Chapter 7
Additional Topics

This chapter contains a collection of miscellaneous topics. Section 7.1 pertains to M. Dehn's algorithmic problems for finitely generated nilpotent groups. In Section 7.2, we prove that finitely generated nilpotent groups are Hopfian. Section 7.3 contains useful facts about groups of upper unitriangular matrices over a commutative ring with unity R. In Section 7.4, we study certain groups of automorphisms that are themselves nilpotent. In particular, we prove that if G is a nilpotent group of class c, then the group of those automorphisms of G that induce the identity on the abelianization of G is nilpotent of class $c - 1$. Section 7.5 ends the chapter with an overview of the Frattini subgroup $\Phi(G)$ and Fitting subgroup $Fit(G)$ of a group G. Among other results, we prove that if G is a finite group, then $\Phi(G)$ is nilpotent, $Fit(G/\Phi(G)) = Fit(G)/\Phi(G)$, and $\Phi(G) \trianglelefteq Fit(G)$.

7.1 Decision Problems

In 1911, M. Dehn raised three decision problems about finitely presented groups. In what follows, let G be a group given by a finite presentation $G = \langle X \mid R \rangle$. An arbitrary (not necessarily reduced) word in the generators X is termed an *X-word*. We now state the problems.

The Word Problem: *Is there an algorithm which determines whether or not an X-word is the identity in G?*

The Conjugacy Problem: *Is there an algorithm which determines whether or not any pair of X-words g and h of G are conjugate in G? In other words, does an X-word k exist in G such that $g = k^{-1}hk$ in G?*

The Isomorphism Problem: *Let H be another finitely presented group with presentation*

$$H = \langle Y \mid S \rangle .$$

A.E. Clement et al., *The Theory of Nilpotent Groups*,
DOI 10.1007/978-3-319-66213-8_7

Is there an algorithm which determines whether or not G and H are isomorphic?

If the answer to any such problem is "yes," then we say that the problem is *solvable*. In this section, we prove that the word problem and the conjugacy problem for finitely generated nilpotent groups are solvable. In fact, we prove the more general result that every finitely presented residually finite group has a solvable word problem.

The isomorphism problem was solved in the positive by Grunewald and Segal [4]. This is by far the most complicated of all of the decision problems for finitely generated nilpotent groups. The algorithms associated with it are quite lengthy and take up about 30 pages in the cited paper.

7.1.1 The Word Problem

We present the solution of the word problem given in [1] which uses residual finiteness. We begin with a key theorem.

Theorem 7.1 (J. C. C. McKinsey) *Every finitely presented residually finite group has a solvable word problem.*

Proof Let G be a finitely presented residually finite group and assume that G is given by an explicit finite presentation. Let w be a given word in the generators. We begin by describing two separate effective procedures.

- The first procedure simply enumerates all consequences of the defining relators. If w appears in this enumeration, then the procedure stops.
- The second procedure begins by enumerating all finite groups, say by constructing their multiplication tables. For each finite group F, the procedure then constructs the (finitely many) homomorphisms θ from G to F. This is done by assigning an element of F to each generator of G, then checking that each defining relator of G maps to the identity element in F. For each such homomorphism θ, the procedure then computes $\theta(w)$ in F. If there exists a finite group F and a homomorphism $\theta : G \to F$ such that $\theta(w) \neq 1$ in F, then the procedure stops.

Now, if $w = 1$ in G, then w will turn up as a consequence of the defining relators and the first procedure will stop. On the other hand, if $w \neq 1$, then the residual finiteness of G guarantees that $w \notin N$ for some $N \lhd G$ with $F = G/N$ finite. Thus, the image of w in F will be a nonidentity element and the second procedure will stop. We conclude that if the first procedure stops, then $w = 1$ in G, whereas if the second one stops, then $w \neq 1$ in G. □

Polycyclic groups are always finitely presentable. This is an immediate result of the next theorem due to P. Hall.

Theorem 7.2 *If G is a group with $N \trianglelefteq G$, and both N and G/N are finitely presented, then G is finitely presented.*

Proof Let $1_{G/N}$ denote the identity element of G/N. Suppose that

$$N = \langle x_1, \ldots, x_n \mid r_1(x_1, \ldots, x_n) = 1, \ldots, r_m(x_1, \ldots, x_n) = 1 \rangle$$

and

$$G/N = \langle g_1N, \ldots, g_lN \mid s_1(g_1N, \ldots, g_lN) = 1_{G/N}, \ldots, s_k(g_1N, \ldots, g_lN) = 1_{G/N} \rangle.$$

Clearly,

$$G = gp(x_1, \ldots, x_n, g_1, \ldots, g_l),$$

and thus G is finitely generated. The relations in these generators are given by

$$r_i(x_1, \ldots, x_n) = 1 \quad (i = 1, \ldots, m),$$
$$s_j(g_1, \ldots, g_l) = t_j(x_1, \ldots, x_n) \quad (j = 1, \ldots, k),$$
$$g_jx_ig_j^{-1} = u_{ij}(x_1, \ldots, x_n) \quad (i = 1, \ldots, n \text{ and } j = 1, \ldots, l), \text{ and}$$
$$g_j^{-1}x_ig_j = v_{ij}(x_1, \ldots, x_n) \quad (i = 1, \ldots, n \text{ and } j = 1, \ldots, l).$$

Define $\overline{G}$ to be the group presented by generators $\overline{x}_1, \ldots, \overline{x}_n, \overline{g}_1, \ldots, \overline{g}_l$ and subject to the above relations in these generators. We claim that $G \cong \overline{G}$.

By Von Dyck's Lemma (see 2.2.1 in [11]), there exists a surjective homomorphism $\varphi : \overline{G} \to G$ determined by

$$\overline{x}_i \mapsto x_i \text{ and } \overline{g}_j \mapsto g_j \quad (i = 1, \ldots, n \text{ and } j = 1, \ldots, l).$$

Let $K = ker\ \varphi$. We need to show that $K = 1$. Put $\overline{N} = gp(\overline{x}_1, \ldots, \overline{x}_n) < \overline{G}$. The restriction of φ to $\overline{N}$ determines an isomorphism with N because all of the relations in the elements x_i are consequences of the relations $r_i(x_1, \ldots, x_n) = 1$ for $i = 1, \ldots, m$. Consequently, $K \cap \overline{N} = 1$. Since

$$\overline{g}_j\overline{x}_i\overline{g}_j^{-1} \in \overline{N} \text{ and } \overline{g}_j^{-1}\overline{x}_i\overline{g}_j \in \overline{N}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, l$, we have that $\overline{N} \lhd \overline{G}$. Furthermore, φ induces an injective map

$$\overline{\varphi} : \overline{G}/\overline{N} \to G/N \text{ such that } \overline{g}_i\overline{N} \mapsto g_iN$$

because $\varphi(\overline{N}) = N$. Now, all relations in the elements g_iN are consequences of the relations

$$s_j(g_1N, \ldots, g_lN) = 1_{G/N} \quad (j = 1, \ldots, k).$$

Hence, $\overline{\varphi}$ is an isomorphism and thus has trivial kernel. However, $ker\ \overline{\varphi} = K\overline{N}/\overline{N}$. And so, $K \leq \overline{N}$. Since $K \cap \overline{N} = 1$, we conclude that $K = 1$. □

Corollary 7.1 *Every polycyclic group is finitely presentable.*

By Theorems 4.4 and 7.1, together with Corollaries 5.21 and 7.1, we now have:

Theorem 7.3 *Polycyclic groups have solvable word problem. In particular, finitely generated nilpotent groups have solvable word problem.*

7.1.2 The Conjugacy Problem

In order to prove that the conjugacy problem for finitely generated nilpotent groups is solvable, we need a theorem of N. Blackburn [2].

Definition 7.1 A group G is called *conjugacy separable* if, whenever two elements g and h of G are conjugate, then the images of g and h in every finite homomorphic image of G are conjugate.

Theorem 7.4 (N. Blackburn) *Every finitely generated nilpotent group is conjugacy separable.*

We introduce some notation: if u and v are conjugate elements of a group, then we write $u \sim v$; otherwise, we write $u \nsim v$.

Proof The proof, which is adopted from [1], is done by induction on the Hirsch length r of G. If $r = 0$, then G is finite and the result is immediate.

Suppose that $r > 0$ and assume the theorem is true for every finitely generated nilpotent group of Hirsch length less than r. Let g and h be elements of G such that $gN \sim hN$ for every $N \lhd G$ with G/N finite. We claim that $g \sim h$. Assume on the contrary, that $g \nsim h$. Since $r > 0$, G must be infinite. By Lemma 2.27, there exists of an element $a \in Z(G)$ of infinite order. Let

$$H_i = gp\left(a^{i!}\right) \quad (i = 1,\ 2,\ \ldots).$$

By Theorem 4.7, the Hirsch length of each G/H_i is $r - 1$. Suppose that $gH_i \nsim hH_i$ for some i. By induction and the Third Isomorphism Theorem, there exists a normal subgroup N_i/H_i of G/H_i such that

$$(G/H_i)\ /\ (N_i/H_i) \cong G/N_i$$

is finite and the images of g and h under the natural homomorphism $G \to G/N_i$ are not conjugate in G/N_i. As this contradicts our earlier assumption that the images of g and h are conjugate in every finite quotient of G, it must be the case that $gH_i \sim hH_i$ for all $i \in \mathbb{N}$. In particular, $gH_1 \sim hH_1$ in G/H_1. Since $H_1 = gp(a)$, we can find a nonzero integer m and an element $k \in G$ such that $h^k = ga^m$. It immediately follows that for all $i \in N$,

$$gH_i \sim ga^m H_i.$$

Thus, for each $i \in \mathbb{N}$, we can find $d_i \in G$ and $s_i \in \mathbb{Z}$ such that

$$g^{d_i} = ga^m \left(a^{i!}\right)^{s_i}. \tag{7.1}$$

Put $L = gp(d_1, d_2, \dots, g, a) \leq G$. Since $a \in Z(G)$, it is clear that $a \in Z(L)$. Moreover, it follows from (7.1) that

$$[d_i, g] \in gp(a)$$

for all $i \in \mathbb{N}$. Since $gp(a) \leq Z(L)$, we apply the commutator calculus to conclude that $[x, g] \in gp(a)$ for all $x \in L$. As a result, we obtain a homomorphism

$$\varphi : L \to gp(a) \text{ defined by } \varphi(x) = [x, g].$$

It follows immediately that $L/ker\ \varphi$ is cyclic. Hence, there exists $b \in L$ such that

$$L = gp(ker\ \varphi, b). \tag{7.2}$$

We henceforth assume, on replacing b by b^{-1} if necessary, that

$$[b, g] = a^\alpha \quad (\alpha \geq 0). \tag{7.3}$$

We now argue that α cannot be zero. If it were the case that $\alpha = 0$, then $[b, g] = 1$. Since $ker\ \varphi = C_L(g)$, this would imply that $g \in Z(L)$. Thus, by (7.1),

$$1 = a^m \left(a^{i!}\right)^{s_i}$$

for all $i \in \mathbb{N}$. Since a has infinite order, it must be that $m + i!s_i = 0$ for all $i \in \mathbb{N}$. This is impossible since $m \neq 0$ and the sequence $\{i! \mid i \in \mathbb{N}\}$ is strictly increasing. We conclude that $\alpha > 0$ in (7.3).

