$. Using (ii) gives $K' = G'\phi$, as required. □

Lemma 6.7 *Subgroups and homomorphic images of soluble groups are themselves soluble.*

Proof: Let G be a soluble group and H be a subgroup of G.

Claim: $H^{(i)} \leqslant G^{(i)}$ for all i.

We prove the claim by induction on i. The case $i = 0$ is the inclusion $H \leqslant G$ which holds by assumption.

Now suppose $H^{(i)} \leqslant G^{(i)}$. Apply Lemma 6.6(i) to give

$$(H^{(i)})' \leqslant (G^{(i)})';$$

that is,

$$H^{(i+1)} \leqslant G^{(i+1)}.$$

This completes the induction.

Now since G is soluble, $G^{(d)} = \mathbf{1}$ for some d. Therefore, as $H^{(d)} \leqslant G^{(d)}$, we have $H^{(d)} = \mathbf{1}$ and so we deduce that H is soluble.

Now let K be a homomorphic image of G. Thus there exists a surjective homomorphism $\phi \colon G \to K$.

Claim: $K^{(i)} = G^{(i)}\phi$ for all i.

We prove the claim by induction on i. The case $i = 0$ is the equation $K = G\phi$ which holds by assumption.

Now suppose $K^{(i)} = G^{(i)}\phi$. Thus ϕ induces a surjective homomorphism $G^{(i)} \to K^{(i)}$ and Lemma 6.6(iii) gives

$$(K^{(i)})' = (G^{(i)})'\phi;$$

that is,

$$K^{(i+1)} = G^{(i+1)}\phi.$$

This completes the induction.

Now as G is soluble we have $G^{(d)} = \mathbf{1}$ and thus

$$K^{(d)} = G^{(d)}\phi = \mathbf{1}\phi = \mathbf{1}.$$

Hence K is soluble. □

It follows that quotient groups (which are the same as homomorphic images) of soluble groups are themselves soluble. There is a rather strong converse to the above lemma as well.

Proposition 6.8 *Let G be a group and N be a normal subgroup of G such that both G/N and N are soluble. Then G is soluble.*

PROOF: Let $\pi\colon G \to G/N$ be the natural map. By assumption $(G/N)^{(d)} = \mathbf{1}$ and $N^{(e)} = \mathbf{1}$ for some d and e. Now, by the second claim in Lemma 6.7, we have

$$G^{(d)}\pi = (G/N)^{(d)} = \mathbf{1}.$$

Hence

$$G^{(d)} \leqslant \ker\pi = N.$$

Therefore, by the first claim in Lemma 6.7,

$$(G^{(d)})^{(e)} \leqslant N^{(e)} = \mathbf{1};$$

that is,

$$G^{(d+e)} = \mathbf{1}.$$

Thus G is soluble. $\square$

We have observed that if $\phi\colon G \to K$ is a surjective homomorphism then $G'\phi = K'$. In particular, if ϕ is an automorphism of G (that is, an isomorphism $G \to G$), then $G'\phi = G'$. We give the following special name to subgroups satisfying this property.

Definition 6.9 A subgroup H of a group G is said to be a *characteristic subgroup* of G if $x\phi \in H$ for all $x \in H$ and all automorphisms ϕ of G.

The definition requires that $H\phi \leqslant H$ for all automorphisms ϕ of G. But then we have $H\phi^{-1} \leqslant H$ and applying ϕ then yields $H \leqslant H\phi$. Thus H is a characteristic subgroup if and only if $H\phi = H$ for all automorphisms ϕ of G.

The notation for being a characteristic subgroup is less consistently developed than for, say, being a normal subgroup. I shall write

$$H \operatorname{char} G$$

to indicate that H is a characteristic subgroup of G.

Our observation above then is that

$$G' \operatorname{char} G$$

for all groups G and we shall soon see that all terms in the derived series are also characteristic.

Lemma 6.10 *Let G be a group.*

(i) *If $H \operatorname{char} G$, then $H \trianglelefteqslant G$.*

(ii) *If $K \operatorname{char} H$ and $H \operatorname{char} G$, then $K \operatorname{char} G$.*

(iii) *If $K \operatorname{char} H$ and $H \trianglelefteqslant G$, then $K \trianglelefteqslant G$.*

Thus there is considerable difference between characteristic subgroups and normal subgroups. For example, note that in general

- $K \trianglelefteqslant H \trianglelefteqslant G$ does not imply $K \trianglelefteqslant G$.
- If $\phi \colon G \to K$ is a homomorphism and $H \operatorname{char} G$, then it does not follow necessarily that $H\phi \operatorname{char} G\phi$. (Consequently the Correspondence Theorem does not work well with characteristic subgroups.)
- If $H \leqslant L \leqslant G$ and $H \operatorname{char} G$, then it does not necessarily follow that $H \operatorname{char} L$.

Proof of Lemma 6.10: (i) If $x \in G$, then $\tau_x \colon g \mapsto g^x$ is an automorphism of G. Hence if $H \operatorname{char} G$, then

$$H^x = H\tau_x = H \qquad \text{for all } x \in G,$$

so $H \trianglelefteqslant G$.

