9

Polycyclic groups

Polycyclic groups form a broad class of finitely presented groups in which extensive computation is possible. This chapter discusses the basic structure and properties of polycyclic groups and presents algorithms for computing with elements and subgroups of polycyclic groups. All finite solvable groups are polycyclic. The literature on algorithms for computing with finite solvable groups is too extensive to cover in detail here. See (Laue, Neubüser, & Shoenwaelder 1984), (Mecky & Neubüser 1989), and (Glasby & Slattery 1990). Emphasis will be placed on those algorithms which apply to infinite as well as finite polycyclic groups. Various computations with polycyclic groups have been shown to be possible in principle but have not yet been shown to be practical for interesting groups. Some of these algorithms will be mentioned, but details will not be given. By combining the rewriting techniques of Chapter 2 with the methods developed in this chapter, one obtains algorithms for solving a wide range of problems in *polycyclic-by-finite* groups. These are groups which have a polycyclic subgroup of finite index. The polycyclic-by-finite groups make up the largest class of finitely presented groups in which most computational problems concerning elements and subgroups have algorithmic solutions.

A word of caution is in order. In coset enumeration, right cosets of subgroups are used almost exclusively. However, in studying polycyclic groups it is traditional to use left cosets. It would be possible to present both subjects using the same type of cosets, but it seems best to remain consistent with other authors.

This chapter assumes that the reader has had a good introduction to the theory of groups. The material required is summarized, but for the most part proofs are omitted. This is particularly true for Sections 9.1 and 9.2, which review results about commutator subgroups and elementary properties of solvable and nilpotent groups. Details can be found in most standard texts on group theory.

9.1 Commutator subgroups

Let h and k be elements of a group G. The *commutator* of h and k is the element $[h,k] = h^{-1}k^{-1}hk$. The conjugate $h^k = k^{-1}hk$ of h by k is $h[h,k]$, and $hk = kh[h,k]$. Thus h and k commute if and only if their commutator is trivial. Suppose H and K are subgroups of G. The *commutator subgroup* of H and K is the group

$$[H,K] = \mathrm{Grp}\,\langle [h,k] \mid h \in H, k \in K \rangle\,.$$

Since $[k,h] = [h,k]^{-1}$, it follows that $[K,H] = [H,K]$. The *commutator subgroup* of G is $[G,G]$, which is also called the *derived subgroup* of G and denoted G'. The derived subgroup of G is trivial if and only if G is abelian.

Proposition 1.1. *The derived subgroup G' is normal in G and the quotient G/G' is abelian. If N is any normal subgroup of G such that G/N is abelian, then $N \supseteq G'$.*

Thus G/G' is the largest abelian quotient group of G. The following propositions present some more basic facts about commutator subgroups.

Proposition 1.2. *Suppose that H_1, K_1, H_2, and K_2 are subgroups of G such that $H_1 \subseteq H_2$ and $K_1 \subseteq K_2$. Then $[H_1, K_1] \subseteq [H_2, K_2]$.*

Proposition 1.3. *If $f\colon G \to Q$ is a homomorphism of groups, then $[f(H), f(K)] = f([H,K])$ for all subgroups H and K of G.*

Corollary 1.4. *Suppose that N is a normal subgroup of G and $\bar{\ }$ is the natural homomorphism from G to G/N. Then $[\overline{H}, \overline{K}] = \overline{[H,K]}$.*

Proposition 1.5. *If H and K are normal subgroups of G, then $[H,K]$ is normal in G and $[H,K] \subseteq H \cap K$.*

If x, y, and z are elements of G, then $[x,y,z]$ is defined to be $[[x,y],z]$. In general, $[x_1, x_2, \ldots, x_n]$ is defined recursively to be $[[x_1, x_2, \ldots, x_{n-1}], x_n]$. Such commutators are said to be *left-normed.* If $H_1, \ldots, H_n$ are subgroups of G, then $[H_1, \ldots, H_n] = [[H_1, \ldots, H_{n-1}], H_n]$.

Proposition 1.6. *For any elements x, y, and z of G, the following hold:*

(a) $[xy, z] = [x,z][x,z,y][y,z]$.
(b) $[x, yz] = [x,z][x,y][x,y,z]$.
(c) $[x, y^{-1}, z]^y [y, z^{-1}, x]^z [z, x^{-1}, y]^x = 1$.

Proposition 1.7. *If H, K, and L are subgroups of G and N is a normal subgroup of G containing $[K,L,H]$ and $[L,H,K]$, then N contains $[H,K,L]$.*

Proof. In view of Corollary 1.4, we may pass to the quotient G/N and assume that N is trivial. Thus it suffices to prove that $[H,K,L]=1$ whenever $[K,L,H]=[L,H,K]=1$.

To show that $[H,K,L]=1$, we must prove that each element of $[H,K]$ commutes with each element of L. For this it suffices to prove that each element in a generating set for $[H,K]$ commutes with each element of L. Thus it is enough to show that $[x,y,z]=1$ for all x in H, y in K, and z in L. By Proposition 1.6(c),

$$[x,y,z]^{y^{-1}}[y^{-1},z^{-1},x]^z[z,x^{-1},y^{-1}]^x=1.$$

But $[y^{-1},z^{-1},x]$ is in $[K,L,H]$ and $[z,x^{-1},y^{-1}]$ is in $[L,H,K]$. Therefore

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

and hence $[x,y,z]^{y^{-1}}=1$. Conjugating by y, we obtain $[x,y,z]=1$. □

Proposition 1.7 is sometimes called the three subgroups lemma.

There are many ways in which one could form "higher commutator subgroups" in G. Examples of such subgroups are $(G')'=[G',G']=[[G,G],[G,G]]$ and $[G',G]=[G,G,G]$. Two sequences of these higher commutator subgroups have been found to be particularly useful. The *derived series* of G is obtained by taking successive derived subgroups. Thus we have $G^{(0)}=G$, $G^{(1)}=G'$, $G^{(2)}=(G')'$, and, in general,

$$G^{(i+1)}=(G^{(i)})'=[G^{(i)},G^{(i)}].$$

The *lower central series* of G is defined by taking successive commutator subgroups with G. Here $\gamma_1(G)=G$, $\gamma_2(G)=G'=G^{(1)}$, and, in general,

$$\gamma_{i+1}(G)=[\gamma_i(G),G].$$

By Proposition 1.5, all of the terms in the derived series and the lower central series are normal in G. In addition, $G^{(0)}\supseteq G^{(1)}\supseteq G^{(2)}\supseteq\cdots$ and $\gamma_1(G)\supseteq\gamma_2(G)\supseteq\gamma_3(G)\supseteq\cdots$.

Proposition 1.8. *Suppose that $f\colon G\to Q$ is a homomorphism of groups. Then $f(G)^{(i)}=f(G^{(i)})$ and $\gamma_i(f(G))=f(\gamma_i(G))$.*

Proof. Induction and Proposition 1.3. □

Proposition 1.9. *If H is a subgroup of G, then $H^{(i)} \subseteq G^{(i)}$ and $\gamma_i(H) \subseteq \gamma_i(G)$.*

Proof. Induction and Proposition 1.2. □

Proposition 1.10. *For all $i \geq 1$ and $j \geq 1$, $[\gamma_i(G), \gamma_j(G)] \subseteq \gamma_{i+j}(G)$.*

Proof. For $j = 1$ we have the definition of $\gamma_{i+1}(G)$. We proceed by induction on j. Since

$$[\gamma_i(G), \gamma_j(G)] = [\gamma_i(G), [\gamma_{j-1}(G), G]] = [\gamma_{j-1}(G), G, \gamma_i(G)],$$

by Proposition 1.7 it suffices to prove that $[G, \gamma_i(G), \gamma_{j-1}(G)]$ and $[\gamma_i(G), \gamma_{j-1}(G), G]$ are contained in $\gamma_{i+j}(G)$. But $[G, \gamma_i(G), \gamma_{j-1}(G)] = [\gamma_{i+1}(G), \gamma_{j-1}(G)]$ is contained in $\gamma_{i+j}(G)$ by induction. Also by induction, $[\gamma_i(G), \gamma_{j-1}(G)] \subseteq \gamma_{i+j-1}(G)$. Therefore

$$[\gamma_i(G), \gamma_{j-1}(G), G] \subseteq [\gamma_{i+j-1}(G), G] = \gamma_{i+j}(G). \quad \square$$

Corollary 1.11. *$G^{(i)} \subseteq \gamma_{2^i}(G)$ for $i \geq 0$.*

Proof. For $i = 0$ we have $G^{(0)} = G = \gamma_1(G)$. If $i > 1$, then

$$G^{(i)} = [G^{(i-1)}, G^{(i-1)}] \subseteq [\gamma_{2^{i-1}}(G), \gamma_{2^{i-1}}(G)] \subseteq \gamma_{2^{i-1}+2^{i-1}}(G) = \gamma_{2^i}(G). \quad \square$$

Exercises

1.1. Suppose that G is a group generated by a set X. Show that $G' = \gamma_2(G)$ is the normal closure in G of the set $\{[x, y] \mid x, y \in X\}$.

1.2. Generalize Exercise 1.1 by showing that $\gamma_s(G)$ is the normal closure in G of the set $\{[x_1, \ldots, x_s] \mid x_i \in X, 1 \leq i \leq s\}$ for $s \geq 2$.

1.3. Let F be the free group on $X = \{a, b\}$, $a \neq b$. Show that every commutator $[[x_1, x_2], [x_3, x_4]]$ with each x_i in X is trivial but that $F^{(2)}$ is not trivial.

9.2 Solvable and nilpotent groups

Let G be a group. We say that G is *solvable* if $G^{(i)}$ is trivial for some $i \geq 0$. If G is solvable, then the smallest value of i for which $G^{(i)} = 1$ is called the *derived length* of G. Groups of order 1 have derived length 0. Nontrivial abelian groups have derived length 1. The derived length of G is 2 if and only if G' is nontrivial and abelian. A group with derived length at most 2 is called *metabelian*.