Next, we find all of the conjugates of g in L. By (7.2), such a conjugate has the form g^{xb^n}, where $x \in ker\ \varphi$ and $n \in \mathbb{Z}$. Since $[g, b] \in Z(L)$ and $ker\ \varphi = C_L(g)$, it must be the case that

$$g^{xb^n} = g^{b^n} = g\,[g, b^n] = g[g, b]^n = ga^{-n\alpha}.$$

It follows that the conjugates of g in L are:

$$g, ga^{\pm\alpha}, ga^{\pm 2\alpha}, \dots . \tag{7.4}$$

Since g and ga^m are not conjugate in G, they cannot be conjugate in L. Thus, α does not divide m. However, (7.1) yields

$$g^{d_\alpha} = ga^m\left(a^{\alpha!}\right)^{s_\alpha}$$

because $\alpha > 0$. This means that $ga^m\left(a^{\alpha!}\right)$ is one of the conjugates listed in (7.4), so $m + s_\alpha \alpha! = \lambda\alpha$ for some integer λ. Hence, α divides m, a contradiction. This completes the proof. □

Corollary 7.2 *The conjugacy problem for finitely generated nilpotent groups is solvable.*

Proof Let G be a finitely generated nilpotent group. We can assume that G is given by an explicit finite presentation in light of Theorem 4.4 and Corollary 7.1. Let g and h be words in the generators of G. As we did in the proof of Theorem 7.1, we describe two separate procedures.

- The first procedure begins by enumerating the consequences of the finitely many relators. If there exists a word k in the generators of G such that $h^{-1}g^k$ appears in this enumeration, then the procedure stops.
- As before, the second procedure begins by listing all finite groups (up to isomorphism) by means of their multiplication tables. For every finite group F, we compute (the finitely many) homomorphisms from G to F. Once again, this is done by assigning an element of F to each generator of G and then verifying that the defining relators are mapped to 1. Finally, for every homomorphism $\varphi : G \to F$, we compute $\varphi(g)$ and $\varphi(h)$ in F. If there exists a finite group F and a homomorphism $\varphi : G \to F$ in which $\varphi(g)$ and $\varphi(h)$ are not conjugate, then the procedure stops.

Now, if g and h are conjugate in G, then there is a word k in the generators of G such that $h^{-1}g^k = 1$ in G. Hence, the word $h^{-1}g^k$ is a consequence of the defining relations and the first procedure will stop.

Next, assume that g and h are not conjugate in G. By Theorem 7.4, there exists a finite group F and a homomorphism $\varphi : G \to F$ such that $\varphi(g)$ and $\varphi(h)$ are not conjugate in F, so the second procedure stops.

Therefore, if the first procedure stops, g and h are conjugate in G; whereas if the second procedure stops, then g and h are not conjugate in G. This completes the solution to the conjugacy problem. □

7.2 The Hopfian Property

In the 1930s, H. Hopf asked whether a finitely generated group could be isomorphic to a proper quotient of itself. Groups which are *not* isomorphic to proper quotients of themselves are known as *Hopfian*. In this section, we prove that every finitely generated nilpotent group is Hopfian. An example of a non-Hopfian group will be presented as well.

Definition 7.2 A group G is *Hopfian* if $G/N \cong G$ for some $N \trianglelefteq G$ implies that $N = 1$. Equivalently, every epimorphism of G is an isomorphism of G.

Recall that a group G satisfies Max if all of its subgroups are finitely generated (see Definition 2.14). By Theorem 2.16, G satisfies Max if and only if every ascending series of subgroups stabilizes.

Proposition 7.1 *If G satisfies Max, then G is Hopfian.*

Proof The proof is done by contradiction. Suppose that G satisfies Max and assume that there is an epimorphism $\Phi : G \to G$ with nontrivial kernel. For $j \geq 1$, set

$$\Phi^{\circ j} = \underbrace{\Phi \circ \cdots \circ \Phi}_{j}.$$

Clearly, $\Phi^{\circ j} : G \to G$ is an epimorphism for each $j \geq 1$. Let $K_j = ker\left(\Phi^{\circ j}\right)$. Observe that $\Phi^{\circ j}\left(K_j\right) = 1$ and $\Phi^{\circ j}\left(K_{j+1}\right) = ker\, \Phi$. We claim that

$$K_1 < K_2 < \cdots$$

is an infinite properly ascending chain of subgroups of G. Indeed, since $ker\, \Phi$ is nontrivial, there exists $k_{j+1} \in K_{j+1}$ such that

$$1 \neq \Phi^{\circ j}\left(k_{j+1}\right) \in ker\, \Phi.$$

Thus, $k_{j+1} \notin K_j$. This contradicts the fact that G satisfies Max. Therefore, the kernel of Φ must be trivial and thus Φ is a monomorphism. Consequently, Φ is an isomorphism. And so, G is Hopfian. □

Corollary 7.3 *Every finitely generated nilpotent group is Hopfian.*

Proof Finitely generated nilpotent groups satisfy Max by Theorem 2.18. Apply Proposition 7.1. □

Remark 7.1 More generally, A. I. Mal'cev proved that every finitely generated residually finite group is Hopfian.

An example of a non-Hopfian group is the *Baumslag-Solitar* group $BS(2,\ 3)$, where

$$BS(2,\ 3) = \left\langle a,\ b \;\middle|\; b^{-1}a^2b = a^3 \right\rangle.$$

Consider the map

$$\theta : BS(2,\ 3) \to BS(2,\ 3) \text{ induced by } a \mapsto a^2 \text{ and } b \mapsto b.$$

The fact that θ is indeed a well-defined homomorphism follows from Von Dyck's Lemma (see 2.2.1 in [11]). Furthermore, b is clearly in the image of θ, and $a \in im\, \theta$ since

$$\theta\left(a^{-1}b^{-1}ab\right) = \theta\left(a^{-1}\right)\theta\left(b^{-1}\right)\theta\left(a\right)\theta\left(b\right)$$
$$= a^{-2}b^{-1}a^2b = a^{-2}a^3 = a.$$

Thus, θ is an epimorphism. We claim that θ has a nontrivial kernel. First, observe that the element $\left[b^{-1}ab,\ a\right]$ can be written as such:

$$\left[b^{-1}ab,\ a\right] = \left(b^{-1}ab\right)^{-1}a^{-1}b^{-1}aba$$
$$= b^{-1}a^{-1}ba^{-1}b^{-1}aba.$$

Hence, $\left[b^{-1}ab,\ a\right] \neq 1$ by Britton's Lemma (see Chapter IV in [9]). The element $\left[b^{-1}ab,\ a\right]$ belongs to the kernel of θ because

$$\theta\left(\left[b^{-1}ab,\ a\right]\right) = \left[\theta\left(b^{-1}ab\right),\ \theta(a)\right]$$
$$= \left[b^{-1}a^2b,\ a\right] = \left[a^3,\ a\right] = 1.$$

Since θ is an epimorphism of $BS(2,\ 3)$ onto itself,

$$BS(2,\ 3)/ker\ \theta \cong BS(2,\ 3)$$

by the First Isomorphism Theorem; that is, $BS(2,\ 3)$ is isomorphic to a proper quotient of itself.

7.3 The (Upper) Unitriangular Groups

In this section, R will always be a commutative ring with unity. An important collection of nilpotent groups are the *(upper) unitriangular groups of degree n over R*, denoted by $UT_n(R)$. These groups were first encountered in Example 2.17 of Section 1.3. The purpose of this section is to acquaint the reader with some of their fundamental properties.

Recall from Example 2.17 that if S is the set of $n \times n$ upper triangular matrix over R whose main diagonal entries are all 0, then

$$UT_n(R) = \{I_n + M \mid M \in S\}$$

is a nilpotent group of class less than n. For any $1 \leq m \leq n$, let $UT_n^m(R)$ be the normal subgroup of $UT_n(R)$ consisting of those matrices whose $m-1$ superdiagonals have 0's in their entries. For example,

$$UT_4^2(R) = \left\{ \begin{pmatrix} 1 & 0 & r & s \\ 0 & 1 & 0 & t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \middle| \; r,\ s,\ t \in R \right\} \quad \text{and} \quad UT_4^3(R) = \left\{ \begin{pmatrix} 1 & 0 & 0 & r \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \middle| \; r \in R \right\}.$$

In particular, $UT_n^1(R) = UT_n(R)$ and $UT_n^n(R) = I_n$. Note that we have the inclusions

$$UT_n(R) > UT_n^2(R) > \cdots > UT_n^{n-1}(R) > I_n.$$

Many of the properties of $UT_n(R)$ can be derived using certain matrices called transvections. Let E_{ij} denote the $n \times n$ matrix with 1 in the (i,j) entry and 0's elsewhere. It is easy to see that

$$E_{ij} \cdot E_{kl} = \begin{cases} E_{il} & \text{if } j = k, \\ 0 & \text{otherwise.} \end{cases} \tag{7.5}$$

Let $r \in R$, and set

$$t_{i,\,j}(r) = I_n + rE_{ij} \quad (1 \le i,\ j \le n \ \text{ and } \ i \ne j).$$

If $r \ne 0$, then $t_{i,\,j}(r)$ is called a *transvection*. We abbreviate $t_{i,\,j}(1) = t_{i,\,j}$.

Lemma 7.1 *Let $r,\ s \in R$, and assume that $i \ne j$ and $k \ne l$. Then*

(i) $t_{i,\,j}(r) \cdot t_{i,\,j}(s) = t_{i,\,j}(r + s)$.

(ii) $\left(t_{i,\,j}(r)\right)^{-1} = t_{i,\,j}(-r)$.

(iii) $\left[t_{i,\,j}(r),\ t_{k,\,l}(s)\right] = \begin{cases} t_{i,\,l}(rs) & \text{if } j = k,\ i \ne l, \\ t_{k,\,j}(-rs) & \text{if } j \ne k,\ i = l, \\ I_n & \text{if } j \ne k,\ i \ne l. \end{cases}$

In case $R = \mathbb{Z}$, we also have

(iv) $t_{i,\,j}(r) = t_{i,\,j}(1)^r = t_{i,\,j}^r$.

Proof Observe that

$$\begin{aligned} t_{i,\,j}(r) \cdot t_{i,\,j}(s) &= \left(I_n + rE_{ij}\right)\left(I_n + sE_{ij}\right) \\ &= I_n + (r + s)E_{ij} + rsE_{ij}^2 \\ &= I_n + (r + s)E_{ij} \\ &= t_{i,\,j}(r + s). \end{aligned}$$

This proves (i). In particular,

$$t_{i,\,j}(r) \cdot t_{i,\,j}(-r) = t_{i,\,j}(r + (-r)) = I_n.$$

This gives (ii). To obtain (iii), we use the fact that

$$\begin{aligned}\left[t_{i,\,j}(r),\ t_{k,\,l}(s)\right] &= t_{i,\,j}(-r)\cdot\ t_{k,\,l}(-s)\cdot t_{i,\,j}(r)\cdot t_{k,\,l}(-s)\\ &= \left(I_n - rE_{ij}\right)\left(I_n - sE_{kl}\right)\left(I_n + rE_{ij}\right)\left(I_n + sE_{kl}\right)\end{aligned}$$

by (ii). For instance, if $j = k$ but $i \neq l$, then a straightforward computation, together with (7.5), give

$$\left[t_{i,\,j}(r),\ t_{k,\,l}(s)\right] = I_n + rsE_{il} = t_{i,\,l}(rs).$$

The rest of (iii) follows similarly.