(ii) Let ϕ be an automorphism of G. Then $H\phi = H$ (as $H \operatorname{char} G$). Hence the restriction $\phi|_H$ of ϕ to H is an automorphism of H and we deduce

$$x\phi \in K \qquad \text{for all } x \in K$$

(since this is the effect that the restriction $\phi|_H$ has when applied to elements of K). Thus $K \operatorname{char} G$.

(iii) Let $x \in G$. Then $H^x = H$ (as $H \trianglelefteqslant G$) and therefore $\tau_x \colon g \mapsto g^x$ (for $g \in H$) is a bijective homomorphism $H \to H$; that is, τ_x is an automorphism of H. Since $K \operatorname{char} H$, we deduce that $K^x = K\tau_x = K$. Thus $K \trianglelefteqslant G$. $\square$

We have seen that $G' \operatorname{char} G$ holds. Recall the definition of the derived series:

$$G^{(0)} = G, \qquad G^{(i+1)} = (G^{(i)})' \quad \text{for } i \geqslant 0.$$

Therefore

$$G^{(i)} \operatorname{char} G^{(i-1)} \operatorname{char} G^{(i-2)} \operatorname{char} \cdots \operatorname{char} G^{(1)} \operatorname{char} G^{(0)} = G.$$

Applying Lemma 6.10(ii) we see that each $G^{(i)}$ is a characteristic subgroup (and hence a normal subgroup) of G for each i.

Proposition 6.11 *The derived series*

$$G = G^{(0)} \geqslant G^{(1)} \geqslant G^{(2)} \geqslant \cdots$$

is a chain of subgroups each of which is a characteristic subgroup of G and hence each of which is a normal subgroup of G. □

In particular, if G is a soluble group of derived length d then we have

$$G = G^{(0)} > G^{(1)} > \cdots > G^{(d)} = \mathbf{1}$$

and this is a normal series (each term is normal in G). In particular, we can consider the factors

$$G^{(0)}/G^{(1)},\ G^{(1)}/G^{(2)}, \ldots, G^{(d-1)}/G^{(d)};$$

i.e., the quotient groups $G^{(i)}/(G^{(i)})'$ for $i = 0, 1, \ldots, d-1$.

We now seek to elucidate information about these factors.

Lemma 6.12 *Let G be a group and N be a normal subgroup of G. Then G/N is abelian if and only if $G' \leqslant N$.*

In particular, G/G' is an abelian group and it is the largest quotient group of G which is abelian. We often call G/G' the *abelianisation* of G.

PROOF: Suppose G/N is abelian. Then

$$Nx \cdot Ny = Ny \cdot Nx \qquad \text{for all } x, y \in G,$$

so

$$N[x,y] = (Nx)^{-1}(Ny)^{-1}(Nx)(Ny) = N1 \qquad \text{for all } x, y \in G.$$

Thus $[x,y] \in N$ for all $x, y \in G$ and we obtain $G' \leqslant N$.

Conversely if $G' \leqslant N$, then $[x,y] \in N$ for all $x, y \in G$ and reversing the above steps shows that G/N is abelian. □

In particular, the factors occurring in the derived series are all abelian. So if G is a soluble group, it has the derived series as a normal series with all factors abelian. The following result strengthens this and puts it into context.

Theorem 6.13 *Let G be a group. The following conditions are equivalent:*

(i) *G is soluble;*

(ii) *G has a chain of subgroups*

$$G = G_0 \geqslant G_1 \geqslant G_2 \geqslant \cdots \geqslant G_n = \mathbf{1}$$

such that G_i is a normal subgroup of G and G_{i-1}/G_i is abelian for $i = 1, 2, \ldots, n$;

(iii) *G has a chain of subgroups*

$$G = G_0 \geqslant G_1 \geqslant G_2 \geqslant \cdots \geqslant G_n = \mathbf{1}$$

such that G_i is a normal subgroup of G_{i-1} and G_{i-1}/G_i is abelian for $i = 1, 2, \ldots, n$.

We describe Condition (ii) as saying that G has a normal series with abelian factors, while (iii) says that G has a series (or subnormal series) with abelian factors.

PROOF: (i) $\Rightarrow$ (ii): The derived series is such a chain of subgroups.

(ii) $\Rightarrow$ (iii): Immediate: If $G_i \trianglelefteqslant G$ for $G_i \leqslant G_{i-1} \leqslant G$, then $G_i \trianglelefteqslant G_{i-1}$.

(iii) $\Rightarrow$ (i): Suppose

$$G = G_0 \geqslant G_1 \geqslant G_2 \geqslant \ldots \geqslant G_n = \mathbf{1}$$

is a series where G_{i-1}/G_i is abelian for all i.

Claim: $G^{(i)} \leqslant G_i$ for all i.

We prove the claim by induction on i. Since $G^{(0)} = G = G_0$, the claim holds for $i = 0$.