Our group G is *nilpotent* if some term in the lower central series is trivial. In this case, the smallest integer c such that $\gamma_{c+1}(G) = 1$ is called the *nilpotency class*, or simply the *class*, of G. Trivial groups have class 0 and nontrivial abelian groups have class 1.

Proposition 2.1. *Subgroups and quotient groups of solvable groups are solvable. Subgroups and quotient groups of nilpotent groups are nilpotent.*

Proof. Let H and N be subgroups of G with N normal. Suppose $G^{(i)} = 1$. Then $H^{(i)} = 1$ by Proposition 1.9 and $(G/N)^{(i)} = 1$ by Proposition 1.8. Similarly, if $\gamma_j(G) = 1$, then $\gamma_j(H) = 1$ and $\gamma_j(G/N) = 1$. □

Proposition 2.2. *The group $G/G^{(i)}$ is solvable with derived length at most i. The group $G/\gamma_j(G)$ is nilpotent of class at most $j-1$.*

Proof. By Proposition 1.9, $(G/G^{(i)})^{(i)} = G^{(i)}/G^{(i)} = 1$. Similarly $\gamma_j(G/\gamma_j(G)) = \gamma_j(G)/\gamma_j(G) = 1$. □

Proposition 2.3. *If N is a normal subgroup of G and both N and G/N are solvable, then G is solvable.*

Proof. Since G/N is solvable, there is an integer i such that $(G/N)^{(i)} = (G^{(i)}N)/N = 1$. This means that $G^{(i)} \subseteq N$. There is an integer j such that $N^{(j)} = 1$. Hence $G^{(i+j)} = (G^{(i)})^{(j)} \subseteq N^{(j)} = 1$. □

At this point nilpotent groups differ from solvable groups. Proposition 2.3 is false if "nilpotent" is substituted for "solvable". Let G be the symmetric group Sym(3), which has order 6, and let N be the alternating subgroup of G. Then N is generated by the 3-cycle (1,2,3). Both N and G/N are abelian and hence nilpotent. Now $[(1,2,3),(1,2)] = (1,2,3)$. Therefore $N = G' = [N,G]$. It follows easily that $N = \gamma_i(G)$ for all $i \geq 2$. Hence G is not nilpotent.

Proposition 2.4. *Nilpotent groups are solvable.*

Proof. Suppose $\gamma_j(G) = 1$. By Corollary 1.11, $G^{(i)} = 1$ provided $2^i \geq j$. □

The symmetric group of degree 3 is an example of a solvable group which is not nilpotent.

Example 2.1. Let R be a commutative ring with $1 \neq 0$ and let n be a positive integer. For $r \geq 1$ let $U_n^{(r)}(R)$ consist of those n-by-n matrices A over R such that $A_{ij} = 0$ if $j < i+r$. Thus A is in $U_n^{(r)}(R)$ if and only if all entries on or below the main diagonal are 0 and all entries on the first $r-1$ diagonals above the main diagonal are also 0. Thus an element of $U_4^{(2)}(R)$ has the form

$$\begin{bmatrix} 0 & 0 & * & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $*$ denotes any element of R. If $r \geq n$, then $U_n^{(r)}(R)$ contains only the 0 matrix. Suppose that A is in $U_n^{(r)}(R)$ and B is in $U_n^{(s)}(R)$. A simple argument shows that AB is in $U_n^{(r+s)}(R)$. In particular, $A^n = 0$. Let $D_n^{(r)}(R)$ consist of all matrices $I + A$, where I is the n-by-n identity matrix and A is in $U_n^{(r)}(R)$. Since

$$(I + A)(I + B) = I + A + B + AB$$

and

$$(I + A)^{-1} = I - A + A^2 - \cdots + (-A)^{n-1},$$

the sets $D_n^{(r)}(R)$ are subgroups of $\mathrm{GL}(n, R)$. If A is in $U_n^{(r)}(R)$ and B is in $U_n^{(s)}(R)$, then the commutator $[I + A, I + B]$ is

$$(I - A + \cdots + (-A)^{n-1})(I - B + \cdots + (-B)^{n-1})(I + A)(I + B).$$

Direct computation shows that this matrix has the form $I + C$, where C is in $U_n^{(r+s)}$. Thus $[D_n^{(r)}(R), D_n^{(s)}(R)]$ is contained in $D_n^{(r+s)}(R)$. Therefore, if $D = D_n^{(1)}(R)$, then $\gamma_r(D) \subseteq D_n^{(r)}(R)$. In particular, $\gamma_n(D)$ is trivial, so D is nilpotent of class at most $n - 1$.

The *quotients* of the derived series of a group G are the groups $G^{(i)}/G^{(i+1)}$, $i = 0, 1, \ldots$. They are all abelian groups. The quotients of the lower central series of G are the groups $\gamma_i(G)/\gamma_{i+1}(G), i = 1, 2, \ldots$. These groups are also abelian. In fact $\gamma_i(G)/\gamma_{i+1}(G)$ is in the center of $G/\gamma_{i+1}(G)$.

If G is finitely generated, then $G/G' = G/\gamma_2(G)$ is finitely generated. However, the remaining quotients in the derived series may not be finitely generated, as the following example illustrates.

Example 2.2. Let x be the permutation of the integers which maps i to $i+2$ for all i. Thus x has "cycles" $(\ldots, -3, -1, 1, 3, 5, \ldots)(\ldots, -4, -2, 0, 2, 4, \ldots)$. Let $y = (0, 1)$ and let G be the group of permutations of the integers generated by x and y. The conjugates of y by powers of x are the 2-cycles $y_i = (2i, 2i+1), i = 0, \pm 1, \ldots$. The subgroup N generated by the y_i is abelian and normal in G. The commutator $[y, x]$ is $z = (0, 1)(2, 3)$. The conjugates of z by powers of x are the elements $z_i = (2i, 2i+1)(2i+2, 2i+3) = y_i y_{i+1}$. The subgroup M generated by the z_i is normal. By Exercise 1.1, $M = G'$. Each element of M moves at most a finite number of integers. Therefore any finitely generated subgroup of M moves only finitely many integers. However, M moves all the integers, so M is not finitely generated. Since N is abelian, so is M. Therefore $G'' = 1$ and G'/G'' is not finitely generated.

Again we have a difference between solvable and nilpotent groups.

Proposition 2.5. *If G is generated modulo G' by $x_1, \ldots, x_n$, then $\gamma_2(G)/\gamma_3(G)$ is generated by the images of $[x_j, x_i]$ with $1 \le i < j \le n$.*

Proof. We may work in $G/\gamma_3(G)$ and so we may assume $\gamma_3(G) = 1$. Thus elements of G' are in the center of G. By Proposition 1.6, $[xy, z] = [x, z][y, z]$ and $[x, yz] = [x, y][x, z]$. By induction one concludes that

$$[y_1 \ldots y_s, z_1 \ldots z_t] = \prod_{i,j} [y_i, z_j].$$

Now $1 = [xx^{-1}, y] = [x, y][x^{-1}, y]$, so $[x^{-1}, y] = [x, y]^{-1}$. Similarly, $[x, y^{-1}] = [x, y]^{-1}$ and $[x^{-1}, y^{-1}] = [x, y]$. It follows that the commutator of any two elements of $H = \operatorname{Grp}\langle x_1, \ldots, x_n \rangle$ is in the subgroup generated by the commutators $[x_j, x_i]$. Since $[x_i, x_i] = 1$ and $[x_i, x_j] = [x_j, x_i]^{-1}$, we may assume that $i < j$. By assumption, any element of G can be written as uz, where u is in H and z is in G'. If v is also in H and w is in G', then $[uz, vw] = [u, v]$. Therefore $G' = H'$. $\square$

Proposition 2.6. *Suppose G is a group generated modulo G' by a set X and Y is a subset of $\gamma_i(G)$, $i \ge 2$, whose image in $\gamma_i(G)/\gamma_{i+1}(G)$ generates that group. Then $\gamma_{i+1}(G)/\gamma_{i+2}(G)$ is generated by the image of $Z = \{[y, x] \mid y \in Y, x \in X\}$.*

Proof. We may assume that $\gamma_{i+2}(G) = 1$. All of the elements of Z are in $\gamma_{i+1}(G)$. If u is in $\gamma_i(G)$ and v and w are in G, then by Proposition 1.6, $[u, vw] = [u, w][u, v][u, v, w]$. But $[u, v, w]$ is in $\gamma_{i+2}(G)$ and hence $[u, v, w]$ is trivial. Also $[[u, w], [u, v]]$ is in $\gamma_{2i+2}(G)$ and so is trivial. Therefore $[u, vw] = [u, v][u, w]$. If w is in G', then $[u, w]$ is in $\gamma_{i+2}(G)$ and so $[u, vw] = [u, v]$. By a similar argument one shows that $[uv, w] = [u, w][v, w]$ whenever u and v are in $\gamma_i(G)$ and w is in G. If v is in $\gamma_{i+1}(G)$, then $[uv, w] = [u, w]$. In addition $[u^{-1}, v] = [u, v]^{-1} = [u, v^{-1}]$ if u is in $\gamma_i(G)$ and v is in G. By an argument similar to the one in the proof of Proposition 2.5, $\gamma_{i+1}(G)$ is generated by the commutators $[u, v]$, where u ranges over a set of generators of $\gamma_i(G)$ modulo $\gamma_{i+1}(G)$ and v ranges over a set of generators of G modulo G'. $\square$

Corollary 2.7. *If a group G is generated by n elements and $i \geq 2$, then $\gamma_i(G)/\gamma_{i+1}(G)$ is generated by $(n-1)n^{i-1}/2$ elements.*

Proof. The case $i = 2$ follows from Proposition 2.5. Now induction on i and Proposition 2.6 complete the proof. □

The bound in Corollary 2.7 can be improved somewhat. See Section 9.9.