We prove (iv). The result is immediate for $r = 0$ since $t_{i,\,j}(0) = I_n$. Let $r > 0$. Using (i), we obtain

$$t_{i,\,j}(r) = t_{i,\,j}\left(\sum_{n=1}^{r} 1\right) = \prod_{n=1}^{r} t_{i,\,j}(1) = t_{i,\,j}^{r}. \tag{7.6}$$

If $r < 0$, then (7.6), together with (ii), give

$$t_{i,\,j}(r) = \left(t_{i,\,j}(-r)\right)^{-1} = t_{i,\,j}^{r}$$

since $-r > 0$. □

Next, we find a convenient collection of transvections which generate $UT_n^m(R)$. First, we make a simple observation. Let $A = (a_{kl})$ be an $n \times n$ matrix with entries in R, so that

$$A = \sum_{1\le k,\ l\le n} a_{kl}E_{kl}.$$

For $i \neq j$, we have

$$\begin{aligned}At_{i,\,j}(r) &= A\left(I_n + rE_{ij}\right) = A + rAE_{ij}\\ &= A + r\left(\sum_{1\le k,\ l\le n} a_{kl}E_{kl}\right)E_{ij}\\ &= A + r\sum_{k=1}^{n} a_{ki}E_{kj}\end{aligned}$$

by (7.5). Thus, the product $At_{i,\,j}(r)$ is the matrix obtained by adding r times the ith column of A to the jth column of A.

Theorem 7.5 *The set of transvections* $\left\{t_{i,\,j}(r) \mid j{-}i \ge m,\ r \in R\right\}$ *generates* $UT_n^m(R)$. *If* $R = \mathbb{Z}$, *then the set* $\left\{t_{i,\,j} \mid j - i \ge m\right\}$ *generates* $UT_n^m(\mathbb{Z})$.

Proof Suppose that $A = (a_{ij}) \in UT_n^m(R)$. By the observation above, A can be reduced to the identity matrix by post-multiplying A by a suitable sequence of transvections. More specifically,

$$A\big(t_{1,\,m+1}(-a_{1,\,m+1})\cdots t_{1,\,n}(-a_{1,\,n})\big)\big(t_{2,\,m+2}(-a_{2,\,m+2})\cdots t_{2,\,n}(-a_{2,\,n})\big) \\ \cdots t_{n-m,\,n}(-a_{n-m,\,n}) = I_n.$$

After taking inverses of both sides and applying Lemma 7.1 (ii), the matrix A equals the product of transvections

$$t_{n-m,\,n}(a_{n-m,\,n})\cdots\big(t_{2,\,n}(a_{2,\,n})\cdots t_{2,\,m+2}(a_{2,\,m+2})\big)\big(t_{1,\,n}(a_{1,\,n})\cdots t_{1,\,m+1}(a_{1,\,m+1})\big).$$

Thus, $\{t_{i,\,j}(r) \mid j-i \geq m,\ r \in R\}$ is a generating set for $UT_n^m(R)$. The case $R = \mathbb{Z}$ follows immediately from Lemma 7.1 (iv). □

Setting $m = 1$ in Theorem 7.5 and applying Lemma 7.1 give:

Corollary 7.4 *The set* $\{t_{k-1,\,k}(r) \mid 2 \leq k \leq n,\ r \in R\}$ *generates* $UT_n(R)$, *and the set* $\{t_{k-1,\,k} \mid 2 \leq k \leq n\}$ *generates* $UT_n(\mathbb{Z})$.

Observe that the set $\{t_{k-1,\,k}(r) \mid 2 \leq k \leq n,\ r \in R\}$, together with the identities in Lemma 7.1, give a presentation for $UT_n(R)$.

The lower and upper central series of $UT_n(R)$ coincide. This is the point of the following theorem.

Theorem 7.6 *The series*

$$UT_n(R) > UT_n^2(R) > \cdots > UT_n^{n-1}(R) > 1 \tag{7.7}$$

is both the lower and the upper central series for $UT_n(R)$. *Hence,*

- *The nilpotency class of* $UT_n(R)$ *is exactly* $n-1$,
- $UT_n^i(R) = \gamma_i UT_n(R) = \zeta_{n-i} UT_n(R)$ *for* $i = 1, 2, \ldots, n$ *and*
- $UT_n^j(R)/UT_n^{j+1}(R)$ *is abelian for* $j = 1, 2, \ldots, n-1$.

Proof Put $UT_n(R) = U$ and $UT_n^i(R) = U_i$ for $i = 1, 2, \ldots, n$. We claim that $[U_i, U] = U_{i+1}$ for each $i = 1, 2, \ldots, n-1$.

1. Let $[g, h] \in [U_m, U]$, where $g \in U_m$ and $h \in U$. By Theorem 7.5 and Lemma 7.1, g is a product of transvections of the form $t_{i,\,j}(r)$ with $j-i \geq m$, $r \in R$, and h is a product of transvections of the form $t_{k,\,l}(s)$ with $l-k \geq 1$ and $s \in R$. We prove the claim by induction on the number of transvections occurring in g and h combined.

   - For the basis of induction, assume that $g = t_{i,\,j}(r)$ and $h = t_{k,\,l}(s)$. Suppose that $j = k$ but $i \neq l$. By Lemma 7.1 (iii),

$$[g, h] = t_{i,\,l}(rs).$$

Since in this case $l - i = (j - i) + (l - k) \geq m + 1$, we have $[g, h] \in U_{m+1}$. The other cases are handled in a similar way.

- Suppose that $g = t_{i,j}(r)\bar{g}$ and $h = t_{k,l}(s)$, where $\bar{g} \in U_m$. By Lemma 1.4 (v),

$$[g, h] = \left[t_{i,j}(r)\bar{g},\ t_{k,l}(s)\right] = \left[t_{i,j}(r),\ t_{k,l}(s)\right]^{\bar{g}}\left[\bar{g},\ t_{k,l}(s)\right].$$

Now, $\left[\bar{g},\ t_{k,l}(s)\right] \in U_{m+1}$ and $\left[t_{i,j}(r), t_{k,l}(s)\right] \in U_{m+1}$ by induction. Hence,

$$\left[t_{i,j}(r),\ t_{k,l}(s)\right]^{\bar{g}} \in U_{m+1},$$

and thus $[g, h] \in U_{m+1}$.

- Suppose that $g = t_{i,j}(r)\bar{g}$ and $h = t_{k,l}(s)\bar{h}$, where $\bar{g} \in U_m$ and $\bar{h} \in U$. By Lemma 1.4 (vi),

$$[g, h] = \left[t_{i,j}(r)\bar{g},\ t_{k,l}(s)\bar{h}\right] = \left[t_{i,j}(r)\bar{g},\ \bar{h}\right]\left[t_{i,j}(r)\bar{g},\ t_{k,l}(s)\right]^{\bar{h}}.$$

Since $\left[t_{i,j}(r)\bar{g},\ \bar{h}\right] \in U_{m+1}$ by induction and $\left[t_{i,j}(r)\bar{g},\ t_{k,l}(s)\right] \in U_{m+1}$ by the previous case, we have $[g, h] \in U_{m+1}$.

Therefore, $[U_m, U] \subseteq U_{m+1}$.

2. In order to establish the reverse inclusion $U_{m+1} \subseteq [U_m, U]$, it suffices to show that $[U_m, U]$ contains every transvection $t_{i,j}(r)$ with $j - i \geq m + 1$ and $r \in R$. Consider the transvections $t_{i,i+m}(r) \in U_m$ and $t_{i+m,j} = t_{i+m,j}(1) \in U$. By Lemma 7.1 (iii),

$$t_{i,j}(r) = \left[t_{i,i+m}(r),\ t_{i+m,j}\right] \in [U_m, U].$$

Thus, $U_{m+1} \subseteq [U_m, U]$.

This proves the claim that $[U_i, U] = U_{i+1}$ for $i = 1, \ldots, n - 1$. Therefore, $\gamma_i U = U_i$, and thus (7.7) is the lower central series for $UT_n(R)$. In order to prove that (7.7) is the upper central series for $UT_n(R)$, one can use induction on $n - i$ to show that $\zeta_{n-i}UT_n(R) = \gamma_i UT_n(R)$. We omit the details. □

One can easily find a set of generators for the factor groups of (7.7) by using Theorem 7.5.

Theorem 7.7 *For $m = 1, 2, \ldots, n - 1$, each factor group $UT_n^m(R)/UT_n^{m+1}(R)$ is generated, modulo $UT_n^{m+1}(R)$, by the set $\{t_{i,m+i}(r) \mid 1 \leq i \leq n - m,\ r \in R\}$. If $R = \mathbb{Z}$, then $UT_n^m(\mathbb{Z})/UT_n^{m+1}(\mathbb{Z})$ is generated, modulo $UT_n^{m+1}(\mathbb{Z})$, by the set $\{t_{i,m+i} \mid 1 \leq i \leq n - m\}$.*

We show next that $UT_n(\mathbb{Z})$ is torsion-free. In fact, we prove this for any ring whose characteristic is zero.

Theorem 7.8 *If R has characteristic zero, then $UT_n(R)$ is torsion-free. In particular, $UT_n(\mathbb{Z})$ is torsion-free.*

Proof By Theorem 7.6, the center of $UT_n(R)$ equals the subgroup $UT_n^{n-1}(R)$ consisting of those matrices with an arbitrary entry of R in the uppermost right corner and 0's elsewhere. It follows that $Z(UT_n(R))$ is isomorphic to the additive group of R. Since R has characteristic zero, $Z(UT_n(R))$ is torsion-free. The result follows from Corollary 2.21. □

Remark 7.2 If the characteristic of R is different than zero, then $UT_n(R)$ need not be torsion-free. For example, let R be the polynomial ring in one variable over a finite field of characteristic a prime p and let $2 \leq k \leq n$. By Corollary 7.4, a typical generator for $UT_n(R)$ is a transvection of the form $t_{k-1,\,k}(r)$, where $r \in R$. Since R has characteristic p, we apply Lemma 7.1 (i) repeatedly to obtain

$$(t_{k-1,\,k}(r))^p = t_{k-1,\,k}(rp) = I.$$

Thus, every generator of $UT_n(R)$ is a torsion element.