Suppose $G^{(i)} \leqslant G_i$. Now $G_{i+1} \trianglelefteqslant G_i$ and G_i/G_{i+1} is abelian. Hence $(G_i)' \leqslant G_{i+1}$ by Lemma 6.12. Further, by Lemma 6.6(i), $(G^{(i)})' \leqslant (G_i)'$. Hence

$$G^{(i+1)} = (G^{(i)})' \leqslant (G_i)' \leqslant G_{i+1}.$$

Hence by induction $G^{(n)} \leqslant G_n = \mathbf{1}$, so $G^{(n)} = \mathbf{1}$ and G is soluble. $\square$

We now have a characterisation that a group is soluble if and only if it has a series with abelian factors. We shall obtain a further such equivalence by linking solubility to composition series. First, however, we note the following:

Example 6.14 An abelian group G is soluble. Indeed in an abelian group G

$$[x, y] = x^{-1}y^{-1}xy = 1 \qquad \text{for all } x \text{ and } y,$$

so $G' = \mathbf{1}$. (Of course, the condition $G' = \mathbf{1}$ is equivalent to G being abelian.)

In particular, the infinite cyclic group is soluble, though we know (Example 4.5) that this group does not have a composition series. Accordingly we cannot hope for composition series to give us complete information about soluble groups.

It turns out that as long as we avoid the infinite soluble groups, composition series do tell us whether or not our group is soluble.

Theorem 6.15 *Let G be a group. Then the following conditions are equivalent:*

(i) *G is a finite soluble group;*

(ii) *G has a composition series with all composition factors cyclic of prime order.*

Recall that the abelian simple groups are precisely the cyclic groups of (various) prime orders. Thus part (ii) describes the groups with abelian composition factors.

PROOF: (ii) $\Rightarrow$ (i): Let

$$G = G_0 > G_1 > G_2 > \cdots > G_n = \mathbf{1}$$

be a composition series for G and suppose that all the factors are cyclic. Then $G_i \trianglelefteqslant G_{i-1}$ and G_{i-1}/G_i is abelian for each i. Thus this is a chain of subgroups as in part (iii) of Theorem 6.13 and therefore G is soluble by that result. Further

$$|G| = |G_0/G_1| \cdot |G_1/G_2| \cdot \ldots \cdot |G_{n-1}/G_n|,$$

a product of finitely many primes, so G is finite.

(i) $\Rightarrow$ (ii): Let G be a finite soluble group. Then by Theorem 6.13, G possesses a chain of subgroups

$$G = G_0 > G_1 > G_2 > \cdots > G_n = \mathbf{1} \tag{6.2}$$

such that $G_i \trianglelefteqslant G_{i-1}$ and G_{i-1}/G_i is abelian for all i. Note that G can only have at most finitely many such series. Thus we may assume that (6.2) is the longest series for G with abelian factors. Such a series must then be a composition series: for if some G_{i-1}/G_i is not simple, then there exists $N \trianglelefteqslant G_{i-1}$ with $G_i < N < G_{i-1}$. We then obtain a series

$$G = G_0 > \cdots > G_{i-1} > N > G_i > \cdots > G_n = \mathbf{1}$$

which is longer than (6.2) and the new factors occurring here satisfy

$$N/G_i \leqslant G_{i-1}/G_i \qquad \text{and} \qquad G_{i-1}/N \cong \frac{G_{i-1}/G_i}{N/G_i}$$

(by the Third Isomorphism Theorem). Since G_{i-1}/G_i is abelian, we see that N/G_i and G_{i-1}/N are abelian. This contradicts the assumption that (6.2) is the longest series with abelian factors.

We now deduce that (6.2) is indeed a composition series and hence the composition factors of G are abelian. Since the only abelian simple groups are cyclic of prime order, we deduce that all the composition factors of G are cyclic of prime order (for various primes). $\square$

So far this section has been rather devoid of examples. We have observed that all abelian groups are soluble, but this is not particularly far reaching. On the other hand, our two characterisation theorems, Theorems 6.13 and 6.15 do far better for helping us recognise (finite) soluble groups. These theorems tell us that a soluble group is one that is built from abelian groups.

Example 6.16 The symmetric group S_4 of degree 4 is soluble. Indeed in Example 4.3 we observed that

$$S_4 > A_4 > V_4 > \langle (1\ 2)(3\ 4) \rangle > \mathbf{1}$$

is a composition series for S_4 and the composition factors are C_2, C_3, C_2 and C_2. Hence S_4 is soluble by Theorem 6.15.

Example 6.17 The dihedral group D_{2n} of order $2n$ is soluble. Indeed D_{2n} contains an element α of order n, so $\langle \alpha \rangle$ has index 2 so is normal. Thus

$$D_{2n} > \langle \alpha \rangle > \mathbf{1}$$

is a series for D_{2n} with both factors cyclic. Hence D_{2n} is soluble by Theorem 6.13.

Example 6.18 The symmetric group S_n of degree n is insoluble if $n \geqslant 5$. Indeed we know that A_n is non-abelian simple group, so is insoluble by Theorem 6.15. As a subgroup of a soluble group is always soluble, it must be the case that S_n is insoluble also.

Careful analysis of the examples in Section 3 shows that the groups we considered that were not simple in 3.6–3.9 are also soluble groups.

Finite soluble groups

For the rest of this section we shall work only with finite groups. Our goal is to study finite soluble groups in much greater detail and in particular prove Hall's Theorem concerning finite soluble groups. We shall prove these by induction on the group order. The method will involve working with quotients and so we begin by studying normal subgroups which are as small as possible.