Proposition 2.8. *If N is a subgroup of the center of G and G/N is nilpotent, then G is nilpotent.*

Proof. By assumption, $\gamma_i(G)$ is contained in N for some i. Since N is central, $[N, G] = 1$. But $\gamma_{i+1}(G) = [\gamma_i(G), G]$ is contained in $[N, G] = 1$, so $\gamma_{i+1}(G) = 1$. □

Exercises

2.1. Suppose that G is a finitely generated nilpotent group. Prove that all of the terms in the lower central series of G are finitely generated.
2.2. Let G be a nilpotent group and let X be a subset of G whose image in G/G' generates G/G'. Show that X generates G.
2.3. Let $D_n^{(r)}(R)$ be as in Example 2.1. Show that $[D_n^{(r)}(R), D_n^{(s)}(R)] = D_n^{(r+s)}(R)$.

9.3 Polycyclic groups

The class of polycyclic groups contains the class of finitely generated nilpotent groups and is contained in the class of finitely generated solvable groups. Polycyclic groups have finite presentations of a form which makes many types of computations practical.

Let G be a group. A *polycyclic series* of length n for G is a sequence of subgroups

$$G = G_1 \supseteq G_2 \supseteq \cdots \supseteq G_n \supseteq G_{n+1} = 1$$

such that for $1 \leq i \leq n$ the subgroup G_{i+1} is normal in G_i and G_i/G_{i+1} is cyclic. Note that we do not require that each G_i be normal in G. A group is *polycyclic* if it has a polycyclic series.

Proposition 3.1. *Polycyclic groups are solvable.*

Proof. Let $G = G_1 \supseteq G_2 \supseteq \cdots \supseteq G_n \supseteq G_{n+1} = 1$ be a polycyclic series for a group G. We proceed by induction on n. If $n = 0$, then G is trivial and hence solvable. Assume that $n > 0$. Then G_2 is normal in G and G/G_2 is cyclic and therefore solvable. The sequence $G_2 \supseteq \cdots \supseteq G_n \supseteq G_{n+1} = 1$ is a polycyclic series for G_2 of length $n-1$. By induction, G_2 is solvable. Thus G is solvable by Proposition 2.3. □

Proposition 3.2. *Finitely generated abelian groups are polycyclic.*

Proof. Let G be an abelian group generated by $a_1, \ldots, a_n$. Set $G_i = \mathrm{Grp}\langle a_i, \ldots, a_n\rangle$, $1 \le i \le n+1$. Then G_{i+1} is contained in G_i and G_{i+1} is normal in G, so G_{i+1} is certainly normal in G_i. The quotient G_i/G_{i+1} is generated by the coset a_iG_{i+1}. Therefore G_i/G_{i+1} is cyclic and the G_i form a polycyclic series for G. □

Proposition 3.3. *If N is a normal subgroup of a group G and both N and G/N are polycyclic, then G is polycyclic.*

Proof. A subgroup of G/N has the form H/N, where H is a subgroup of G containing N. If K is another subgroup of G containing N, then $K/N \subseteq H/N$ if and only if $K \subseteq H$, and K/N is normal in H/N if and only if K is normal in H. In this case, $(H/N)/(K/N)$ is isomorphic to H/K. Thus we can pull back a polycyclic series of G/N to a sequence of subgroups of G from G to N such that each subgroup is normal in the preceding one and the quotients are cyclic. Following this by a polycyclic series for N, we get a polycyclic series for G. □

Proposition 3.4. *Finitely generated nilpotent groups are polycyclic.*

Proof. Let G be a finitely generated nilpotent group of class c. If $c \le 1$, then G is abelian, and hence G is polycyclic by Proposition 3.2. Assume that $c > 1$. The last nontrivial term in the lower central series of G is $\gamma_c(G)$. By Proposition 1.10 and Corollary 2.7, $\gamma_c(G)$ is abelian and finitely generated. Therefore $\gamma_c(G)$ is polycyclic. The quotient $G/\gamma_c(G)$ is nilpotent of class $c-1$. By induction on c, $G/\gamma_c(G)$ is polycyclic. Thus G is polycyclic by Proposition 3.3. □

Proposition 3.5. *If G is a group with a polycyclic series of length n, then G can be generated by n elements.*

Proof. Let $G = G_1 \supseteq \cdots \supseteq G_{n+1} = 1$ be a polycyclic series. For $1 \le i \le n$, let a_i be an element of G_i such that a_iG_{i+1} generates G_i/G_{i+1}. Then every coset of G_{i+1} in G_i contains a power of a_i. Thus if g is in G, then $g = a_1^{\alpha_1}g_2$, where g_2 is in G_2. But $g_2 = a_2^{\alpha_2}g_3$, where g_3 is in G_3. Therefore $g = a_1^{\alpha_1}a_2^{\alpha_2}g_3$. Continuing in this manner, we find that $g = a_1^{\alpha_1} \ldots a_n^{\alpha_n}g_{n+1}$, where g_{n+1} is in G_{n+1}. But $G_{n+1} = 1$, so $g = a_1^{\alpha_1} \ldots a_n^{\alpha_n}$. Hence G is generated by $a_1, \ldots, a_n$. □

Proposition 3.6. *Quotient groups of polycyclic groups are polycyclic.*

Proof. Let $G = G_1 \supseteq \cdots \supseteq G_{n+1} = 1$ be a polycyclic series for a group G and let N be a normal subgroup of G. For $1 \le i \le n+1$, the product G_iN

is a subgroup of G and $G_{i+1}N$ is normal in G_iN if $i \le n$. Also

$$G_iN/G_{i+1}N \cong G_i/G_{i+1}(G_i \cap N) \cong (G_i/G_{i+1})/(G_{i+1}(G_i \cap N)/G_{i+1})$$

is a quotient of the cyclic group G_i/G_{i+1}. Thus $G_iN/G_{i+1}N$ is cyclic.

Define H_i to be G_iN/N. Then $G/N = H_1 \supseteq \cdots \supseteq H_{n+1} = N/N = 1$. Moreover, H_{i+1} is normal in H_i and $H_i/H_{i+1} \cong G_iN/G_{i+1}N$ is cyclic. Therefore G/N is polycyclic. □

Proposition 3.7. *Subgroups of polycyclic groups are polycyclic.*

Proof. Let $G = G_1 \supseteq \cdots \supseteq G_{n+1} = 1$ be a polycyclic series for G and let H be a subgroup of G. Set $H_i = G_i \cap H$. It is easy to check that H_{i+1} is a normal subgroup of H_i. By the second isomorphism theorem,

$$H_i/H_{i+1} \cong H_i/(H_i \cap G_{i+1}) \cong (H_iG_{i+1})/G_{i+1},$$

which is a subgroup of the cyclic group G_i/G_{i+1}. Therefore $H_1 \supseteq \cdots \supseteq H_{n+1} = 1$ is a polycyclic series for H. □

Corollary 3.8. *If G has a polycyclic series of length n, then every subgroup of G can be generated by n or fewer elements.*

Proof. The proof of Proposition 3.7 showed that every subgroup of G has a polycyclic series of length n. Thus Proposition 3.5 applies. □

The following characterization of polycyclic groups is one of the main reasons that extensive computation in polycyclic groups is possible.

Proposition 3.9. *A group is polycyclic if and only if it is solvable and all subgroups are finitely generated.*

Proof. Let G be a group. We have already shown that G polycyclic implies that G is solvable and that subgroups are finitely generated. Now suppose that G is solvable and all subgroups of G are finitely generated. Let G have derived length k. If $k \le 1$, then G is abelian and finitely generated. Therefore G is polycyclic by Proposition 3.2. Suppose $k > 1$. Then $G^{(k-1)}$ is a finitely generated abelian normal subgroup of G and $G/G^{(k-1)}$ has derived length $k-1$. Subgroups of $G/G^{(k-1)}$ are images of subgroups of G and hence are finitely generated. By induction on k, $G/G^{(k-1)}$ is polycyclic. Therefore G is polycyclic by Proposition 3.3. □

Example 3.1. Let D be the group $D_4^{(1)}(\mathbb{Z})$ defined in Example 2.1. Thus D is the subgroup of $\mathrm{GL}(4,\mathbb{Z})$ consisting of all matrices

$$A = \begin{bmatrix} 1 & x_1 & x_4 & x_6 \\ 0 & 1 & x_2 & x_5 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where the x_i are integers. If

$$B = \begin{bmatrix} 1 & y_1 & y_4 & y_6 \\ 0 & 1 & y_2 & y_5 \\ 0 & 0 & 1 & y_3 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

is another element of D, then AB has the form

$$\begin{bmatrix} 1 & x_1 + y_1 & * & * \\ 0 & 1 & x_2 + y_2 & * \\ 0 & 0 & 1 & x_3 + y_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus the map from D to $\mathbb{Z}^3$ taking A to (x_1, x_2, x_3) is a surjective homomorphism with kernel $D_4^{(2)}(\mathbb{Z})$. Similarly, if A is in $D_4^{(2)}(\mathbb{Z})$, then mapping A to (x_4, x_5) defines a homomorphism of $D_4^{(2)}(\mathbb{Z})$ onto $\mathbb{Z}^2$ with kernel $D_4^{(3)}(\mathbb{Z})$. Finally, $D_4^{(3)}(\mathbb{Z})$ is generated by

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

and hence is cyclic. By Proposition 3.3, $D_4^{(2)}(\mathbb{Z})$ is polycyclic. Using the same proposition again, we see that D is polycyclic. If E_i consists of the matrices A given earlier with $x_1 = \cdots = x_{i-1} = 0$, then $D = E_1 \supset E_2 \supset \cdots \supset E_7 = 1$ is a polycyclic series for D.

A group G is *hopfian* if whenever $f: G \to G$ is a surjective homomorphism, then f is an isomorphism. This is equivalent to saying that if N is a normal subgroup of G such that G/N is isomorphic to G, then $N = 1$.