There is a natural Mal'cev basis for $UT_n(\mathbb{Z})$. By Corollary 7.4 and Theorem 7.8, $UT_n(\mathbb{Z})$ is finitely generated and torsion-free. Each of the factor groups $UT_n^m(\mathbb{Z})/UT_n^{m+1}(\mathbb{Z})$ is torsion-free and finitely generated, modulo $UT_n^{m+1}(\mathbb{Z})$, by transvections of the form $t_{i,\,m+i}$, where $1 \leq i \leq n-m$. This follows directly from Corollary 2.20 and Theorem 7.7. Taking this into consideration, we have:

Lemma 7.2 *The set of transvections*

$$\left\{t_{1,\,2},\ t_{2,\,3},\ \ldots,\ t_{n-1,\,n},\ t_{1,\,3},\ t_{2,\,4},\ \ldots,\ t_{n-2,\,n},\ \ldots,\ t_{1,\,n}\right\}$$

is a Mal'cev basis for $UT_n(\mathbb{Z})$.

7.4 Nilpotent Groups of Automorphisms

In this section, we prove that certain subgroups of the automorphism group of a given group are nilpotent. We begin by defining the holomorph of a group.

Let G be a group and define the set

$$\widetilde{G} = \{\varphi g \mid \varphi \in Aut(G),\ g \in G\}.$$

We can regard $\widetilde{G}$ as the Cartesian product $Aut(G) \times G$. This set becomes a group under the operation

$$(\varphi g)(\varphi' g') = \varphi\varphi' g^{\varphi'} g',$$

where $g^{\varphi'} = \varphi'(g) \in G$. Thus, we have a semi-direct product of G by $Aut(G)$. This group is called the *holomorph* of G, written as $Hol(G)$.

7.4.1 The Stability Group

Definition 7.3 Let G be any group and let

$$G = G_0 \geq G_1 \geq \cdots \geq G_r = 1 \tag{7.8}$$

be a series of subgroups of G. The *stability group* of G relative to the series (7.8) is the group A of all automorphisms φ of G such that

$$\varphi\big(xG_i\big) = xG_i \text{ for every } i = 0, \ldots, r-1 \text{ and } x \in G_{i+1}. \tag{7.9}$$

The group A and the automorphism $\varphi \in A$ are said to *stabilize* the series (7.8).

Putting $x = 1$ in (7.9) shows that each G_i is invariant under φ. In [7], L. Kalužnin proved that the stability group of a group is always solvable of "solvability length" at most m. If the series happens to be a normal series, then the stability group is, in fact, nilpotent.

Theorem 7.9 *Let*

$$G = G_0 \geq G_1 \geq \cdots \geq G_r = 1 \tag{7.10}$$

be a normal series of a group G. If A is the stability group of G relative to (7.10), then A and $[G, A]$ are nilpotent of class less than r. Here, $[G, A]$ is viewed as a subgroup of $Hol(G)$.

In this situation, the stability group acts nilpotently on G according to Definition 6.5. The proof relies on a property of commutator subgroups.

Theorem 7.10 *Let H and K be subgroups of a group G and suppose that*

$$H = H_0 \geq H_1 \geq \cdots$$

is a descending series of normal subgroups of H such that $[H_i, K] \leq H_{i+1}$ for each integer $i \geq 0$. Put $K = K_1$, and for $j > 1$, define

$$K_j = \{x \in K \mid [H_i, x] \leq H_{i+j} \text{ for all } i \geq 0\}.$$

Then $\big[K_j, K_l\big] \leq K_{j+l}$ for all $j, l \geq 1$, and $\big[H_i, \gamma_j K\big] \leq H_{i+j}$ for all $i \geq 0$ and $j \geq 1$.

Proof We show first that $\big[K_j, K_l\big] \leq K_{j+l}$. By definition, $\big[H_i, K_j\big] \leq H_{i+j}$ and $\big[H_{i+j}, K_l\big] \leq H_{i+j+l}$. Hence,

$$\big[H_i, K_j, K_l\big] \leq \big[H_{i+j}, K_l\big] \leq H_{i+j+l}$$

by Proposition 1.1 (iii). Likewise,

$$\big[K_l, H_i, K_j\big] = \big[H_i, K_l, K_j\big] \leq H_{i+j+l}$$

by Proposition 1.1 (i) and (iii). By hypothesis, H_{i+j+l} is normal in H. Moreover, for $x \in K$,

$$x^{-1}H_{i+j+l}x = H_{i+j+l}\big[H_{i+j+l},\ x\big].$$

Since $\big[H_{i+j+l},\ x\big] \le H_{i+j+l+1}$, we conclude that

$$x^{-1}H_{i+j+l}x \le H_{i+j+l}.$$

This shows that H_{i+j+l} is normal in the subgroup of G generated by H and K. Thus,

$$\big[K_j,\ K_l,\ H_i\big] = \big[H_i,\ \big[K_j,\ K_l\big]\big] \le H_{i+j+l}$$

by Proposition 1.1 (i) and Lemma 2.18. Therefore, by definition, $\big[K_j,\ K_l\big] \le K_{j+l}$.

In particular, $\big[K_j,\ K\big] \le K_{j+1}$. Hence, $\gamma_j K \le K_j$ and thus

$$\big[H_i,\ \gamma_j K\big] \le \big[H_i,\ K_j\big] \le H_{i+j}$$

by Proposition 1.1 (iii). This completes the proof. □

Remark 7.3 Note that Theorem 2.14 (i) can be obtained from Theorem 7.10 by setting $H_i = \gamma_{i+1}G$ and $K = G$.

We now prove Theorem 7.9. To begin with, we assert that $[G_i,\ A] \le G_{i+1}$ for $i = 0,\ 1,\ \ldots,\ r-1$. If $x \in G_i$ and $\alpha \in A$, then

$$[x,\ \alpha] = x^{-1}\alpha^{-1}x\alpha = x^{-1}x^{\alpha} \in G_{i+1}$$

since α induces the identity on G_i/G_{i+1}. Put $H_i = G_i$ ($i = 0,\ 1,\ \ldots,\ r$) and $K = A$ in Theorem 7.10 and regard all of these as subgroups of the $Hol(G)$. Then

$$[G,\ \gamma_r A] = [G_0,\ \gamma_r A] \le G_r = 1,$$

and thus $[G,\ \gamma_r A] = 1$. Now, let $\alpha \in \gamma_r A$ and $x \in G$. Then $[x,\ \alpha] = x^{-1}x^{\alpha} = 1$, which shows that $\alpha = 1$. Consequently, $\gamma_r A = 1$ and A is nilpotent of class less than r as asserted.

Next, we prove that $[G,\ A]$ is nilpotent of class less than r. Since $G_i \trianglelefteq G$ for $i = 0,\ 1,\ \ldots,\ r$, we have $[G_i,\ G] \le G_i$. Hence,

$$[G_{i-1},\ G,\ A] \le [G_{i-1},\ A] \le G_i.$$

Furthermore,

$$[A,\ G_{i-1},\ G] = [G_{i-1},\ A,\ G] \le [G_i,\ G] \le G_i$$

by Proposition 1.1 (i). By assumption, $x^\alpha \in G_i$ for $x \in G_i$ and $\alpha \in A$. Thus, G_i is normal in the subgroup of $Hol(G)$ generated by A and G. Therefore,

$$[G, A, G_{i-1}] = [G_{i-1}, [G, A]] \leq G_i$$

by Proposition 1.1 (i). By letting $K = [G, A]$ and $G_i = H_i$ in Theorem 7.10, we have

$$[G_1, \gamma_{r-1}[G, A]] \leq G_r = 1.$$

However, $[G, A] = [G_0, A] \leq G_1$. Therefore,

$$[[G, A], \gamma_{r-1}[G, A]] = 1,$$

and thus $\gamma_r[G, A] = 1$. This completes the proof of Theorem 7.9.

Remark 7.4 By repeating what was done in Example 6.1 with a free R-module of finite rank over a ring with unity, one can conclude from Theorem 7.9 that $UT_n(R)$ is a nilpotent group of class less than n. This was established in Chapter 1 using a different method (see Example 2.17).

The next theorem is a more general result due to P. Hall [5].

Theorem 7.11 *The stability group A relative to any series of subgroups of length $m \geq 1$ of a group G is nilpotent of class at most $m(m-1)/2$.*

We begin by proving Theorem 7.12, from which Theorem 7.11 will follow. Regard both $Aut(G)$ and G as subgroups of $Hol(G)$ and note that the stability group A relative to the series (7.8) also becomes a subgroup of $Hol(G)$. Thus, A is characterized by the property

$$[G_{i-1}, A] \leq G_i \quad (i = 1, \cdots, m)$$

since $\alpha(g) = g^\alpha = \alpha^{-1} g \alpha$ in $Hol(G)$ for every $\alpha \in Aut(G)$. We have already encountered this property in the proof of Theorem 7.9.

Theorem 7.12 *Let H and K be two subgroups of a group G. If*

$$[H, \underbrace{K, \cdots, K}_{m}] = 1$$

for some $m \in \mathbb{N}$, then $[\gamma_{n+1}K, H] = 1$, where $n = m(m-1)/2$.

To obtain Theorem 7.11 from Theorem 7.12, observe that $[G_{i-1}, A] \leq G_i$ implies that

$$[G, \underbrace{A, \cdots, A}_{m}] \leq [G_1, \underbrace{A, \cdots, A}_{m-1}] \leq \cdots \leq [G_{m-1}, A] \leq G_m = 1,$$

so that

$$[G, \underbrace{A, \cdots, A}_{m}] = 1.$$

Theorem 7.12 now implies that $[\gamma_{n+1}A, G] = 1$, where $n = m(m-1)/2$. However, this means that $\alpha(g) = g^{\alpha} = g$ for every $\alpha \in \gamma_{n+1}A$ and $g \in G$. Therefore, $\gamma_{n+1}A$ contains only the identity isomorphism and thus A is nilpotent of class at most n.

The proof of Theorem 7.12 relies on the next lemma.

Lemma 7.3 *Let H, K, and L be subgroups of a group G such that $G = gp(H, K)$. If $[H, K, L] = 1$, then $[L, H, K] = [K, L, H] \trianglelefteq G$.*

Proof Let C be the centralizer of $[H, K]$ in G. By Corollary 1.4, $[H, K] \trianglelefteq gp(H, K)$ and thus $[H, K] \trianglelefteq G$ by assumption. Since $L \leq C$ by hypothesis, a direct computation shows that $[L, H] \leq C$.

Let $x \in H$, $y \in K$, and $t \in [L, H]$. Set $z = \left[x, y^{-1}\right]$. Then $t \in C$ and $z \in [H, K]$, so that t and z commute. Since $y^x = zy$, Lemma 1.4 (vi) gives

$$\left[t, y^x\right] = \left[t, zy\right] = \left[t, y\right]\left[t, z\right]^y = \left[t, y\right].$$

This means that $[L, H, K^x] = [L, H, K]$. Since $[L, H] \trianglelefteq gp(L, H)$ by Corollary 1.4 and $x \in H$, we also have $[L, H] = [L, H]^x$. Hence,

$$[L, H, K]^x = [[L, H]^x, K^x] = [L, H, K^x] = [L, H, K],$$

which means that H normalizes $[L, H, K]$. Since $[L, H, K] \trianglelefteq gp([L, H], K)$ by Corollary 1.4, K also normalizes $[L, H, K]$. Therefore, $[L, H, K]$ is normal in G.