Minimal normal subgroups

For the moment we shall work with arbitrary finite groups without assuming that they are also soluble. Solubility will return in due course.

Definition 6.19 Let G be a finite group. A *minimal normal subgroup* of G is a non-trivial normal subgroup of G which has no non-trivial proper subgroup which is also normal in G.

Thus M is a minimal normal subgroup of G if

(i) $\mathbf{1} < M \trianglelefteqslant G$;

(ii) if $\mathbf{1} \leqslant N \leqslant M$ and $N \trianglelefteqslant G$, then either $N = \mathbf{1}$ or $N = M$.

Note that, apart from the trivial group, all finite groups have minimal normal subgroups. To see this, we start with the group G itself. If this isn't a minimal normal subgroup, then there is a proper subgroup below it which is normal. If this isn't minimal, then there is a proper subgroup below it which is normal in G. Repeating this process must eventually stop (since G is finite) and yield a minimal normal subgroup.

We shall prove the following description of minimal normal subgroups.

Theorem 6.20 *A minimal normal subgroup of a finite group G is a direct product of isomorphic simple groups.*

In the case of a minimal normal subgroup of a finite soluble group, these simple groups will be cyclic of prime order. We shall work towards the proof of this theorem next. First we make the following definition.

Definition 6.21 A non-trivial group G is called *characteristically simple* if the only characteristic subgroups it has are $\mathbf{1}$ and G.

(Recall, from Definition 6.9, that a characteristic subgroup of G is a subgroup which is closed under applying all automorphisms of G.)

Lemma 6.22 *A minimal normal subgroup of a group is characteristically simple.*

PROOF: Let M be a minimal normal subgroup of the group G. Let K be a characteristic subgroup of M. Then

$$K \operatorname{char} M \trianglelefteqslant G,$$

so $K \trianglelefteqslant G$ by Lemma 6.10(iii). Thus minimality of M forces $K = \mathbf{1}$ or $K = M$. Hence M is indeed characteristically simple. □

Theorem 6.20 then follows immediately from the following result. (The advantage of proving Theorem 6.23 over a direct attempt on Theorem 6.20 is that we can concentrate only on the characteristically simple group rather than having to juggle both the minimal normal subgroup and its embedding in the larger group.)

Theorem 6.23 *A characteristically simple finite group is a direct product of isomorphic simple groups.*

PROOF: Let G be a finite group which is characteristically simple. Let S be a minimal normal subgroup of G. (So $S \neq \mathbf{1}$. It is possible that $S = G$.) Consider the following set

$$\mathscr{D} = \{\, N \trianglelefteq G \mid N = S_1 \times S_2 \times \cdots \times S_k \text{ where each } S_i \text{ is a minimal normal subgroup of } G \text{ isomorphic to } S \,\}.$$

(Recall what we mean by the direct product here: it is an internal direct product, so we need $S_i \cap S_1 \dots S_{i-1} S_{i+1} \dots S_k = \mathbf{1}$ for each i, as well as $N = S_1 S_2 \dots S_k$. We already assume $S_i \trianglelefteq G$, so the requirement $S_i \trianglelefteq N$ comes for free.)

Note that $S \in \mathscr{D}$, so $\mathscr{D}$ certainly contains non-trivial members. Choose $N \in \mathscr{D}$ of largest possible order.

Claim: $N = G$.

Suppose our maximal member N of $\mathscr{D}$ is not equal to G. Then as G is characteristically simple, N cannot be a characteristic subgroup of G. Hence there exists an automorphism ϕ of G such that

$$N\phi \nleqslant N.$$

Let $N = S_1 \times S_2 \times \cdots \times S_k$ where each S_i is a minimal normal subgroup of G isomorphic to S. Therefore there exists i such that

$$S_i\phi \nleqslant N.$$

Now ϕ is an automorphism of G, so $S_i\phi$ is a minimal normal subgroup of G. Now $N \cap S_i\phi \trianglelefteq G$ and $N \cap S_i\phi$ is properly contained in $S_i\phi$ (as $S_i\phi \nleqslant N$). Therefore, by minimality, $N \cap S_i\phi = \mathbf{1}$. It follows that

$$N \cdot S_i\phi = N \times S_i\phi = S_1 \times S_2 \times \cdots \times S_k \times S_i\phi$$

and

$$N \cdot S_i\phi \trianglelefteq G.$$

This shows that $N \cdot S_i\phi \in \mathscr{D}$. This contradicts N being a maximal member of $\mathscr{D}$.

Therefore

$$G = N = S_1 \times S_2 \times \cdots \times S_k,$$

where each S_i is a minimal normal subgroup of G isomorphic to our original minimal normal subgroup S.

It remains to check that S is simple. If $J \trianglelefteqslant S_1$, then

$$J \trianglelefteqslant S_1 \times S_2 \times \cdots \times S_k = G.$$

Therefore, as S_1 is a minimal normal subgroup of G, we must have $J = \mathbf{1}$ or $J = S_1$. Hence S_1 (and accordingly S) is simple.