Proposition 3.10. *If G is a group which satisfies the ascending chain condition on subgroups, then G is hopfian.*

Proof. Let $f: G \to G$ be a surjective homomorphism such that the kernel N of f is nontrivial. For $i \geq 1$, the i-fold composition f^i of f with itself is a homomorphism of G onto itself. Set N_i equal to the kernel of f^i. Then N_i is the inverse image of 1 under f^i and N_{i+1} is the inverse image of N under f^i. Since N is nontrivial and f^i is surjective, N_{i+1} properly contains N_i. Therefore $N_1 \subset N_2 \subset \ldots$ is a strictly increasing, infinite sequence of subgroups of G. This is impossible by the ascending chain condition. Therefore G is hopfian. □

Corollary 3.11. *Polycyclic groups are hopfian.*

Proof. By Corollary 3.8, subgroups of polycyclic groups are finitely generated, so the ascending chain condition holds.

□

Proposition 3.12. *Suppose that H and K are subgroups of a polycyclic group G. If $H \subseteq K$ and H is conjugate to K in G, then $H = K$.*

Proof. Suppose that $H \subset K = H^g$, where g is in G. Conjugating repeatedly by g, we see that the sequence $H \subset H^g \subset H^{g^2} \subset \cdots$ is strictly increasing. This cannot happen in a polycyclic group. Thus $H = K$. □

Let G be a polycyclic group. Not all polycyclic series for G have the same length. However, the number of infinite quotients in a polycyclic series is the same for all series. This number is called the *Hirsch number* of G. It is possible to choose the polycyclic series so that all the infinite factors come after the finite factors. See Proposition 2 in Chapter 1 of [Segal 1983].

Exercises

3.1. Show that the order of a finite subgroup of a polycyclic group G divides the product of the orders of the finite quotients in any polycyclic series for G.

3.2. Generalize the discussion in Example 3.1 to $D_n^{(1)}(\mathbb{Z})$ for any $n > 1$.

9.4 Polycyclic presentations

Let $G = G_1 \supseteq \cdots \supseteq G_{n+1} = 1$ be a polycyclic series for a group G. For $1 \leq i \leq n$, let a_i be an element of G_i whose image in G_i/G_{i+1} generates that group. The sequence $a_1, \ldots, a_n$ will be called a *polycyclic generating sequence* for G. Note that the order is important. (For finite solvable groups, the term *AG-system* was introduced in (Jürgensen 1970).) By the proof of Proposition 3.5, $G_i = \mathrm{Grp}\langle a_i, \ldots, a_n\rangle$ and every element g of G_i can be expressed in the form $a_i^{\alpha_i} \ldots a_n^{\alpha_n}$, where the exponents α_j are integers. Let $I = I(a_1, \ldots, a_n)$ denote the set of subscripts i such that G_i/G_{i+1} is finite, and let $m_i = |G_i : G_{i+1}|$, the order of a_i relative to G_{i+1}, if i is in I. We shall normally assume that the generating sequence is not redundant in the sense that no a_i is in G_{i+1}. Thus $m_i > 1$ for each i in I. We shall say that the expression for g is a *collected word* if $0 \le \alpha_j < m_j$ for j in I. Each element of G can be described by a unique collected word in the generators $a_1, \ldots, a_n$. If $a_1^{\alpha_1} \ldots a_n^{\alpha_n}$ is the collected word representing a nontrivial element g of G and α_i is the first nonzero exponent, then $a_i^{\alpha_i}$ and α_i will be called the *leading term* and the *leading exponent* of g, respectively.

Suppose i is in I. Then $a_i^{m_i}$ is in G_{i+1} and can be expressed as a collected word in the generators $a_{i+1}, \ldots, a_n$. The collected word representing a_i^{-1} has the form $a_i^{m_i-1}u$, where u is in G_{i+1}. Thus a_i^{-1} can be eliminated from any word representing an element of G. Now suppose that $1 \le i < j \le n$. Then a_j is in G_{i+1}, which is normal in G_i. Thus the conjugate $a_i^{-1}a_ja_i$ is in G_{i+1}, so a_ja_i can be expressed as the product of a_i and a collected word involving $a_{i+1}, \ldots, a_n$. Similarly, $a_j^{-1}a_i$ can be expressed in this form as well, although this is necessary only if j is not in I. Thus there are unique relations

$$\begin{aligned}
a_ja_i &= a_ia_{i+1}^{\alpha_{iji+1}} \ldots a_n^{\alpha_{ijn}}, && j > i,\\
a_j^{-1}a_i &= a_ia_{i+1}^{\beta_{iji+1}} \ldots a_n^{\beta_{ijn}}, && j > i,\ j \notin I,\\
a_ja_i^{-1} &= a_i^{-1}a_{i+1}^{\gamma_{iji+1}} \ldots a_n^{\gamma_{ijn}}, && j > i,\ i \notin I,\\
a_j^{-1}a_i^{-1} &= a_i^{-1}a_{i+1}^{\delta_{iji+1}} \ldots a_n^{\delta_{ijn}}, && j > i,\ i,j \notin I,\\
a_i^{m_i} &= a_{i+1}^{\mu_{ii+1}} \ldots a_n^{\mu_{in}}, && i \in I,\\
a_i^{-1} &= a_i^{m_i-1}a_{i+1}^{\nu_{ii+1}} \ldots a_n^{\nu_{in}}, && i \in I,
\end{aligned} \qquad (*)$$

where the right sides are collected words. The relations $(*)$ constitute a group presentation for G, the *standard polycyclic presentation* relative to $a_1, \ldots, a_n$. If i is in I, then a_i^{-1} occurs only in one relation, the one giving the collected form of a_i^{-1}. We can get a monoid presentation for G in terms of the a_i and the a_i^{-1} with i not in I by adding the relations $a_ia_i^{-1} = a_i^{-1}a_i = 1$ for i not in I and deleting the relation with left side a_i^{-1} if i is in I. The result will be called the *standard monoid polycyclic presentation.* However, keeping all of the a_i^{-1} is often useful since it facilitates the computation of inverses of elements in G.

Let $X = \{a_1, \ldots, a_n\}$. Interpreted as pairs of words in $X^{\pm *}$, the relations $(*)$ are rewriting rules with respect to the basic wreath-product ordering with $a_n \prec a_n^{-1} \prec \cdots \prec a_1 \prec a_1^{-1}$. Let $\mathcal{R}$ be the set of these rules, together with the monoid rules $a_ia_i^{-1} \to \varepsilon$ and $a_i^{-1}a_i \to \varepsilon$ with i not in I. Using

$\mathcal{R}$, we can rewrite any word in $X^{\pm *}$ into collected form. Since collected forms are unique, the rewriting system is confluent. Since a finite, confluent rewriting system exists, if we start with any monoid presentation for G on the monoid generators in $X^{\pm}$, then, using the ordering $\prec$, the Knuth-Bendix procedure for strings will construct $\mathcal{R}$, which will be called the *standard polycyclic rewriting system* for G relative to the polycyclic generating sequence $a_1, \ldots, a_n$.

Any group presentation of the form $(*)$ defines a polycyclic group G. However, the order of G_i/G_{i+1} may be finite even if i is not in I, and the order of G_i/G_{i+1} may be less than m_i when i is in I. If a presentation $(*)$ is the standard polycyclic presentation for the group it defines, then the presentation is said to be *consistent*. In this context, "consistent" and "confluent" are essentially synonyms.

Example 4.1. Let D be the group $D_4^{(1)}(\mathbb{Z})$ studied in Example 3.1. Define $a_1, \ldots, a_6$ as follows:

$$a_1 = \begin{bmatrix} 1&1&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 1&0&0&0\\ 0&1&1&0\\ 0&0&1&0\\ 0&0&0&1 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&1\\ 0&0&0&1 \end{bmatrix},$$

$$a_4 = \begin{bmatrix} 1&0&1&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{bmatrix}, \quad a_5 \begin{bmatrix} 1&0&0&0\\ 0&1&0&1\\ 0&0&1&0\\ 0&0&0&1 \end{bmatrix}, \quad a_6 \begin{bmatrix} 1&0&0&1\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{bmatrix}.$$

It is easy to check that $a_1, \ldots, a_6$ is a polycyclic generating sequence for D. The set I is empty.

Given an element of D, it is not difficult to determine the collected word $a_1^{\alpha_1} \ldots a_6^{\alpha_6}$ which represents the element. For example, let

$$u = \begin{bmatrix} 1&2&0&-2\\ 0&1&-1&3\\ 0&0&1&1\\ 0&0&0&1 \end{bmatrix}.$$

Then α_1, α_2, and α_3 are the entries of u just above the main diagonal. That is, $\alpha_1 = 2$, $\alpha_2 = -1$, and $\alpha_3 = 1$. Multiplying u on the left by $(a_1^2 a_2^{-1} a_3)^{-1}$

gives

$$\begin{bmatrix} 1 & 0 & 2 & -8 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The entries on the second diagonal above the main diagonal give $\alpha_4 = 2$ and $\alpha_5 = 4$. Multiplying on the left by $(a_4^2 a_5^4)^{-1}$ yields

$$\begin{bmatrix} 1 & 0 & 0 & -8 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

from which we see that $\alpha_6 = -8$. Thus $u = a_1^2 a_2^{-1} a_3 a_4^2 a_5^4 a_6^{-8}$.