By Proposition 1.1 (i), $[K, H, L] = [H, K, L] = 1$. Thus, we interchange the roles of H and K and conclude that $[L, K, H] = [K, L, H]$ is also normal in G. It follows that $[H, K, L] = 1$ and both $[L, H, K]$ and $[K, L, H]$ are normal in G. By Lemma 2.18, the subgroups $[L, H, K]$ and $[K, L, H]$ are contained in one another. The result now follows. ☐

We now prove Theorem 7.12 by induction on m. If $m = 1$, then $[H, K] = 1$ and $n = 0$. Thus, $[K, H] = [H, K] = 1$ by Proposition 1.1 (i). This gives the basis of induction.

Let $m > 1$, and put $n = m(m-1)/2$. Let $H_1 = [H, K]$, and notice that

$$[H_1, \underbrace{K, \cdots, K}_{m-1}] = 1$$

since, by hypothesis, $[H, \underbrace{K, \cdots, K}_{m}] = 1$. Given that $[H_1, \underbrace{K, \cdots, K}_{m-1}] = 1$, we may assume inductively that

$$[\gamma_{l+1}K, H_1] = 1,$$

where $l = (m-1)(m-2)/2$. A direct calculation shows that $l = n - m + 1$, where $n = m(m-1)/2$. Let $r \geq l + 1$, and observe that

$$[\gamma_r K, H_1] \leq [\gamma_{l+1} K,\ H_1] = 1.$$

By Proposition 1.1 (i),

$$[H,\ K,\ \gamma_r K] = [H_1,\ \gamma_r K] = [\gamma_r K,\ H_1] = 1.$$

By Lemma 7.3 and Proposition 1.1 (i),

$$[\gamma_r K,\ H,\ K] = [K,\ \gamma_r K,\ H] = [\gamma_r K,\ K,\ H] = [\gamma_{r+1} K,\ H]$$

for $r > l$. Hence,

$$[\gamma_{l+1} K, H, \underbrace{K, \cdots, K}_{m-1}] = [\gamma_{l+2} K, H, \underbrace{K, \cdots, K}_{m-2}] = \cdots = [\gamma_n K, H, K] = [\gamma_{n+1} K, H].$$

Since $\gamma_{l+1} K \leq K$, it follows from Proposition 1.1 (i) and (iii) that

$$[\gamma_{l+1} K,\ H] \leq [K,\ H] = [H,\ K].$$

Thus,

$$[\gamma_{n+1} K,\ H] = [\gamma_{l+1} K,\ H,\ \underbrace{K,\ \cdots,\ K}_{m-1}] \leq [H,\ \underbrace{K,\ \cdots,\ K}_{m}].$$

However, $[H,\ \underbrace{K,\ \cdots,\ K}_{m}] = 1$. This completes the proof of Theorem 7.12.

7.4.2 *The IA-Group of a Nilpotent Group*

If G is a nilpotent group of class c, then we conclude from Corollary 1.1 and Lemma 2.12 that $Inn(G)$ is nilpotent of class $c-1$. We turn our attention to a certain nilpotent subgroup of $Aut(G)$ that contains $Inn(G)$ and whose class is also $c-1$. The material that appears in this section is based on [6] and Section 1.2 of [13].

Definition 7.4 An *IA-automorphism* of a group G is an automorphism of G that induces the identity on the abelianization of G.

Thus, α is an *IA*-automorphism of G if and only if α belongs to the kernel of the natural homomorphism

$$Aut(G) \to Aut\,(G/\gamma_2 G)\,.$$

It is easy to show that the set of all *IA*-automorphisms of G is a subgroup of $Aut(G)$. This subgroup is called the *IA-group* of G and is denoted by $IA(G)$. Hence,

$$IA(G) = \{\alpha \in Aut(G) \mid g^{-1}\alpha(g) \in \gamma_2 G \text{ for all } g \in G\}.$$

Clearly, $IA(G) = 1$ whenever G is abelian. Furthermore, $IA(G)$ contains $Inn(G)$. For if $\varphi_h \in Inn(G)$, then

$$\varphi_h(g) = g^h = g[g, h]$$

for every $g \in G$. Since $[g, h] \in \gamma_2 G$, $\varphi_h \in IA(G)$.

Our ultimate goal in this section is to prove that the *IA*-group of a finitely generated torsion-free nilpotent group of class c is finitely generated torsion-free nilpotent of class $c-1$. This is essentially a result due to P. Hall [6]. We begin by proving a lemma which is the main ingredient in establishing that the *IA*-group of a nilpotent group of class c is nilpotent of class $c-1$.

Lemma 7.4 *Let H and K be subgroups of a group G such that $[H, K] \leq \gamma_2 H$. Then $\left[\gamma_i H, \gamma_j K\right] \leq \gamma_{i+j} H$ for all $i, j \in \mathbb{N}$.*

Proof We do “double induction.” Let $j = 1$. We prove that $[\gamma_i H, K] \leq \gamma_{i+1} H$ by induction on i. If $i = 1$, then the result holds by hypothesis. Assume that $[\gamma_{i-1} H, K] \leq \gamma_i H$ for $i > 1$. By the induction hypothesis and Proposition 1.1 (iii),

$$[\gamma_{i-1} H, K, H] \leq [\gamma_i H, H] = \gamma_{i+1} H.$$

Furthermore,

$$[K, H, \gamma_{i-1} H] \leq [\gamma_2 H, \gamma_{i-1} H] \leq \gamma_{i+1} H$$

by hypothesis, Proposition 1.1 (i) and (iii), and Theorem 2.14 (i). Using Lemma 2.18, we have

$$[\gamma_i H, K] = [H, \gamma_{i-1} H, K] \leq \gamma_{i+1} H.$$

The basis of induction for j now follows.

The induction hypothesis for j is $\left[\gamma_i H, \gamma_{j-1} K\right] \leq \gamma_{i+j-1} H$ for all i. By Proposition 1.1 (i),

$$\left[\gamma_i H, \gamma_j K\right] = \left[\gamma_i H, \left[K, \gamma_{j-1} K\right]\right] = \left[K, \gamma_{j-1} K, \gamma_i H\right].$$

By Proposition 1.1 (iii) and the induction hypothesis on j, we have

$$\left[\gamma_i H, K, \gamma_{j-1} K\right] \leq \left[\gamma_{i+1} H, \gamma_{j-1} K\right] \leq \gamma_{i+j} H$$

and

$$\left[\gamma_{j-1}K,\ \gamma_i H,\ K\right] \le \left[\gamma_{i+j-1}H,\ K\right] \le \gamma_{i+j}H.$$

Invoking Lemma 2.18 once again, we conclude that $\left[\gamma_i H,\ \gamma_j K\right] \le \gamma_{i+j}H$. □

Theorem 7.13 *For all $j \in \mathbb{N}$ and any group G, the elements of $\gamma_j IA(G)$ induce the identity on each $\gamma_i G/\gamma_{i+j}G$. If G is nilpotent of class c, then $IA(G)$ is nilpotent of class $c-1$.*

In particular, $IA(G)$ is abelian when G is nilpotent of class 2.

Proof To prove the first assertion, notice that if we choose $x \in G$ and $\psi \in IA(G)$, then

$$[x,\ \psi] = x^{-1}x^{\psi} \in \gamma_2 G$$

in $Hol(G)$. Hence, $[G,\ IA(G)] \le \gamma_2 G$, and by Lemma 7.4,

$$\left[\gamma_i G,\ \gamma_j IA(G)\right] \le \gamma_{i+j}G.$$

This means that the elements of $\gamma_j IA(G)$ induce the identity on $\gamma_i G/\gamma_{i+j}G$ as asserted.

Next, suppose that G is nilpotent of class c. We claim that $IA(G)$ is nilpotent of class $c-1$. Let $\alpha \in \gamma_c IA(G)$. By the previous result, α induces the identity on $\gamma_1 G/\gamma_{c+1}G$. Since $\gamma_1 G/\gamma_{c+1}G \cong G$, α is the identity automorphism and thus $\gamma_c IA(G) = 1$. By Corollary 2.3, $IA(G)$ is nilpotent of class less than c. However,

$$IA(G) \ge Inn(G) \cong G/Z(G),$$

and $G/Z(G)$ is of class $c-1$ by Lemma 2.12. Thus, $IA(G)$ is of class $c-1$. □

Remark 7.5

(i) Theorem 7.13 implies that if G is a nilpotent group, then $IA(G)$ is the stability group of G relative to its lower central series.
(ii) While Theorem 7.11 already implies that $IA(G)$ is nilpotent whenever G is nilpotent, Theorem 7.13 guarantees that the class of $IA(G)$ is exactly one less than the class of G.

If G is a torsion-free nilpotent group, then so is $G/Z(G)$ by Corollary 2.22. Thus, $Inn(G)$ is also torsion-free nilpotent. The fact that every inner automorphism is an IA-automorphism suggests that $IA(G)$ may be torsion-free as well. This is indeed the case.

Lemma 7.5 *If G is a torsion-free nilpotent group, then $IA(G)$ is torsion-free.*

First, we prove an auxiliary result. Let $\overline{\gamma}_c G$ denote the isolator of $\gamma_c G$ in G.

Lemma 7.6 *If G is a torsion-free nilpotent group of class c, then $\overline{\gamma}_c G$ is a central subgroup of G and $G/\overline{\gamma}_c G$ is torsion-free.*

Proof We first prove that $\overline{\gamma}_c G$ is central in G. If $g \in \overline{\gamma}_c G$, then there exists $m \in \mathbb{N}$ such that $g^m \in \gamma_c G$. Since G is of class c, $\gamma_c G \leq Z(G)$, and thus $g^m \in Z(G)$. By Theorem 2.7 and Lemma 5.15, $g \in Z(G)$ as desired. To establish that $G/\overline{\gamma}_c G$ is torsion-free, observe that $\tau(G/\gamma_c G) = \overline{\gamma}_c G/\gamma_c G$ and thus

$$G/\overline{\gamma}_c G \cong (G/\gamma_c G)/\tau(G/\gamma_c G)$$

by the Third Isomorphism Theorem. The result immediately follows from Corollary 2.15. □

We now prove Lemma 7.5 by induction on the class c of G. If $c = 2$, then $IA(G)$ is abelian by Theorem 7.13. Let $\varphi \in IA(G)$ and $x \in G$. Then $\varphi(x) = xd$, where $d \in \gamma_2 G$. Suppose that $\varphi^m = 1$, where $m > 0$. Since G is of class 2, φ acts as the identity on $\gamma_2 G$ by Theorem 7.13. Thus,

$$\varphi^m(x) = xd^m = x,$$

and consequently, $d^m = 1$. Since $\gamma_2 G$ is torsion-free, $d = 1$. This completes the basis of induction.