We have shown that, indeed, G is a direct product of isomorphic simple groups. $\square$

We have now established Theorems 6.20 and 6.23 in a general setting. We are, however, only interested in soluble groups in the current section and Theorem 6.15 tells us that the only simple groups which can be occurring in this world are cyclic groups of prime order. Thus in a finite soluble group, a minimal normal subgroup is a direct product of cyclic groups of order p (for some prime p). We give a special name to these groups:

Definition 6.24 Suppose that p is a prime number. An *elementary abelian p-group* G is an abelian group such that

$$x^p = 1 \qquad \text{for all } x \in G.$$

Recall that a finite abelian group is a direct product of cyclic groups. It follows that a finite group is an elementary abelian p-group if and only if

$$G \cong \underbrace{C_p \times C_p \times \cdots \times C_p}_{d \text{ times}}$$

for some d.

Putting together Theorem 6.15 and Theorem 6.20 gives:

Theorem 6.25 *A minimal normal subgroup of a finite soluble group is an elementary abelian p-group for some prime number p.* $\square$

This result will be used in the induction step of our proof of Hall's Theorem. We now move on to describe the type of subgroup this theorem concerns.

Hall subgroups

Definition 6.26 Let π be a set of prime numbers and let G be a finite group. A *Hall π-subgroup* of G is a subgroup H of G such that $|H|$ is a product involving only the primes in π and $|G : H|$ is a product involving only primes not in π.

If p is a prime number, then a Hall $\{p\}$-subgroup is precisely the same thing as a Sylow p-subgroup.

Example 6.27 Consider the alternating group A_5 of degree 5. Here

$$|A_5| = 60 = 2^2 \cdot 3 \cdot 5.$$

So a Hall $\{2,3\}$-subgroup of A_5 has order 12. We already know of a subgroup with this order: thus, A_4 is a Hall $\{2,3\}$-subgroup of A_5.

A Hall $\{2,5\}$-subgroup of A_5 would have order 20 and index 3, while a Hall $\{3,5\}$-subgroup of A_5 would have order 15 and index 4. If H were one of these, then we could let A_5 act on the cosets of H and obtain a homomorphism $\rho\colon A_5 \to S_r$ (where $r = 3$ or 4). Here $\ker\rho \neq \mathbf{1}$ and $\ker\rho \neq A_5$ (as $\ker\rho \leqslant H$), which would contradict the fact that A_5 is simple.

Hence A_5 does not have any Hall π-subgroups for $\pi = \{2,5\}$ or $\pi = \{3,5\}$.

So in insoluble groups, some Hall π-subgroups might exist, while others might not (in fact, it is a theorem that some definitely do not!). This is in stark contrast to soluble groups where we shall observe that Hall π-subgroups always do exist:

Theorem 6.28 (P. Hall, 1928) *Let G be a finite soluble group and let π be a set of prime numbers. Then*

(i) *G has a Hall π-subgroup;*

(ii) *any two Hall π-subgroups of G are conjugate;*

(iii) *any π-subgroup of G is contained in a Hall π-subgroup.*

A subgroup of G is called a *π-subgroup* if its order is a product involving only the primes in π. There is a clear analogy between this theorem of Hall and Sylow's Theorem (Theorem 3.4).

Hall subgroups and this theorem are named after Philip Hall (1904–1982), a British mathematician who did groundbreaking research into the theory of finite and infinite groups in the early and mid-parts of the twentieth century.

A number of tools are needed in the course of this theorem. The one remaining fact that has not already been established is the following result. This first appeared in the context of nilpotent groups, and we shall use it in that context in the next section, but it is also needed for the hardest part of the proof of Hall's Theorem.

Lemma 6.29 (Frattini Argument) *Let G be a finite group, N be a normal subgroup of G and P be a Sylow p-subgroup of N. Then*

$$G = \mathrm{N}_G(P)\, N.$$

The name of the lemma suggests (correctly) that it is the method of proof that is actually most important here. The idea can be adapted to many situations and turns out to be very useful.

Proof: Let $x \in G$. Since $N \trianglelefteqslant G$, we have

$$P^x \leqslant N^x = N,$$

so P^x is a Sylow p-subgroup of N. Sylow's Theorem then tells us that P^x and P are conjugate in N:

$$P^x = P^n \qquad \text{for some } n \in N.$$

Therefore

$$P^{xn^{-1}} = P,$$

so $y = xn^{-1} \in \mathrm{N}_G(P)$. Hence $x = yn \in \mathrm{N}_G(P)N$. The reverse inclusion is obvious, so

$$G = \mathrm{N}_G(P)N.$$

$\square$

Proof of Theorem 6.28: Our strategy is to prove part (i) and deduce part (iii) by showing that a π-subgroup is contained in a conjugate of the Hall π-subgroup found already. We shall then deduce part (ii) at the end. The argument is by induction on the group order dividing into several cases. Since the same cases arise in the proof of both (i) and (iii), we shall actually step through doing both parts together. (A proof of (i) can be extracted by just deleting the second half of each case.)

Thus we shall prove the following:

- G has a Hall π-subgroup H;
- if L is a π-subgroup of G then L is contained in some conjugate of H.