Using this technique, we can determine the standard polycyclic presentation for D with respect to $a_1, \ldots, a_6$. It consists of 60 relations. In the following description, α and β range independently over $\{1, -1\}$.

$$\begin{aligned}
a_6^\alpha a_i^\beta &= a_i^\beta a_6^\alpha, \quad 1 \le i \le 5, \\
a_5^\alpha a_1^\beta &= a_1^\beta a_5^\alpha a_6^{-\alpha\beta}, \\
a_5^\alpha a_i^\beta &= a_i^\beta a_5^\alpha, \quad 2 \le i \le 4, \\
a_4^\alpha a_3^\beta &= a_3^\beta a_4^\alpha a_6^{\alpha\beta}, \\
a_4^\alpha a_i^\beta &= a_i^\beta a_4^\alpha, \quad 1 \le i \le 2, \\
a_3^\alpha a_2^\beta &= a_2^\beta a_3^\alpha a_5^{-\alpha\beta}, \\
a_3^\alpha a_1^\beta &= a_1^\beta a_3^\alpha, \\
a_2^\alpha a_1^\beta &= a_1^\beta a_2^\alpha a_4^{-\alpha\beta}.
\end{aligned}$$

In order to recognize that $a_1, \ldots, a_n$ is a polycyclic generating sequence for a group G, we do not have to be given the entire standard polycyclic presentation for G. It would be nice to be able to define a general polycyclic presentation on generators $a_1, \ldots, a_n$ to be any presentation which makes it obvious that the group G defined is polycyclic and that $a_1, \ldots, a_n$ is a polycyclic generating sequence for G. Unfortunately, "obvious" cannot be defined precisely, so we must resort to a somewhat more cumbersome definition, one which describes several cases where the group is obviously polycyclic. The reader should be aware that the terminology related to these presentations is not standard.

Let $a_1, \ldots, a_n$ be a sequence of generators and set $X_i = \{a_i, \ldots, a_n\}^{\pm}$. A *polycyclic presentation* on $a_1, \ldots, a_n$ is a group presentation on these generators such that the following conditions hold:

(i) For $1 \le i < j \le n$ there is a relation $a_j a_i = a_i S_{ij}$, where S_{ij} is in X_{i+1}^*.
(ii) One of the following holds:
  (a) All of the words S_{ij} in (i) have the form $a_j A_{ij}$, where A_{ij} is in X_{j+1}^*.
  (b) For $1 \le i < n$, either there is a relation $a_i^{m_i} = W_i$, where $m_i > 0$ and W_i is in X_{i+1}^*, or for each j with $i < j \le n$ there is a relation $a_j a_i^{-1} = a_i^{-1} U_{ij}$, where U_{ij} is in X_{i+1}^*.

A relation $a_j a_i = a_i S_{ij}$, $i \ne j$, will be called a *commutation relation.* It is *trivial* if $S_{ij} = a_j$. A relation $a_i^{m_i} = W_i$ will be called a *power relation.*

Condition (i) in the definition of a polycyclic presentation does not, by itself, imply that the group is polycyclic, as the following example shows.

Example 4.2. Let G be the group generated by a and b subject to the single defining relation $ba = ab^2$. With respect to the basic wreath-product ordering of $\{a, b\}^{\pm}$ in which $b \prec b^{-1} \prec a \prec a^{-1}$ the group G has the following confluent rewriting system:

$$aa^{-1} \to \varepsilon, \quad a^{-1}a \to \varepsilon, \quad bb^{-1} \to \varepsilon, \quad b^{-1}b \to \varepsilon,$$
$$b^2a^{-1} \to a^{-1}b, \quad ba \to ab^2, \quad b^{-1}a \to ab^{-2}, \quad b^{-1}a^{-1} \to ba^{-1}b^{-1}.$$

The words b^i and aba^{-1} are all irreducible with respect to this system. Thus aba^{-1} is not equal to a power of b, so $H = \text{Grp}\langle b \rangle$ is not equal to aHa^{-1}. Since $(aba^{-1})^2 = ab^2a^{-1} = b$, we have $H \subset aHa^{-1}$. By Proposition 3.12, G is not polycyclic.

Proposition 4.1. *Suppose that G is a group defined by a polycyclic presentation on generators $a_1, \ldots, a_n$. Then G is polycyclic and $a_1, \ldots, a_n$ is a polycyclic generating sequence for G.*

Proof. If condition (iia) holds, then we can prove that G is nilpotent. Clearly this is always the case if $n = 1$. In general, taking $j = n$ in (iia), we see that $a_n a_i = a_i a_n$ for all i. Thus $N = \text{Grp}\langle a_n \rangle$ is in the center of G and hence N is normal in G. A presentation for G/N is obtained by setting a_n equal to 1 in the relations for G. This presentation also satisfies conditions (i) and (iia). Therefore, by induction on n, G/N is nilpotent, and G is nilpotent by Proposition 2.8. By Proposition 3.4, G is polycyclic.

Now suppose that condition (iib) holds. For $1 \le i \le n$, let G_i be the subgroup of G generated by $a_i, \ldots, a_n$. We must show that G_{i+1} is normal in G_i, $1 \le i < n$. It suffices to prove that $a_i^{-1}G_{i+1}a_i = G_{i+1}$. By condition (i), $a_i^{-1}G_{i+1}a_i \subseteq G_{i+1}$. If $a_i^{m_i}$ is in G_{i+1} for some positive integer m_i, then

$$G_{i+1} \supseteq a_i^{-1}G_{i+1}a_i \supseteq a_i^{-2}G_{i+1}a_i^2 \supseteq \cdots \supseteq a_i^{-m_i}G_{i+1}a_i^{m_i} = G_{i+1}.$$

Thus $G_{i+1} = a_i^{-1}G_{i+1}a_i$.

If there is no relation $a_i^{m_i} = W_i$ with W_i in X_{i+1}^*, then condition (iib) says that $a_iG_{i+1}a_i^{-1} \subseteq G_{i+1}$, or $G_{i+1} \subseteq a_i^{-1}G_{i+1}a_i$. Thus $G_{i+1} = a_i^{-1}G_{i+1}a_i$ in this case too. □

By pushing the analysis in the proof of Proposition 4.1 a little further, one can show that the exponents β_{ijk}, δ_{ijk}, and ν_{ijk} in $(*)$ are determined by the α_{ijk}, γ_{ijk}, and μ_{ik}. See Exercise 4.2. If $(*)$ is consistent, then the γ_{ijk} are actually determined by the α_{ijk} and μ_{ik}. See Proposition 8.2. Given a polycyclic presentation for a group, the corresponding standard polycyclic presentation can be obtained using the Knuth-Bendix procedure for strings. It can also be constructed using more specialized techniques. (See Section 9.8.)

Example 4.3. The presentation

$$a_2a_1 = a_1a_2^3, \quad a_2^{-1}a_1 = a_1a_2^{-3}, \quad a_2a_1^{-1} = a_1^{-1}a_2^4, \quad a_2^{-1}a_1^{-1} = a_1^{-1}a_2^{-4}$$

has the form $(*)$ with $I = \emptyset$. As in the Knuth-Bendix procedure, we can rewrite the overlap $a_2a_1a_1^{-1}$ in two ways:

$$a_2\underline{a_1a_1^{-1}} = a_2$$

and

$$\underline{a_2a_1}a_1^{-1} = a_1a_2^3a_1^{-1} = a_1a_2^2a_1^{-1}a_2^4 = a_1a_2a_1^{-1}a_2^8 = a_1a_1^{-1}a_2^{12} = a_2^{12}.$$

Thus $a_2 = a_2^{12}$. Since we are in a group, this implies that $a_2^{11} = 1$. If this relation and $a_2^{-1} = a_2^{10}$ are added, the resulting presentation is consistent.

A polycyclic presentation for which condition (iia) holds will be called a *nilpotent presentation.* A presentation

$$\begin{aligned} a_i^{m_i} &= W_i, \quad 1 \le i \le n, \\ a_ja_i &= a_iS_{ij}, \quad 1 \le i < j \le n \end{aligned} \tag{$**$}$$

where the words W_i and S_{ij} are in $\{a_{i+1}, \ldots, a_n\}^*$, is called a *power-conjugate* presentation and defines a finite solvable group G with order dividing $m_1 \ldots m_m$. The presentation $(**)$ is actually a monoid presentation for G. It is also a rewriting system with respect to the basic wreath-product ordering with $a_n \prec \cdots \prec a_1$. Nilpotent power-conjugate presentations are sometimes called *power-commutator presentations*, although this term is often reserved for the *prime-exponent* case, in which all of the exponents m_i are equal to a fixed prime.

The abbreviation "pc-presentation" is frequently used. However, "pc" could reasonably stand for "polycyclic", "power-conjugate", or "power-commutator". To avoid confusion, the abbreviation "pc" will not be used in this book.

Example 4.4. The following consistent power-conjugate presentation on generators a, b, c, d, e, f, g is a modification of one in (Felsch 1976):

$$\begin{gathered} g^2 = 1, \\ f^4 = 1, \quad gf = fg, \\ e^2 = 1, \quad ge = ef^2g, \quad fe = ef^3, \\ d^6 = f^2, \quad gd = def^3, \quad fd = def^2g, \quad ed = df^2g, \\ c^3 = 1, \quad gc = cd^3f^2g, \quad fc = cf, \quad ec = ced^3, \quad dc = cd^4f^3g, \\ b^2 = 1, \quad gb = bg, \quad fb = bf^3, \quad eb = bef^3, \quad db = bd^5ef, \quad cb = bc^2, \\ a^2 = 1, \quad ga = ag, \quad fa = afg, \quad ea = ad^3f, \quad da = acd^3ef^2, \\ ca = ad^4ef^2g, \quad ba = ab. \end{gathered}$$

The group defined has order $2 \cdot 4 \cdot 2 \cdot 6 \cdot 3 \cdot 2 \cdot 2 = 1152$.

In Section 9.9 and in Chapter 11 we shall be working with a class of nilpotent polycyclic presentations which have additional structure. A γ-*weighted* presentation for a group G is a nilpotent polycyclic presentation $\mathcal{R}$ on generators $a_1, \ldots, a_n$ such that the following hold:

(a) Each a_i has associated with it a positive integer weight w_i such that $w_1 = 1$ and $w_i \le w_{i+1}$ for $1 \le i < n$.
(b) If there is a power relation $a_i^{m_1} = W_i$, then the generators occurring in W_i all have weight at least $w_i + 1$.
(c) For each commutation relation $a_j a_i = a_i a_j A_{ij}$, the generators occurring in A_{ij} all have weight at least $w_i + w_j$.
(d) If $w_k = e > 1$, then there are integers i and j with $w_i = 1$ and $w_j = e - 1$ such that $A_{ij} = a_k$. One such pair is fixed, and the relation $a_j a_i = a_i a_j a_k$ is called the *definition* of a_k.