Assume that the IA-group of a torsion-free nilpotent group of class less than c is always torsion-free. By Lemma 7.6, $G/\overline{\gamma}_c G$ is torsion-free nilpotent and clearly of class less than c. The induction hypothesis gives that $IA(G/\overline{\gamma}_c G)$ is torsion-free. To prove that $IA(G)$ is torsion-free, let $\varphi \in IA(G)$ and assume that $\varphi^m = 1$ where $m > 0$. For $g \in G$, let $[g]$ denote the equivalence class of g in $G/\overline{\gamma}_c G$. Consider the natural homomorphism

$$\pi : IA(G) \to IA\big(G/\overline{\gamma}_c G\big) \text{ defined by } \chi \mapsto \widehat{\chi}, \text{ where } \widehat{\chi}([g]) = [\chi(g)].$$

Since φ^m is the identity and $IA\big(G/\overline{\gamma}_c G\big)$ is torsion-free, $\widehat{\varphi}$ is the identity on $G/\overline{\gamma}_c G$. Let $x \in G$ and write $\varphi(x) = xd$, where $d \in \gamma_2 G$. Then

$$\widehat{\varphi}([x]) = [\varphi(x)] = [xd] = [x].$$

Hence, $d \in \overline{\gamma}_c G$. Since G is of class c, φ acts trivially on $\gamma_c G$ by Theorem 7.13. We claim that φ also acts trivially on $\overline{\gamma}_c G$. To see this, let $y \in \overline{\gamma}_c G$. There exists $m \in \mathbb{N}$ such that $y^m \in \gamma_c G$. Moreover, there exists $y_1 \in \gamma_2 G$ such that $\varphi(y) = yy_1$. Hence,

$$y^m = \varphi(y^m) = \varphi(y)^m = (yy_1)^m = y^m y_1^m$$

since y is central by Lemma 7.6. And so, $y_1^m = 1$. Since G is torsion-free, $y_1 = 1$, and thus φ acts trivially on $\overline{\gamma}_c G$ as claimed. This, together with $\varphi(x) = xd$ and $\varphi^m = 1$, gives

$$\varphi^m(x) = xd^m = x.$$

Therefore, $d^m = 1$, and consequently, $d = 1$. This completes the proof of Lemma 7.5.

Lemma 7.7 *If G is a finitely generated nilpotent group, then $IA(G)$ is finitely generated.*

Proof The proof is done by induction on the class c of G. If $c = 2$, then $IA(G)$ is abelian by Theorem 7.13. Since G is finitely generated, so is $\gamma_2 G$ by Theorem 2.18. A typical member of a generating set for $IA(G)$ can be constructed as follows: let X be a finite set of generators for G and Y a finite set of generators for $\gamma_2 G$. For every $x \in X$ and $y \in Y$, construct the IA-automorphism that sends x to xy and each remaining generator of G to itself. It is clear that the set of all such IA-automorphisms generates $IA(G)$. Since X and Y are finite, $IA(G)$ is finitely generated.

Assume that the IA-group of any finitely generated nilpotent group of class less than c is finitely generated. Consider the natural homomorphism

$$\pi : IA(G) \to IA(G/\gamma_c G) \text{ defined by } \chi \mapsto \widehat{\chi}, \text{ where } \widehat{\chi}([g]) = [\chi(g)]$$

and $[g]$ now denotes the equivalence class of g in $G/\gamma_c G$. The kernel of π is the subgroup

$$I_c = \left\{\alpha \in IA(G) \;\middle|\; g^{-1}\alpha(g) \in \gamma_c G \text{ for all } g \in G\right\}$$

of $IA(G)$, and it is finitely generated. This can be established by an analogous construction as before, where Y is taken to be a finite generating set for $\gamma_c G$. By the induction hypothesis, $IA\big(G/\gamma_c G\big)$ is finitely generated. Thus, the image of π, being a subgroup of a finitely generated nilpotent group, is also finitely generated by Theorem 2.18, as well as isomorphic to $IA(G)/I_c$. Since $IA(G)$ is an extension of a finitely generated group by another, it must be finitely generated. □

Theorem 7.13, together with Lemmas 7.5 and 7.7, gives:

Theorem 7.14 *If G is a finitely generated torsion-free nilpotent group of class c, then $IA(G)$ is finitely generated, torsion-free nilpotent of class $c - 1$.*

Using basic sequences and the commutator calculus, M. Zyman [13] has shown that if G is a finitely generated nilpotent group such that $\gamma_2 G$ is abelian, then $IA(G_p) \cong \big(IA(G)\big)_p$, where p is any prime and G_p and $\big(IA(G)\big)_p$ denote the p-localizations of G and $IA(G)$ respectively. A group G for which $\gamma_2 G$ is abelian is termed *metabelian*.

7.5 The Frattini and Fitting Subgroups

Two subgroups which provide useful information about the structure of a group are the Frattini and Fitting subgroups. In this section, we give a brief overview of the properties of these subgroups and describe their connection to nilpotent groups.

7.5.1 The Frattini Subgroup

Definition 7.5 Let G be any group. The *Frattini subgroup* of G, denoted by $\Phi(G)$, is the intersection of all of the maximal subgroups of G.

By convention, $\Phi(G) = G$ if G has no maximal subgroups. Thus, $\Phi(\mathbb{Q}) = \mathbb{Q}$ and $\Phi\left(\mathbb{Z}_{p^\infty}\right) = \mathbb{Z}_{p^\infty}$. Clearly, the Frattini subgroup of a group is characteristic since every group automorphism maps maximal subgroups to maximal subgroups.

Example 7.1 We give the Frattini subgroup of certain groups.

1. For each prime p, the subgroup $gp(p)$ of $\mathbb{Z}$ is maximal. Thus, $\Phi(\mathbb{Z}) = \{0\}$.
2. If p is a prime and $G = gp(g)$ is a cyclic group of order p^2, then $\Phi(G) = gp(g^p)$.
3. $\Phi(S_3) = \{e\}$. To see this, notice that the distinct subgroups

$$H = gp((1 \quad 2)) \text{ and } K = gp((1 \quad 2 \quad 3))$$

of S_3 each have prime index in S_3. Thus, H and K are maximal subgroups. Therefore, $\Phi(S_3)$ is a subgroup of $H \cap K = \{e\}$.
4. This example appears in [8]. Let p be a prime and let U_{ip} be the subgroup of $UT_n(\mathbb{Z})$ consisting of all matrices $A = (a_{ij})$ whose superdiagonal entries $a_{i,\, i+1}$ are contained in $\langle p \rangle$ for $i = 1, 2, \ldots, n-1$. Then U_{ip} is maximal in $UT_n(\mathbb{Z})$. Since the intersection of all U_{ip} lies in $UT_n^2(\mathbb{Z})$, we have $\Phi(UT_n(\mathbb{Z})) \leq UT_n^2(\mathbb{Z})$.

Another way to define the Frattini subgroup of a group is in terms of its set of non-generators.

Definition 7.6 Let G be a group. An element $g \in G$ is called a *non-generator* of G if $G = gp(g, X)$ implies that $G = gp(X)$ whenever $X \subset G$.

Thus, the set of non-generators of a group are precisely the elements that can be excluded from any generating set.

Theorem 7.15 (G. Frattini) *If G is any group, then $\Phi(G)$ is the set of all non-generating elements of G.*

The proof uses the next lemma.

Lemma 7.8 *Let G be any group and suppose that $H < G$ with $g \in G \setminus H$. There exists a subgroup $K < G$ which is maximal in G with respect to the properties $H \leq K$ and $g \notin K$.*

Proof Let $R = \{J < G \mid H \leq J$ and $g \notin J\}$. Clearly, $R \neq \emptyset$ since $H \in R$. Furthermore, R is partially ordered by inclusion and the union of any chain in R is again in R. By Zorn's lemma, R has a maximal element. □

We now prove Theorem 7.15. Let $g \in \Phi(G)$. We prove by contradiction that g is a non-generating element. Assume that there exists $X \subset G$ such that $G = gp(g, X)$, but $G \neq gp(X)$. Clearly, $g \notin gp(X)$. By Lemma 7.8, there exists a subgroup M of G which is maximal in G with respect to the properties $gp(X) \leq M$ and $g \notin M$. If $M < H \leq G$ for some H, then $g \in H$. Consequently, $H = G$. Hence, M is maximal in G. This implies that $g \in \Phi(G) \leq M$, which is a contradiction.

Conversely, let g be a non-generator of G. We prove by contradiction that $g \in \Phi(G)$. Suppose on the contrary, that $g \notin \Phi(G)$. There exists a maximal subgroup M of G for which $g \notin M$. Hence, $M \neq gp(g, M)$, and thus $G = gp(g, M)$ by the maximality of M in G. Since g is a non-generator of G, we have $G = M$, a contradiction. This completes the proof of Theorem 7.15.

The next two corollaries are immediate.

Corollary 7.5 *If G is a finitely generated group and $G = H\Phi(G)$ for some $H \leq G$, then $G = H$.*

Corollary 7.6 *If G is any group and $\Phi(G)$ is finitely generated, then the only subgroup H of G such that $G = H\Phi(G)$ is $H = G$.*

Lemma 7.9 *Let G be any group and suppose that H is a finitely generated subgroup of G. If $N \lhd G$ and $N \leq \Phi(H)$, then $N \leq \Phi(G)$.*

Proof Assume on the contrary, that N is not a subgroup of $\Phi(G)$. There exists a maximal subgroup M of G that does not contain N as a subgroup and thus, satisfies $G = MN$. Hence,

$$H = H \cap G = H \cap (MN) = (H \cap M)N.$$

Since $N \leq \Phi(H)$, we have $H = (H \cap M)\Phi(H)$. By Corollary 7.5, $H = H \cap M$, and thus $H \leq M$. However, H contains N. And so, $N \leq M$, a contradiction. □

Corollary 7.7 *If G is any group and H is a finitely generated normal subgroup of G, then $\Phi(H) \leq \Phi(G)$.*

Proof By Lemma 1.8, $\Phi(H) \trianglelefteq G$ because $\Phi(H)$ is characteristic in H and $H \trianglelefteq G$. Put $N = \Phi(H)$ in Lemma 7.9. □

The next two lemmas deal with the Frattini subgroup of a direct product. We give the proofs which appear in [3].

Lemma 7.10 *If $G = H \times K$, then $\Phi(G) \leq \Phi(H) \times \Phi(K)$.*

Proof If M is a maximal subgroup of H, then $M \times K$ is a maximal subgroup of $H \times K$. Hence, $\Phi(G) \leq \Phi(H) \times K$. Similarly, $\Phi(G) \leq H \times \Phi(K)$. Therefore,

$$\Phi(G) \leq (\Phi(H) \times K) \cap (H \times \Phi(K)) = \Phi(H) \times \Phi(K),$$

as desired. □

Lemma 7.11 *If G is a finitely generated group where $G = H \times K$ for some subgroups H and K of G, then $\Phi(G) = \Phi(H) \times \Phi(K)$.*

Proof Clearly, H and K are finitely generated and normal in G since G is finitely generated and $G = H \times K$. Thus, $\Phi(H) \leq \Phi(G)$ and $\Phi(K) \leq \Phi(G)$ by Corollary 7.7. It follows that $\Phi(G) \geq \Phi(H) \times \Phi(K)$. Lemma 7.10 takes care of the reverse inclusion. □

Lemma 7.12 *Let G be any group and $N \leq G$. If $N \leq \Phi(G)$, then N is normal in G and $\Phi(G/N) = \Phi(G)/N$.*

Proof Since $\Phi(G)$ is characteristic in G and $N \leq \Phi(G)$, N must be normal in G. Furthermore, $N \leq \Phi(G)$ implies that N is contained in every maximal subgroup of G. If $gN \in \Phi(G/N)$, then $gN \in M/N$ for every maximal subgroup M/N in G/N. Thus, for each maximal subgroup M in G, there exists $h_m \in M$ such that $gN = h_m N$; that is, $gh_m^{-1} \in N$. It follows that $gN \in \Phi(G)/N$. We reverse the above steps to conclude that $\Phi(G)/N \leq \Phi(G/N)$. □

Theorem 7.16 *Let G be a finite group. If H is a normal subgroup of G containing $\Phi(G)$ and $H/\Phi(G)$ is nilpotent, then H is nilpotent.*

In particular, if G is finite and $G/\Phi(G)$ is nilpotent, then G is nilpotent.