We prove these statements by induction on the order of G. Both are trivial if $|G| = 1$. Assume then that $|G| > 1$ and that these statements hold for soluble groups of order smaller than G. Write $|G| = mn$ where m is a product involving primes in π and n is a product involving primes not in π. (A Hall π-subgroup of G is then a subgroup of order m.) We can assume that $m > 1$ since otherwise the statements are trivially true.

Let M be a minimal normal subgroup of G. By Theorem 6.25, M is elementary abelian. We consider two cases according to the prime dividing the order of M.

Case 1: M is an elementary abelian p-group where $p \in \pi$. Write $|M| = p^\alpha$.

Then

$$|G/M| = mn/p^\alpha = m_1 n,$$

where $m = m_1 p^\alpha$. By induction, the above statements hold for G/M. The Correspondence Theorem tells us that a Hall π-subgroup of G/M has the form H/M where H is a subgroup of G containing M. Then

$$|H/M| = m_1$$

so

$$|H| = m_1 |M| = m_1 p^\alpha = m.$$

Hence H is a Hall π-subgroup of G.

Now let L be any π-subgroup of G. Then image LM/M ($\cong L/(L \cap M)$) of L in the quotient group is a π-subgroup of G/M. Hence, by induction, some conjugate of H/M contains LM/M, say

$$LM/M \leqslant (H/M)^{Mx} = H^x/M$$

where $x \in G$. Thus

$$L \leqslant LM \leqslant H^x.$$

This completes Case 1.

Case 2: No minimal normal subgroup of G is an elementary abelian p-group with $p \in \pi$. In particular, our minimal normal subgroup M of G satisfies $|M| = q^\beta$ where $q \notin \pi$.

Then

$$|G/M| = mn/q^\beta = mn_1$$

where $n = n_1 q^\beta$. We now further subdivide according to n_1.

Subcase 2A: $n_1 \neq 1$.

By induction, G/M has a Hall π-subgroup, which has the form K/M where K is a subgroup of G containing M and

$$|K/M| = m.$$

Then

$$|K| = m|M| = mq^\beta = mn/n_1 < mn.$$

We shall further apply induction to K. This has smaller order than G and hence possesses a Hall π-subgroup. Let H be a Hall π-subgroup of K. Then $|H| = m$, so H is also a Hall π-subgroup of G.

Now let L be a π-subgroup of G. Now the image LM/M of L in the quotient group is a π-subgroup of G/M. Hence, by induction, LM/M is contained in some conjugate of K/M; say

$$LM/M \leqslant (K/M)^{Mx} = K^x/M$$

where $x \in G$. Hence $L \leqslant LM \leqslant K^x$, so $L^{x^{-1}} \leqslant K$. Then $L^{x^{-1}}$ is a π-subgroup of K and by induction (again) we deduce $L^{x^{-1}} \leqslant H^y$ for some $y \in K$. Hence

$$L \leqslant H^{yx}$$

and we have completed Subcase 2A.

Subcase 2B: $n_1 = 1$, so $|G| = mq^\beta$.

Note also that the general assumption of Case 2 still applies: G has no minimal normal subgroup which is elementary abelian-p for $p \in \pi$.

Now $|G/M| = m > 1$. Let N/M be a minimal normal subgroup of G/M. Then N/M is an elementary abelian p-group for some $p \in \pi$ (since m is a product involving only primes in π), say $|N/M| = p^\alpha$. Then $N \trianglelefteqslant G$ and

$$|N| = p^\alpha q^\beta.$$

Let P be a Sylow p-subgroup of N. Let us now apply the Frattini Argument (Lemma 6.29):

$$G = \mathrm{N}_G(P)N.$$

But $N = PM$, so

$$G = \mathrm{N}_G(P)PM = \mathrm{N}_G(P)M$$

(as $P \leqslant \mathrm{N}_G(P)$).

Now consider $J = \mathrm{N}_G(P) \cap M$. Since M is abelian, $J \trianglelefteqslant M$. Also since $M \trianglelefteqslant G$, $J = \mathrm{N}_G(P) \cap M \trianglelefteqslant \mathrm{N}_G(P)$. Hence

$$J \trianglelefteqslant \mathrm{N}_G(P)M = G.$$

But M is a minimal normal subgroup of G, so $J = \mathbf{1}$ or $J = M$.

If $J = \mathrm{N}_G(P) \cap M = M$, then $M \leqslant \mathrm{N}_G(P)$, so $G = \mathrm{N}_G(P)$. Hence P is a normal p-subgroup of G and some subgroup of P is a minimal normal subgroup of G and this is then an elementary abelian p-group with $p \in \pi$. This is contrary to the general assumption made for Case 2.

Thus $J = \mathbf{1}$, so $\mathrm{N}_G(P) \cap M = \mathbf{1}$. Now

$$mq^\beta = |G| = |\mathrm{N}_G(P)M| = |\mathrm{N}_G(P)| \cdot |M|,$$

so $|\mathrm{N}_G(P)| = m$. Hence $H = \mathrm{N}_G(P)$ is our Hall π-subgroup. (We have now completed the existence part of the whole theorem!)