Suppose that $\mathcal{R}$ is a γ-weighted presentation for G on $a_1, \ldots, a_n$. For $e \ge 1$, set $G(e)$ equal to the subgroup of G generated by the a_i with $w_i \ge e$.

By (c), $G(e)$ is normal in G and $G(e-1)/G(e)$ is central in $G/G(e)$. Thus $\gamma_e(G) \subseteq G(e)$. Let $c = w_n$. Condition (d) implies that $G(c) = [G(c-1), G]$. Now $[G(c-2), G]$ contains $[G(c-1), G]$ and by (d) $G(c-1)$ is generated by $[G(c-2), G]$ modulo $G(c)$. Therefore $G(c-1) = [G(c-2), G]$. Continuing in this way, we find that $G(e+1) = [G(e), G]$, $1 \le e < c$. Thus $\gamma_e(G) = G(e)$. Consistency is more easily checked for γ-weighted presentations than it is for general nilpotent presentations. (See Section 9.8.)

Example 4.5. Let us consider the nilpotent presentation on generators $a_1, \ldots, a_7$ with weights $w_1 = w_2 = w_3 = 1$, $w_4 = w_5 = 2$, and $w_6 = w_7 = 3$ in which $a_j a_i = a_i a_j$ when $w_i + w_j > 3$ and

$$a_2a_1 = a_1a_2a_4^4a_5^2a_6^3, \quad a_3a_1 = a_1a_3a_5, \quad a_3a_2 = a_2a_3a_4, \quad a_4a_1 = a_1a_4a_6,$$
$$a_4a_2 = a_2a_4a_6^3a_7^2, \quad a_4a_3 = a_3a_4a_6^4, \quad a_5a_1 = a_1a_5a_7^6,$$
$$a_5a_2 = a_2a_5a_7, \quad a_5a_3 = a_3a_5a_6^{-2}a_7^4.$$

This presentation is γ-weighted. The definitions of a_4 to a_7 are the relations

$$a_3a_2 = a_2a_3a_4, \quad a_3a_1 = a_1a_3a_5, \quad a_4a_1 = a_1a_4a_6, \quad a_5a_2 = a_2a_5a_7,$$

respectively.

Rewriting with respect to a standard polycyclic rewriting system or a power-conjugate system is now called *collection*. The term "collection" was originally introduced in (P. Hall 1934) and referred there to computation in free nilpotent groups as discussed in Section 9.10. A great deal of effort has gone into devising efficient collection strategies. Suppose the generators are $a_1, \ldots, a_n$. Hall used a strategy in which all occurrences of a_1 are moved left to the beginning of the word. Next, all occurrences of a_2 are moved left until they are adjacent to the a_1's. Then the a_3's are moved left, and so on. This collection strategy is called *collection to the left*. It has properties which make it useful in the proofs of various formulas for the collected form of special words, but it is usually not efficient for computation. For some time, a consensus favored *collection from the right*, in which the left side of a rule occurring nearest the end of the word is selected for replacement. However, evidence in (Leedham-Green & Soicher 1990) and (Vaughan-Lee 1990) suggests that in a substantial number of cases *collection from the left* is superior. In this strategy, the left side nearest the beginning of the word is chosen.

Example 4.6. Let us compare these three strategies as applied to collecting the word $a_1a_2a_3a_4a_1a_2a_3a_4$ using the rewriting system of Example 4.1. Collection to the left uses 22 applications of the rules:

$$
\begin{aligned}
a_1a_2a_3\underline{a_4a_1}a_2a_3a_4 &= a_1a_2\underline{a_3a_1}a_4a_2a_3a_4 \\
&= a_1\underline{a_2a_1}a_3a_4a_2a_3a_4 \\
&= a_1a_1a_2a_4^{-1}a_3\underline{a_4a_2}a_3a_4 \\
&= a_1a_1a_2a_4^{-1}\underline{a_3a_2}a_4a_3a_4 \\
&= a_1a_1a_2\underline{a_4^{-1}a_2}a_3a_5^{-1}a_4a_3a_4 \\
&= a_1a_1a_2a_2\underline{a_4^{-1}a_3}a_5^{-1}a_4a_3a_4 \\
&= a_1a_1a_2a_2a_3a_4^{-1}a_6^{-1}a_5^{-1}\underline{a_4a_3}a_4 \\
&= a_1a_1a_2a_2a_3a_4^{-1}a_6^{-1}\underline{a_5^{-1}a_3}a_4a_6a_4 \\
&= a_1a_1a_2a_2a_3a_4^{-1}\underline{a_6^{-1}a_3}a_5^{-1}a_4a_6a_4 \\
&= a_1a_1a_2a_2a_3\underline{a_4^{-1}a_3}a_6^{-1}a_5^{-1}a_4a_6a_4 \\
&= a_1a_1a_2a_2a_3a_3a_4^{-1}a_6^{-1}a_6^{-1}\underline{a_5^{-1}a_4}a_6a_4 \\
&= a_1a_1a_2a_2a_3a_3a_4^{-1}a_6^{-1}\underline{a_6^{-1}a_4}a_5^{-1}a_6a_4 \\
&= a_1a_1a_2a_2a_3a_3a_4^{-1}\underline{a_6^{-1}a_4}a_6^{-1}a_5^{-1}a_6a_4 \\
&= a_1a_1a_2a_2a_3a_3\underline{a_4^{-1}a_4}a_6^{-1}a_6^{-1}a_5^{-1}a_6a_4 \\
&= a_1a_1a_2a_2a_3a_3a_6^{-1}a_6^{-1}a_5^{-1}\underline{a_6a_4} \\
&= a_1a_1a_2a_2a_3a_3a_6^{-1}a_6^{-1}\underline{a_5^{-1}a_4}a_6 \\
&= a_1a_1a_2a_2a_3a_3a_6^{-1}\underline{a_6^{-1}a_4}a_5^{-1}a_6 \\
&= a_1a_1a_2a_2a_3a_3\underline{a_6^{-1}a_4}a_6^{-1}a_5^{-1}a_6 \\
&= a_1a_1a_2a_2a_3a_3a_4a_6^{-1}\underline{a_6^{-1}a_5^{-1}}a_6 \\
&= a_1a_1a_2a_2a_3a_3a_4\underline{a_6^{-1}a_5^{-1}}a_6^{-1}a_6 \\
&= a_1a_1a_2a_2a_3a_3a_4a_5^{-1}a_6^{-1}\underline{a_6^{-1}a_6} \\
&= a_1a_1a_2a_2a_3a_3a_4a_5^{-1}a_6^{-1}.
\end{aligned}
$$

Collection from the right does not do any better:

$$
\begin{aligned}
a_1a_2a_3\underline{a_4a_1}a_2a_3a_4 &= a_1a_2a_3a_1\underline{a_4a_2}a_3a_4 \\
&= a_1a_2a_3a_1a_2\underline{a_4a_3}a_4 \\
&= a_1a_2a_3a_1a_2a_3a_4\underline{a_6a_4} \\
&= a_1a_2\underline{a_3a_1}a_2a_3a_4a_4a_6 \\
&= a_1a_2a_1\underline{a_3a_2}a_3a_4a_4a_6
\end{aligned}
$$

$$
\begin{aligned}
&= a_1a_2a_1a_2a_3\underline{a_5^{-1}a_3}a_4a_4a_6 \\
&= a_1a_2a_1a_2a_3a_3\underline{a_5^{-1}a_4}a_4a_6 \\
&= a_1a_2a_1a_2a_3a_3a_4\underline{a_5^{-1}a_4}a_6 \\
&= a_1\underline{a_2a_1}a_2a_3a_3a_4a_4a_5^{-1}a_6 \\
&= a_1a_1a_2\underline{a_4^{-1}a_2}a_3a_3a_4a_4a_5^{-1}a_6 \\
&= a_1a_1a_2a_2\underline{a_4^{-1}a_3}a_3a_4a_4a_5^{-1}a_6 \\
&= a_1a_1a_2a_2a_3a_4^{-1}\underline{a_6^{-1}a_3}a_4a_4a_5^{-1}a_6 \\
&= a_1a_1a_2a_2a_3a_4^{-1}a_3\underline{a_6^{-1}a_4}a_4a_5^{-1}a_6 \\
&= a_1a_1a_2a_2a_3a_4^{-1}a_3a_4\underline{a_6^{-1}a_4}a_5^{-1}a_6 \\
&= a_1a_1a_2a_2a_3a_4^{-1}a_3a_4a_4\underline{a_6^{-1}a_5^{-1}}a_6 \\
&= a_1a_1a_2a_2a_3a_4^{-1}a_3a_4a_4a_5^{-1}\underline{a_6^{-1}a_6} \\
&= a_1a_1a_2a_2a_3\underline{a_4^{-1}a_3}a_4a_4a_5^{-1} \\
&= a_1a_1a_2a_2a_3a_3a_4^{-1}\underline{a_6^{-1}a_4}a_4a_5^{-1} \\
&= a_1a_1a_2a_2a_3a_3a_4^{-1}a_4\underline{a_6^{-1}a_4}a_5^{-1} \\
&= a_1a_1a_2a_2a_3a_3a_4^{-1}a_4a_4\underline{a_6^{-1}a_5^{-1}} \\
&= a_1a_1a_2a_2a_3a_3\underline{a_4^{-1}a_4}a_4a_5^{-1}a_6^{-1} \\
&= a_1a_1a_2a_2a_3a_3a_4a_5^{-1}a_6^{-1}.
\end{aligned}
$$