Proof In light of Theorem 2.13, it is enough to prove that the Sylow subgroups of H are normal. Let P be a Sylow p-subgroup of H. Since $\Phi(G)$ is contained in H by hypothesis, $\Phi(G) \trianglelefteq H$. By Lemma 2.16 (ii), $P\Phi(G)/\Phi(G)$ is a Sylow p-subgroup of $H/\Phi(G)$. Since $H/\Phi(G)$ is nilpotent, its Sylow p-subgroups are normal by Theorem 2.13. Thus, $P\Phi(G)/\Phi(G) \trianglelefteq H/\Phi(G)$. Furthermore, $P\Phi(G)/\Phi(G)$ is characteristic in $H/\Phi(G)$. This is due to the fact that normal Sylow p-subgroups are unique by Corollary 2.8 and every automorphism maps a subgroup of a given order into a subgroup of the same order. Since $H/\Phi(G) \trianglelefteq G/\Phi(G)$, it follows that $P\Phi(G)/\Phi(G) \trianglelefteq G/\Phi(G)$, and thus $P\Phi(G) \trianglelefteq G$.

Now, P is a Sylow p-subgroup of $P\Phi(G)$ because it is a Sylow p-subgroup of H. Since $P\Phi(G) \trianglelefteq G$, Lemma 2.17 gives $N_G(P)P\Phi(G) = G$. However, $P \leq N_G(P)$, and thus $N_G(P)\Phi(G) = G$. By Theorem 7.15, $N_G(P) = G$. Therefore, $P \trianglelefteq G$, and consequently, $P \trianglelefteq H$. □

Setting $H = \Phi(G)$ in Theorem 7.16 gives:

Corollary 7.8 (G. Frattini) *If G is a finite group, then $\Phi(G)$ is nilpotent.*

Remark 7.6 If G is infinite, then $\Phi(G)$ need not be nilpotent. Consider, for instance, the wreath product $G = \mathbb{Z}_{p^\infty} \wr \mathbb{Z}_{p^\infty}$. By Remark 2.8, G is an infinite non-nilpotent group. Since G has no maximal subgroups, it coincides with its Frattini subgroup as we noted after Definition 7.5. Thus, $\Phi(G)$ is not nilpotent.

Theorem 7.17 (W. Gaschütz) *If G is a finite group, then $\gamma_2 G \cap Z(G) \leq \Phi(G)$.*

Proof It is enough to prove that $\gamma_2 G \cap Z(G)$ is contained in each maximal subgroup of G. If M is a maximal subgroup of G, then either $Z(G) \leq M$ or $G = MZ(G)$. If it is the case that $Z(G) \leq M$, then it is clear that $\gamma_2 G \cap Z(G) \leq M$. If, on the other hand, $G = MZ(G)$, then $M \trianglelefteq G$ and G/M is abelian. By Lemma 1.6, $\gamma_2 G \leq M$. Hence, $\gamma_2 G \cap Z(G) \leq M$. □

We now turn our attention to the Frattini subgroup of nilpotent groups. We prove an important result which states that the commutator subgroup of a nilpotent group G is a set of non-generating elements of G.

Theorem 7.18 (K. A. Hirsch) *Let G be a nilpotent group. If $H \leq G$ and $H\gamma_2 G = G$, then $H = G$.*

Proof The proof is done by induction on the class c of G. If $c = 1$, then $\gamma_2 G = 1$ and the result is immediate.

Suppose that the result holds for all nilpotent groups of class less than c, and consider the natural homomorphism $\pi : G \to G/\gamma_c G$. By Lemma 2.8, $G/\gamma_c G$ is nilpotent of class $c - 1$. If $H\gamma_2 G = G$, then $\pi(H)\pi(\gamma_2 G) = \pi(G)$ holds in $G/\gamma_c G$. By Lemma 2.5, $\pi(\gamma_2 G) = \gamma_2 \pi(G)$. Thus, $\pi(H) = \pi(G) = G/\gamma_c(G)$ by induction; that is, $H\gamma_c G = G$. Using the commutator calculus and the fact that $\gamma_c G \leq Z(G)$, we have

$$\begin{aligned}\gamma_2 G &= [H\gamma_c G,\ H\gamma_c G]\\ &= [H,\ H][H,\ \gamma_c G][\gamma_c G,\ H][\gamma_c G,\ \gamma_c G]\\ &= [H,\ H] = \gamma_2 H.\end{aligned}$$

Therefore, $\gamma_2 G = \gamma_2 H \leq H$, and thus $H = G$. □

Corollary 7.9 *If G is a nilpotent group, then $\gamma_2 G \leq \Phi(G)$.*

Proof This is immediate from Theorems 7.15 and 7.18. □

Remark 7.7 By Theorem 7.6, $\gamma_2 UT_n(\mathbb{Z}) = UT_n^2(\mathbb{Z})$. Thus, $\Phi(UT_n(\mathbb{Z})) = UT_n^2(\mathbb{Z})$ by Example 7.1 and Corollary 7.9.

Corollary 7.10 *Let G be a nilpotent group and $X \subseteq G$. If $\pi : G \to Ab(G)$ is the natural homomorphism, then $G = gp(X)$ if and only if $Ab(G) = gp(\pi(X))$.*

Proof Suppose that $Ab(G) = gp(\pi(X))$, and let $H = gp(X)$. Then $G = H\gamma_2 G$, and thus $G = H = gp(X)$ by Theorem 7.18. The converse is obvious. □

Corollary 7.11 *Let G and H be nilpotent groups. If $\varphi : G \to H$ is a homomorphism, then φ is surjective if and only if $\overline{\varphi} : Ab(G) \to Ab(H)$ is surjective.*

Proof Assume that $\overline{\varphi}$ is surjective and let $h \in H$. There exists an element $g\gamma_2 G$ in $Ab(G)$ such that

$$\overline{\varphi}(g\gamma_2 G) = \varphi(g)\gamma_2 H = h\gamma_2 H.$$

Hence, $h \in \varphi(G)\gamma_2 H$, and thus $H = \varphi(G)\gamma_2 H$. By Theorem 7.18, $H = \varphi(G)$. The converse is clear. □

For finite groups, the converse of Corollary 7.9 holds.

Theorem 7.19 (H. Wielandt) *If G is a finite group and $\gamma_2 G \leq \Phi(G)$, then G is nilpotent.*

Proof If $\gamma_2 G \leq \Phi(G)$, then $\gamma_2 G \leq M$ for every maximal subgroup M of G. By Lemma 1.6, each M is normal in G. Thus, G is nilpotent by Theorem 2.13. □

For any prime p, the quotient of a finite p-group by its Frattini subgroup has a nice structure. It is always abelian and each of its elements has order p. This motivates the next definition.

Definition 7.7 Let p be a prime. A group G is called an *elementary abelian p-group* if it is abelian and every element of G has order p.

Clearly, every finite elementary abelian p-group is isomorphic to a direct product of n copies of $\mathbb{Z}_p$ for some $n < \infty$. Moreover, the Frattini subgroup of such a group is always trivial.

Lemma 7.13 *If G is a finite elementary abelian p-group, then $\Phi(G) = 1$.*

Proof Suppose that G is a direct product of n copies of $\mathbb{Z}_p$. The subgroup

$$G_i = \{(g_1, \ldots, g_{i-1}, 1, g_{i+1}, \ldots, g_n) \mid g_j \in \mathbb{Z}_p\}$$

is maximal in G, and $\cap_{i=1}^{n} G_i = 1$. Hence, $\Phi(G) \leq 1$. □

Lemma 7.14 *If G is a finite p-group and $H \leq G$, then $\Phi(G) \leq H$ if and only if $H \trianglelefteq G$ and G/H is an elementary abelian p-group.*

Proof Suppose that $H \trianglelefteq G$ and G/H is an elementary abelian p-group. Then $\Phi(G/H) = H$ by Lemma 7.13. Thus, $\Phi(G) \leq H$ by Lemma 7.12.

Conversely, suppose that $\Phi(G) \leq H$. By Corollary 7.9, $\gamma_2 G \leq \Phi(G)$. Thus, $H \trianglelefteq G$ and G/H is abelian by Lemma 1.6. Suppose that $g \in G$, and let M be any maximal proper subgroup of G. By Theorems 2.3 and 2.13 (iv), $M \triangleleft G$. Furthermore, $(gM)^p = M$ because $|G/M| = p$. Hence, $g^p \in M$, and thus g^p lies in all maximal proper subgroups of G. Therefore, $g^p \in \Phi(G)$ and G/H is an elementary abelian p-group since $\Phi(G) \leq H$. □

Setting $H = \Phi(G)$ in Lemma 7.14 proves our earlier remark:

Lemma 7.15 *If G is a finite p-group, then $G/\Phi(G)$ is an elementary abelian p-group.*

Lemma 7.16 *If G is a finite p-group, then*

$$\Phi(G) = G^p \gamma_2 G = gp(a^p, [b, c] \mid a, b, c \in G).$$

Hence, $N = \Phi(G)$ is the smallest normal subgroup of a finite p-group G such that G/N is an elementary abelian p-group.

Proof Since G is a finite p-group, $G/\Phi(G)$ is elementary abelian by Lemma 7.15. It follows from this and Lemma 1.6 (ii) that $g^p \in \Phi(G)$ for all $g \in G$, and $\gamma_2 G \leq \Phi(G)$. Therefore, $G^p\gamma_2 G \leq \Phi(G)$.

We establish the reverse inclusion. By Lemma 1.6 (i), $G/G^p\gamma_2 G$ is abelian. Since $G^p \leq G^p\gamma_2 G$, the factor group $G/G^p\gamma_2 G$ is a finite elementary abelian p-group. By the previous observation, $G^p\gamma_2 G \leq \Phi(G)$. Thus,

$$\Phi(G/G^p\gamma_2 G) = \Phi(G)/G^p\gamma_2 G = G^p\gamma_2 G$$

by Lemmas 7.12 and 7.13. And so, $G^p\gamma_2 G \geq \Phi(G)$. □

By Lemma 7.15, the quotient $G/\Phi(G)$ can be viewed as a vector space over $\mathbb{F}_p$, the finite field containing p elements. Furthermore, any set of group generators for $G/\Phi(G)$ is also a spanning set of $G/\Phi(G)$ as a vector space over $\mathbb{F}_p$.