Now consider some π-subgroup L of G. We have $G = HM$ above, so

$$\begin{aligned} LM &= LM \cap G \\ &= LM \cap HM \\ &= (LM \cap H)M \end{aligned}$$

by Dedekind's Modular Law (Lemma 1.7). Now $LM \cap H$ is a π-group (as a

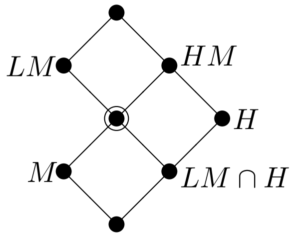

Figure 6.1: Subcase 2B: $LM \cap HM = (LM \cap H)M$

subgroup of H) and

$$\begin{aligned} |LM : LM \cap H| &= \frac{|(LM \cap H)M|}{|LM \cap H|} \\ &= \frac{|M|}{|LM \cap H \cap M|} = |M| \end{aligned}$$

(since $H \cap M = \mathbf{1}$ as they have coprime order). Hence $LM \cap H$ is a Hall π-subgroup of LM.

If $LM < G$, we can apply induction to the group LM to see that some conjugate of the Hall π-subgroup $LM \cap H$ contains the π-subgroup L:

$$L \leqslant (LM \cap H)^x \leqslant H^x$$

for some $x \in G$ (indeed we could pick $x \in LM$). We would then be done.

So suppose $LM = G$. Then as $L \cap M = \mathbf{1}$ (they have coprime order) we see

$$|G| = |LM| = |L| \cdot |M|,$$

so $|L| = mq^\beta/q^\beta = m$. Also, since $M \leqslant N$, we have $G = LN$ (we already know that $G = LM$). Thus

$$|G| = |LN| = \frac{|L| \cdot |N|}{|L \cap N|},$$

so

$$|L \cap N| = \frac{|L| \cdot |N|}{|G|} = \frac{m \cdot p^\alpha q^\beta}{mq^\beta} = p^\alpha.$$

Thus $L \cap N$ is a Sylow p-subgroup of N. By Sylow's Theorem, it is conjugate to the Sylow p-subgroup P which we already know about, say

$$L \cap N = P^x \qquad \text{where } x \in G$$

(indeed we can pick $x \in N$). Now $L \cap N \trianglelefteqslant L$, so

$$L \leqslant \mathrm{N}_G(L \cap N) = \mathrm{N}_G(P^x) = \mathrm{N}_G(P)^x = H^x.$$

Thus L is contained in some conjugate of our Hall π-subgroup H.

This completes the proof of the both statements. We have now shown that if G is a finite soluble group, then G has a Hall π-subgroup H (i.e., part (i) of Theorem 6.28 holds), and every π-subgroup of G is contained in a conjugate of H (and thus part (iii) of Theorem 6.28 holds).

Now let K be any Hall π-subgroup of G. By the second statement, $K \leqslant H^x$ for some $x \in G$. But these subgroups have the same order, so we deduce $K = H^x$ and so part (ii) of Theorem 6.28 holds.

This completes the proof of Hall's Theorem. □

Sylow systems and Sylow bases

We shall now examine some consequences of Hall's Theorem. Specifically we shall see how the Sylow subgroups of a soluble group can be arranged to have special properties. We begin with the following definition.

Definition 6.30 If p is a prime number, we write p' for the set of all primes not equal to p. A Hall p'-subgroup of a finite group G is called a *p-complement*.

Note that $2'$ then denotes the set of all odd primes.

The reason for the above nomenclature is as follows. Let G be a finite group and write $|G| = p^n m$ where p does not divide m. Then a Hall p'-subgroup H has order m, while a Sylow p-subgroup P has order p^n. As they have coprime orders, we see $H \cap P = \mathbf{1}$ and therefore

$$|HP| = |H| \cdot |P| = |G|,$$

so

$$G = HP, \quad H \cap P = \mathbf{1}.$$

This is the situation we referred to as H and P being complements (see Definition 5.8, although neither subgroup is necessarily normal here).

Now let G be a finite soluble group and write

$$|G| = p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$$

where $p_1, p_2, \dots, p_k$ are the distinct prime factors of $|G|$. By Hall's Theorem, G has a Hall p'-subgroup for each prime. Let Q_1, Q_2, $\dots$, Q_k be Hall p_i'-subgroups for $i = 1, 2, \dots, k$, respectively. Thus they are characterised by

$$|Q_i| = |G|/p_i^{n_i} \qquad \text{and} \qquad |G : Q_i| = p_i^{n_i}.$$

Claim: $Q_1 \cap Q_2 \cap \cdots \cap Q_t$ is a Hall $\{p_{t+1}, \dots, p_k\}$-subgroup of G.