However, collection from the left requires only 12 applications of the rules:

$$
\begin{aligned}
a_1a_2a_3\underline{a_4a_1}a_2a_3a_4 &= a_1a_2\underline{a_3a_1}a_4a_2a_3a_4 \\
&= a_1\underline{a_2a_1}a_3a_4a_2a_3a_4 \\
&= a_1a_1a_2\underline{a_4^{-1}a_3}a_4a_2a_3a_4 \\
&= a_1a_1a_2a_3a_4^{-1}\underline{a_6^{-1}a_4}a_2a_3a_4 \\
&= a_1a_1a_2a_3\underline{a_4^{-1}a_4}a_6^{-1}a_2a_3a_4 \\
&= a_1a_1a_2a_3\underline{a_6^{-1}a_2}a_3a_4 \\
&= a_1a_1a_2\underline{a_3a_2}a_6^{-1}a_3a_4 \\
&= a_1a_1a_2a_2a_3a_5^{-1}\underline{a_6^{-1}a_3}a_4 \\
&= a_1a_1a_2a_2a_3\underline{a_5^{-1}a_3}a_6^{-1}a_4 \\
&= a_1a_1a_2a_2a_3a_3a_5^{-1}\underline{a_6^{-1}a_4}
\end{aligned}
$$

$$= a_1a_1a_2a_2a_3a_3\underline{a_5^{-1}a_4}a_6^{-1}$$
$$= a_1a_1a_2a_2a_3a_3a_4a_5^{-1}a_6^{-1}.$$

In its basic form, neither collection from the right nor collection from the left is adequate. There are several ways of speeding up the process further. With either strategy, there is a collected part and an uncollected part of the current word. In collection from the right, the collected part is a suffix, while in collection from the left it is a prefix. The collected part $a_1^{\alpha_1}\ldots a_n^{\alpha_n}$ can be represented by its exponent vector $(\alpha_1,\ldots,\alpha_n)$. Let r be the smallest index such that $G_r = \mathrm{Grp}\langle a_r,\ldots,a_n\rangle$ is abelian. In Example 4.1, $r = 4$. If $i \geq r-1$, then in collection from the left we have

$$(\alpha_1,\ldots,\alpha_n)a_i^\beta = (\alpha_1,\ldots,\alpha_i+\beta,\alpha_{i+1},\ldots,\alpha_n).$$

Note that, if i is in I and $\alpha_i+\beta \geq m_i$ or $\alpha_i+\beta < 0$, then the vector on the right does not represent a collected word. In collection from the right, one has

$$a_i^\beta(0,\ldots,0,\alpha_i,\ldots,\alpha_n) = (0,\ldots,0,\alpha_i+\beta,\alpha_{i+1},\ldots,\alpha_n)$$

for any i, and

$$a_i^\beta(0,\ldots,0,a_r,\ldots,a_n) = (0,\ldots,0,a_r,\ldots,a_{i-1},a_i+\beta,a_{i+1},\ldots,a_n)$$

provided $i \geq r$. When collecting in nilpotent groups, it is important to be able to develop formulas for the collected form of various families of words. The simplest formula is $a_j^\alpha a_i^\beta = a_i^\beta a_j^\alpha$ when $i < j$ and a_i and a_j commute. If $\mathrm{Grp}\langle a_i,a_j\rangle$ is nilpotent of class 2, then $a_j^\alpha a_i^\beta = a_i^\beta a_j^\alpha[a_j,a_i]^{\alpha\beta}$. The formulas defining the rules in Example 4.1 are valid for all integers α and β, not just in the case $|\alpha| = |\beta| = 1$. The use of more complicated formulas in an approach called *combinatorial collection* is described in (Havas & Nicholson 1976).

When formulas cannot easily be derived and memory is not a problem, then one can store the collected form for additional products $a_j^\alpha a_i^\beta$. In a power-conjugate system, this could mean storing up to

$$\sum_{i<j}(m_j-1)(m_i-1)$$

conjugation rules. With the presentation of Example 4.4, adding the 56 extra rules greatly reduces the time needed to collect words. In (Felsch 1976), a novel data structure is described which encodes a power-conjugate presentation as a family of subroutines, which are executed to carry out collection.

Extensive research on collection is in progress as of this writing. Most experimental evidence comes from collection within finite solvable groups. Complicating the situation is the observation that the words which arise in various important algorithms connected with polycyclic groups do not appear to be random. To get the best performance on these algorithms, the collection procedure must be tuned to the words which occur most frequently. In view of the inconclusive results available at the present time, no recommendation will be made here concerning the choice of collection procedures.

Although we may not know the best collection strategy, we certainly can solve the word problem in a group G given by polycyclic presentations. Probably the next problem about elements of G to consider is the conjugacy problem. Conjugacy of elements can be decided in principle, but practical algorithms for infinite polycyclic groups have not yet been developed. If two elements g and h of G are not conjugate, then there is a finite quotient group of G in which the images of g and h are not conjugate. A proof of this result may be found in [Segal 1983]. This leads to the following "algorithm" for deciding whether two elements g and h are conjugate. We start two computers running. The first computer systematically forms conjugates of g in G. The second computer systematically examines the finite quotients of G. The conjugacy problem in a finite group is clearly solvable in principle. Thus either the first computer will find an element u of G such that $u^{-1}gu = h$ or the second computer will find a finite quotient group of G in which the images of g and h are not conjugate. We simply wait to see which computer stops. Useful conjugacy algorithms for finite solvable groups have been developed. See (Mecky & Neubüser 1989). Conjugacy in nilpotent groups is discussed in Section 9.7.

The next two sections discuss computation in subgroups and quotient groups of polycyclic groups.

Exercises

4.1. Suppose that $a_1, \ldots, a_n$ is a polycyclic generating sequence for a group G. Let $a_1^{\alpha_1} \ldots a_n^{\alpha_n}$ be the collected word defining an element g of G, let $a_i^{\beta_i}$ be the leading term of an element h, and let $a_1^{\gamma_1} \ldots a_n^{\gamma_n}$ be the collected word defining gh. Show that $\gamma_j = \alpha_j$, $1 \le j < i$, and that $\gamma_i = \alpha_i + \beta_i$ if i is not in I, while $\gamma_i = (\alpha_i + \beta_i) \bmod m_i$ if i is in I. Here $I = I(a_1, \ldots, a_n)$ and m_i is the relative order of a_i modulo $\mathrm{Grp}\langle a_{i+1}, \ldots, a_n\rangle$. Describe the leading terms of gh and h^{-1}.

4.2. Show that the exponents β_{ijk}, δ_{ijk}, and ν_{ik} in a standard polycyclic presentation $(*)$ are determined by the α_{ijk}, γ_{ijk}, and μ_{ik}.

4.3. Let $(X, \mathcal{R})$ be a nilpotent presentation satisfying conditions (a), (b), and (c) of the definition of a γ-weighted presentation. For $e \ge 1$, let $Q(e)$ be the abelian group generated by the a_i of weight e subject to the relations $a_i^{m_i} = 1$ for i in I and $w_i = e$. Given U in $X^{\pm *}$, let $\overline{U}$ denote the word obtained from U by deleting all generators of weight different from e. We shall say that $(X, \mathcal{R})$ is *weakly γ-weighted* if for $e \ge 2$ the group $Q(e)$ is generated by the images of the words $\overline{A_{ij}}$, where $w_i = 1$ and $w_j = e - 1$. Show that a γ-weighted presentation is weakly γ-weighted and that in a group G defined by a weakly γ-weighted presentation the group $\gamma_e(G)$ is generated by the a_i with $w_i \geq e$.

9.5 Subgroups

In Section 8.1 we discussed how to compute with subgroups of $\mathbb{Z}^n$. Subgroups were represented by integer matrices with n columns, and integer row operations were used to manipulate these matrices. This section uses ideas of M. F. Newman described in (Laue et al. 1984) to generalize the techniques of Section 8.1 to study the subgroups of a group G given by a polycyclic presentation on generators $a_1, \dots, a_n$. The group $\mathrm{Grp}\langle a_i, \dots, a_n\rangle$ will be denoted G_i. We shall assume that the presentation is consistent and that we know the set I of indices i such that G_i/G_{i+1} is finite and the relative order m_i of a_i modulo G_{i+1} for i in I.

A subgroup H of G will be described by a sequence $U = (g_1, \dots, g_s)$ of generating elements. Let the collected form of g_i be $a_1^{\alpha_{i1}} \dots a_n^{\alpha_{in}}$. The s-by-n matrix A of integers α_{ij} will be used to represent U and will be called the *associated exponent matrix*. Corresponding to the elementary row operations of Section 8.1, we have the following *elementary operations* on U:

(1) Interchange g_i and g_j if $i \neq j$.
(2) Replace g_i by g_i^{-1}.
(3) Replace g_i by $g_i g_j^\beta$, where β is an integer and $i \neq j$.
(4) Add as a new component g_{s+1} any element of $\mathrm{Grp}\langle g_1, \dots, g_s\rangle$.
(5) Delete g_s, if $g_s = 1$.

Notice that the length of U may increase or decrease. This was not necessary in Section 8.1, since any finite sequence of generators of an abelian group is a polycyclic generating sequence. In general, a polycyclic group generated by a small set may require long polycyclic generating sequences.

Two sequences $U = (g_1, \dots, g_s)$ and $V = (h_1, \dots, h_t)$ are *equivalent under elementary operations* if one can be transformed into the other by a sequence of these operations. To see that this is an equivalence relation, we must show that the effect of an elementary operation can be undone by a sequence of one or more operations. Applying operations of types (1) and (2) twice to U leaves U unchanged. If an operation of type (3) is applied, then replacing g_i by $g_i g_j^{-\beta}$ restores U to its original form. After an operation of type (4), g_{s+1} is a product of powers of the g_i with $1 \leq i \leq s$. A sequence of operations of type (3) can make g_{s+1} the identity element, and then an operation of type (5) deletes g_{s+1}. Finally, if an operation of type (5) is performed, then undoing it is a special case of an operation of type (4).