Theorem 7.20 (Burnside's Basis Theorem) *Let G be a finite p-group. Consider $\overline{G} = G/\Phi(G)$ to be a vector space over $\mathbb{F}_p$, and suppose that $[G : \Phi(G)] = p^d$.*

(i) The dimension of $\overline{G}$ over $\mathbb{F}_p$ is d.
(ii) If $G = gp(g_1, \ldots, g_k)$, then $k \geq d$. More generally, $G = gp(g_1, \ldots, g_k)$ if and only if $\overline{G} = span\{g_1\Phi(G), \ldots, g_k\Phi(G)\}$.
(iii) G can be generated by exactly d elements. Furthermore, the set $\{g_1, \ldots, g_d\}$ generates G if and only if the set $\{g_1\Phi(G), \ldots, g_d\Phi(G)\}$ is a basis for $\overline{G}$ over $\mathbb{F}_p$.

Proof

(i) The dimension of a vector space over $\mathbb{F}_p$ is d if and only if it contains precisely p^d elements.
(ii) Since $G = gp(g_1, \ldots, g_k)$, the vectors $g_1\Phi(G), \ldots, g_k\Phi(G)$ span $\overline{G}$. By (i), the dimension of $\overline{G}$ over $\mathbb{F}_p$ is d. And so, $k \geq d$.
 If $\overline{G} = span\{g_1\Phi(G), \ldots, g_k\Phi(G)\}$, then $G = gp(g_1, \ldots, g_k, \Phi(G))$. By Theorem 7.15, $G = gp(g_1, \ldots, g_k)$.
(iii) If $G = gp(g_1, \ldots, g_d)$, then $\overline{G}$ is spanned by $g_1\Phi(G), \ldots, g_d\Phi(G)$ by (ii). Since $G/\Phi(G)$ has dimension d by (i), the vectors $g_1\Phi(G), \ldots, g_d\Phi(G)$ are linearly independent. Thus, $\{g_1\Phi(G), \ldots, g_d\Phi(G)\}$ is a basis for $\overline{G}$. The converse follows from (ii). □

Lemma 7.17 *If G is a finite nilpotent group, then $G/\Phi(G)$ is a direct product of elementary abelian p-groups.*

Proof Let P_i denote the Sylow p_i-subgroup of G for $i = 1, \ldots, k$. Since G is finite and nilpotent, $G \cong P_1 \times \cdots \times P_k$ by Theorem 2.13. Thus,

$$\frac{G}{\Phi(G)} \cong \frac{P_1 \times \cdots \times P_k}{\Phi(P_1 \times \cdots \times P_k)} = \frac{P_1}{\Phi(P_1)} \times \cdots \times \frac{P_k}{\Phi(P_k)},$$

where the last equality follows from Lemma 7.11. Since each P_i is a finite p_i-group, each factor group $P_i/\Phi(P_i)$ is an elementary abelian p_i-group by Lemma 7.15. □

7.5.2 The Fitting Subgroup

Definition 7.8 The *Fitting subgroup* of a group G, denoted by $Fit(G)$, is the subgroup of G generated by all of the normal nilpotent subgroups of G.

Clearly, $Fit(G)$ is a characteristic subgroup of G, and thus $Fit(G) \trianglelefteq G$.

Lemma 7.18 *If G is a finite group, then $Fit(G)$ is the nilpotent radical of G.*

Proof The result follows at once from Theorem 2.11. □

Remark 7.8 If G is an infinite group, then $Fit(G)$ is not necessarily nilpotent. For example, choose any prime p, and let A be a countably infinite elementary abelian p-group. One can show that the group $G = \mathbb{Z}_p \wr A$ is not nilpotent and $G = Fit(G)$. See pp. 3–4 in [10] for details.

Lemma 7.19 *If G is a nontrivial finite group, then the centralizer of $Fit(G)$ in G contains every minimal normal subgroup of G.*

Proof Set $F = Fit(G)$, and let N be a minimal normal subgroup of G. There are two cases to consider.

- If N is not a subgroup of F, then $F \cap N = 1$ because $F \cap N$ is a proper subgroup of N and $F \cap N \trianglelefteq G$. By Theorem 1.4, $[F, N] = 1$. And so, $N \leq C_G(F)$.
- If $N \leq F$, then there is a minimal normal subgroup M of F with $M \leq N$. By Theorem 2.29 and Lemma 7.18, we have $M \leq Z(F)$. Hence, $Z(F) \cap N \neq 1$. However, $Z(F) \trianglelefteq G$, and thus $Z(F) \cap N \trianglelefteq G$. Consequently, $Z(F) \cap N = N$ because N is a minimal normal subgroup of G. Therefore, $N \leq Z(F) \leq C_G(F)$. □

Another description of $Fit(G)$ for a finite group G is in terms of its chief factors.

Theorem 7.21 *If G is a finite group, then $Fit(G)$ is the intersection of the centralizers of the chief factors of G.*

Proof We adopt the proof from [12]. Set $F = Fit(G)$, and let

$$1 = G_0 \leq G_1 \leq \cdots \leq G_n = G$$

be a chief series of G. Set

$$I = \bigcap_{i=0}^{n-1} C_G(G_{i+1}/G_i).$$

(The notation used above can be found in Definition 5.6.) It is clear that $I \trianglelefteq G$. Furthermore, $[G_{i+1}, I] \leq G_i$ for each $i = 0, 1, \ldots, n-1$. Consequently, I is nilpotent and thus, $I \leq F$.

To show that $F \leq I$, we prove that $F \leq C_G(G_{i+1}/G_i)$ for each $i = 0, 1, \ldots, n-1$. Consider the factor group $FG_i/G_i \leq G/G_i$. Clearly, FG_i/G_i is normal in G/G_i and $F \cap G_i \trianglelefteq F$. By the Second Isomorphism Theorem, $FG_i/G_i \cong F/(F \cap G_i)$.

Since F is nilpotent by Lemma 7.18, FG_i/G_i is a normal nilpotent subgroup of G/G_i by Corollary 2.5. Hence, $FG_i/G_i \leq Fit(G/G_i)$. Now, G_{i+1}/G_i is a minimal normal subgroup of G/G_i by Lemma 2.23. According to Lemma 7.19, it follows that $Fit(G/G_i)$ must centralize G_{i+1}/G_i. This means that

$$FG_i/G_i \leq C_{G/G_i}(G_{i+1}/G_i),$$

or equivalently, $F \leq C_G(G_{i+1}/G_i)$. □

We record a couple of properties of the Fitting subgroup of a finite group.

Proposition 7.2 *Let G be a finite group.*

(i) If $H \trianglelefteq G$, then $H \cap Fit(G) = Fit(H)$.
(ii) If $K \leq Z(G)$, then $Fit(G/K) = Fit(G)/K$.

Proof

(i) By Lemma 1.8, $Fit(H)$ is a normal subgroup of G because it is characteristic in H. Since $Fit(H)$ is also nilpotent by Lemma 7.18, $Fit(H) \leq H \cap Fit(G)$. Furthermore, $H \cap Fit(G) \trianglelefteq H$ because $Fit(G) \trianglelefteq G$. Since $H \cap Fit(G)$ is nilpotent, we have $H \cap Fit(G) \trianglelefteq Fit(H)$. Hence, $H \cap Fit(G) = Fit(H)$.
(ii) Set $M/K = Fit(G/K)$. We claim that $M = Fit(G)$. By Lemma 7.18, M/K is nilpotent. Thus, M is nilpotent by Theorem 2.6 because K is central in M. Moreover, $M \trianglelefteq G$ since M/K is characteristic, hence normal, in G/K. Thus, $M \leq Fit(G)$. Now, since $Fit(G)$ is nilpotent and $K \leq Fit(G)$, it must be the case that $Fit(G)/K$ is nilpotent. Since $Fit(G)/K$ is normal in G/K, we have

$$Fit(G)/K \leq Fit(G/K) = M/K.$$

Consequently, $Fit(G) \leq M$, and thus $M = Fit(G)$. □

There are natural connections between the Fitting and Frattini subgroups. We state two of them in our last theorem.

Theorem 7.22 (W. Gaschütz) *If G is a finite group, then $\Phi(G) \trianglelefteq Fit(G)$ and $Fit(G/\Phi(G)) = Fit(G)/\Phi(G)$.*

Proof The fact that $\Phi(G) \trianglelefteq Fit(G)$ is immediate from Corollary 7.8. We prove the second assertion. Clearly, $Fit(G)/\Phi(G) \leq Fit(G/\Phi(G))$ since $Fit(G)/\Phi(G)$ is a normal nilpotent subgroup of $G/\Phi(G)$. Suppose that $H/\Phi(G)$ is a normal nilpotent subgroup of $G/\Phi(G)$. Then H is a normal nilpotent subgroup of G by Theorem 7.16. Consequently, $H \leq Fit(G)$, and thus $Fit(G/\Phi(G)) \leq Fit(G)/\Phi(G)$. Therefore, $Fit(G/\Phi(G)) = Fit(G)/\Phi(G)$. □

References

1. G. Baumslag, *Lecture Notes on Nilpotent Groups*. Regional Conference Series in Mathematics, No. 2 (American Mathematical Society, Providence, RI, 1971). MR0283082
2. N. Blackburn, Conjugacy in nilpotent groups. Proc. Am. Math. Soc. **16**, 143–148 (1965). MR0172925
3. J.D. Dixon, *Problems in Group Theory* (Dover, New York, 1973)
4. F. Grunewald, D. Segal, Some general algorithms. II. Nilpotent groups. Ann. Math. (2) **112**(3), 585–617 (1980). MR0595207
5. P. Hall, Some sufficient conditions for a group to be nilpotent. Ill. J. Math. **2**, 787–801 (1958). MR0105441
6. P. Hall, *The Edmonton Notes on Nilpotent Groups*. Queen Mary College Mathematics Notes. Mathematics Department (Queen Mary College, London, 1969). MR0283083
7. L. Kalužnin, Über gewisse Beziehungen zwischen einer Gruppe und ihren Automorphismen. Berliner Mathematische Tagung, pp. 164–172 (1953)
8. M.I. Kargapolov, J.I. Merzljakov, *Fundamentals of the Theory of Groups*. Graduate Texts in Mathematics, vol. 62 (Springer, New York, 1979). MR0551207. Translated from the second Russian edition by Robert G. Burns
9. R.C. Lyndon, P.E. Schupp, *Combinatorial Group Theory* (Springer, Berlin, 1977). MR0577064
10. D.J.S. Robinson, *Finiteness Conditions and Generalized Soluble Groups. Part 2* (Springer, New York, 1972). MR0332990
11. D.J.S. Robinson, *A Course in the Theory of Groups*. Graduate Texts in Mathematics, vol. 80, 2nd edn. (Springer, New York, 1996). MR1357169
12. J.S. Rose, *A Course on Group Theory* (Dover Publications, New York, 1994). MR1298629. Reprint of the 1978 original [Dover, New York; MR0498810 (58 #16847)]
13. M. Zyman, *IA-Automorphisms and Localization of Nilpotent Groups* (ProQuest LLC, Ann Arbor, MI, 2007). MR2711065