We are intersecting subgroups whose indices are coprime. We recall the principal fact from Lemma 1.12 which we need: if $|G : H|$ and $|G : K|$ are coprime integers, then

$$|G : H \cap K| = |G : H| \cdot |G : K|.$$

Proof of Claim: This is certainly true when $t = 1$. Suppose then, as an inductive hypothesis, that the intersection $H = Q_1 \cap Q_2 \cap \cdots \cap Q_t$ is a Hall $\{p_{t+1}, \dots, p_k\}$-subgroup of G. Then

$$|H| = p_{t+1}^{n_{t+1}} \dots p_k^{n_k} \qquad \text{and} \qquad |G : H| = p_1^{n_1} \dots p_t^{n_t}.$$

Now apply Lemma 1.12: H and Q_{t+1} have coprime indices, so

$$|G : H \cap Q_{t+1}| = |G : H| \cdot |G : Q_{t+1}| = p_1^{n_1} \dots p_t^{n_t} p_{t+1}^{n_{t+1}}.$$

Hence $|H \cap Q_{t+1}| = p_{t+2}^{n_{t+2}} \dots p_k^{n_k}$, so $H \cap Q_{t+1} = Q_1 \cap \cdots \cap Q_{t+1}$ is a Hall $\{p_{t+2}, \dots, p_k\}$-subgroup of G. Thus the claim holds by induction. □

In particular, $P_k = Q_1 \cap Q_2 \cap \dots Q_{k-1}$ is a Hall $\{p_k\}$-subgroup of G; that is, a Sylow p_k-subgroup of G.

Generalising in the obvious way, we deduce that

$$P_r = \bigcap_{i \neq r} Q_i$$

is a Sylow p_r-subgroup of G (for $r = 1, 2, \dots, k$).

Now consider the two Sylow subgroups P_{k-1} and P_k. Firstly $P_{k-1} \cap P_k = \mathbf{1}$ (since they have coprime orders), so

$$|P_{k-1}P_k| = |P_{k-1}| \cdot |P_k| = p_{k-1}^{n_{k-1}} p_k^{n_k} = |P_k P_{k-1}|.$$

Further, by construction, both P_{k-1} and P_k are contained in the intersection $Q_1 \cap Q_2 \cap \cdots \cap Q_{k-2}$ and by our claim this intersection is a Hall $\{p_{k-1}, p_k\}$-subgroup of G; that is,

$$|Q_1 \cap Q_2 \cap \cdots \cap Q_{k-2}| = p_{k-1}^{n_{k-1}} p_k^{n_k}.$$

Now since it is a subgroup, this Hall subgroup is closed under products, so we deduce that

$$P_{k-1}P_k,\ P_kP_{k-1} \subseteq Q_1 \cap Q_2 \cap \cdots \cap Q_{k-2}.$$

Finally the subsets occurring in the previous inclusion all have the same size, so we deduce

$$P_{k-1}P_k = Q_1 \cap Q_2 \cap \cdots \cap Q_{k-2} = P_kP_{k-1}.$$

Generalising in the obvious way, we deduce that for all $r \neq s$:

$$P_rP_s = P_sP_r.$$

Definition 6.31 Let G be a finite group and let p_1, p_2, ..., p_k be the distinct prime factors of $|G|$.

(i) A *Sylow system* for G is a collection Q_1, Q_2, ..., Q_k such that Q_i is a Hall p_i'-subgroup of G (for $i = 1, 2, \ldots, k$).

(ii) A *Sylow basis* for G is a collection P_1, P_2, ..., P_k such that P_i is a Sylow p_i-subgroups of G (for $i = 1, 2, \ldots, k$) and such that

$$P_i P_j = P_j P_i \qquad \text{for all } i \text{ and } j.$$

We have shown:

Theorem 6.32 *A finite soluble group possesses a Sylow system and a Sylow basis.* □

Recall that the product HK of two subgroups is a subgroup if and only if $HK = KH$. Consequently, if we start with a Sylow basis P_1, P_2, ..., P_k for a finite soluble group G, then we can form

$$P_{i_1} P_{i_2} \ldots P_{i_s}$$

for any subset $\{i_1, i_2, \ldots, i_s\} \subseteq \{1, 2, \ldots, k\}$. The fact that the Sylow subgroups in our Sylow basis permute ensures that this is a subgroup and it is easy to see that its order is $p_{i_1}^{n_{i_1}} p_{i_2}^{n_{i_2}} \ldots p_{i_s}^{n_{i_s}}$. Thus we have formed a Hall subgroup for the appropriate collection of primes. Hence a Sylow basis is a nice collection of Sylow subgroups from which we may easily construct Hall subgroups.

Philip Hall proved far more than these results. The final two theorems of this section will not be proved (though the first appears, in guided form, on the problem sheet).

Theorem 6.33 (P. Hall) *Let G be a finite soluble group. Then any two Sylow bases for G are conjugate (that is, if P_1, P_2, ..., P_k and R_1, R_2, ..., R_k are two Sylow bases for G, where P_i and R_i are Sylow subgroups for the same prime, then there exists $x \in G$ such that $R_i = P_i^x$ for all i).*

This is much stronger than Sylow's Theorem. The latter tells us that each R_i is a conjugate of P_i. What the above theorem tells us is that when the Sylow subgroups come from a Sylow basis then we can actually choose the same element x to conjugate all the Sylow subgroups simultaneously.

Finally we have the following major converse to Hall's Theorem.

Theorem 6.34 (P. Hall) *Let G be a finite group which possesses a Hall p'-subgroup for every prime p. Then G is soluble.*

Putting Theorems 6.28 and 6.34 together, we see that a group is soluble if and only if it has Hall π-subgroups for all collections π of primes. (In particular, our observation that A_5 was missing some Hall subgroups is no longer surprising.)