Let us say that a sequence $U = (g_1, \dots, g_s)$ of elements of G is in *standard* *form* if the associated exponent matrix A satisfies the following conditions:

(i) All rows of A are nonzero (i.e., no g_i is the identity).
(ii) A is row reduced over $\mathbb{Z}$.
(iii) If A_{ij} is a corner entry and j is in I, then A_{ij} divides m_j.

Suppose that U is in standard form. An *admissible sequence of exponents* for U is a sequence $(\beta_1, \ldots, \beta_s)$ of integers such that, if A_{ij} is a corner entry and j is in I, then $0 \le \beta_i < m_j/A_{ij}$. Let $E(U)$ be the set of admissible sequences of exponents for U and let $S(U)$ be the set of products $g_1^{\beta_1} \ldots g_s^{\beta_s}$, where $(\beta_1, \ldots, \beta_s)$ ranges over $E(U)$.

Example 5.1. Suppose G is $\mathbb{Z}^n$ and $a_1, \ldots, a_n$ is the standard basis. Then a sequence $U = (g_1, \ldots, g_s)$ of elements of G is in standard form if and only if the associated matrix A is row reduced and has rank s. In this case, all s-tuples of integers are admissible for U. The set $S(U)$ is simply the subgroup $S(A)$ of Section 8.1. The components of U form a basis of $S(U)$.

In general, $S(U)$ is not a subgroup, but there is a one-to-one correspondence between elements of $S(U)$ and elements of $E(U)$.

Proposition 5.1. *Suppose that $U = (g_1, \ldots, g_s)$ is a sequence of elements of G in standard form and $(\beta_1, \ldots, \beta_s)$ and $(\gamma_1, \ldots, \gamma_s)$ are in $E(U)$. If $g_1^{\beta_1} \ldots g_s^{\beta_s} = g_1^{\gamma_1} \ldots g_s^{\gamma_s}$, then $\beta_i = \gamma_i$, $1 \le i \le s$.*

Proof. Let A be the matrix of exponents associated with U, and let $g = g_1^{\beta_1} \ldots g_s^{\beta_s} = g_1^{\gamma_1} \ldots g_s^{\gamma_s}$. Suppose A_{1j} is the corner entry in the first row of A. If $a_1^{\delta_1} \ldots a_n^{\delta_n}$ is the collected word defining g, then $\delta_k = 0$, $1 \le k < j$. We have two cases depending on whether j is in I. Assume first that j is not in I. Then G_j/G_{j+1} is isomorphic to $\mathbb{Z}$ and $\delta_j = A_{1j}\beta_1 = A_{1j}\gamma_1$. Since $A_{1j} \ne 0$, this means that $\beta_1 = \gamma_1$. Now assume that j is in I. Then G_j/G_{j+1} is isomorphic to $\mathbb{Z}_{m_j}$ and $A_{1j}\beta_1 \equiv \delta_j \equiv A_{1j}\gamma_1 \pmod{m_j}$. But both β_1 and γ_1 are nonnegative and less than m_j/A_{1j}. Therefore $A_{1j}\beta_1$ and $A_{1j}\gamma_1$ are nonnegative and less than m_j. Thus $A_{1j}\beta_1 = A_{1j}\gamma_1$. Hence $\beta_1 = \gamma_1$ in this case too. Thus we may multiply g on the left by $g_1^{-\beta_1}$ and conclude that $g_2^{\beta_2} \ldots g_s^{\beta_s} = g_2^{\gamma_2} \ldots g_s^{\gamma_s}$. By induction applied to the $(s-1)$-tuple $(g_2, \ldots, g_s)$, we have $\beta_i = \gamma_i$, $2 \le i \le s$. $\square$

The proof of Proposition 5.1 gives us an algorithm for deciding membership in $S(U)$ for a sequence U of elements of G in standard form. It is a straightforward generalization of the algorithm in Section 8.1 for deciding membership in a subgroup of $\mathbb{Z}^n$ given as $S(B)$, where B is a row reduced integer matrix.

Function POLY_MEMBER(U, g): boolean;
Input: U : a sequence $(g_1, \ldots, g_s)$ of elements of G in standard form;
g : an element of G;
(* The value true is returned if g is in $S(U)$, and false is returned otherwise. *)
Begin
Let A be the exponent matrix associated with U; $h := g$;
(* At all times $a_1^{\gamma_1} \ldots a_n^{\gamma_n}$ will be the collected word representing h. *)
$i := 1$; *done* := false;
While $i \le s$ and not *done* do begin
Let A_{ij} be the corner entry of A in the i-th row;
If some $\gamma_k \neq 0$ for $1 \le k < j$ then *done* := true
Else if A_{ij} does not divide γ_j then *done* := true
Else begin
$q := \gamma_j / A_{ij}$; $h := g_i^{-q} h$;
End;
$i := i + 1$
End;
POLY_MEMBER := $(h = 1)$
End.

Example 5.2. Let $D = D_4^{(1)}(\mathbb{Z})$ as given by the presentation on $a_1, \ldots, a_6$ derived in Example 4.1. If

$$g_1 = a_1^2 a_2^{-1} a_4,$$
$$g_2 = a_3^3 a_4 a_6,$$
$$g_3 = a_4^2 a_5 a_6,$$

then $U = (g_1, g_2, g_3)$ is in standard form. The associated exponent matrix is

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

Let us decide whether $u = a_1^{-6} a_2^3 a_3^6 a_4^{21} a_5^5 a_6^{64}$ is in $S(U)$. Since $I = \emptyset$, we simply want to know whether u can be expressed as $g_1^{\beta_1} g_2^{\beta_2} g_3^{\beta_3}$. The leading terms of u and g_1 are a_1^{-6} and a_1^2, respectively. Thus if u is in $S(U)$, then $\beta_1 = (-6)/2 = -3$. The product $g_1^3 u$ is $v = a_3^6 a_4^{12} a_5^5 a_6^{10}$. Therefore β_2 must be $6/3 = 2$. Multiplying g_2^{-2} and v yields $w = a_4^{10} a_5^5 a_6^5$. Now $\beta_3 = 10/2 = 5$. The product $g_3^{-5} w$ is 1, so $u = g_1^{-3} g_2^2 g_3^5$ is in $S(U)$.

Given a sequence $U = (g_1, \ldots, g_s)$ of elements of G in standard form, we can decide whether $S(U)$ is a subgroup of G. Let us say that U is *full* if the following conditions hold:

(i) For $1 \le i < j \le s$ the set $S(U)$ contains $g_i^{-1} g_j g_i$.
(ii) If A_{ij} is a corner entry of the matrix A of exponents associated with U and j is in I, then $S(U)$ contains g_i^q, where $q = m_j / A_{ij}$.

Example 5.3. The triple $U = (g_1, g_2, g_3)$ in Example 5.2 is not full. Let $u = g_1^{-1} g_2 g_1 = a_3^3 a_4 a_5^3 a_6^{-2}$. If u is in $S(U)$ and $u = g_1^{\beta_1} g_2^{\beta_2} g_3^{\beta_3}$, then $\beta_1 = 0$ and $\beta_2 = 1$. But $g_2^{-1} u = a_5^3 a_6^{-3}$. Since no g_i has a power of a_5 as its leading term, u is not in $S(U)$.

Proposition 5.2. *If $U = (g_1, \ldots, g_s)$ is a sequence of elements of G in standard form, then $S(U)$ is a subgroup of G if and only if U is full. If U is full, then $g_1, \ldots, g_s$ is a polycyclic generating sequence for $S(U)$.*

Proof. The elements g_i are in $S(U)$. Thus, if $S(U)$ is a subgroup of G, then U is full. Assume now that U is full. If $s = 0$, then $S(U) = \{1\}$ is a subgroup. We proceed by induction on s and suppose that $s > 0$. Let A_{1k} be the corner entry in the first row of the matrix associated with U. Then $g_2, \ldots, g_s$ are contained in G_{k+1} and $S(U) \cap G_{k+1}$ is $S(V)$, where $V = (g_2, \ldots, g_s)$. If $2 \le i < j \le s$, then $g_i^{-1} g_j g_i$ is in $S(U)$ and is in G_{k+1}. Therefore $g_i^{-1} g_j g_i$ is in $S(V)$. By a similar argument, V satisfies condition (ii) of the definition of full. Therefore V is full and, by induction, $H = S(V)$ is a subgroup of G. If $2 \le j \le s$, then $g_1^{-1} g_j g_1$ is in G_{k+1} by Exercise 4.1. Since U is full, $g_i^{-1} g_j g_1$ is also in $S(U)$, and hence $g_i^{-1} g_j g_1$ is in H. Thus $g_1^{-1} H g_1 \subseteq H$. By Proposition 3.12, $g_1^{-1} H g_1 = H$. This implies that $g_1 H g_1^{-1} = H$. Thus H is normal in $K = \mathrm{Grp}\langle g_1, \ldots, g_s \rangle$ and every element of K can be written in the form $g_1^\alpha h$, where h is in H. If k is not in I, then every element $g_1^\alpha h$ is in $S(U)$, so $K = S(U)$. Suppose that k is in I and $q = m_k / A_{1k}$. Then $u = g_1^q$ is in H, so α can always be chosen so that $0 \le \alpha < q$. Thus $K = S(U)$ in this case too. $\square$

The following result generalizes Proposition 1.1 in Chapter 8.

Proposition 5.3. *Let H be a subgroup of G. There is a unique sequence $U = (g_1, \ldots, g_s)$ in standard form such that $H = S(U)$.*

Proof. Let k be the largest index such that $H \subseteq G_k$. If $k = n+1$, then H is trivial and the empty sequence $U = ()$ is the only sequence in standard form such that $H = S(U)$. Suppose that $k \le n$. Let $^-$ denote the canonical homomorphism from G_k to G_k / G_{k+1}. The image $\overline{H}$ is nontrivial. Let α be the least positive integer such that $(\overline{a_k})^\alpha$ is in $\overline{H}$. Then $\overline{H}$ is generated by $(\overline{a_k})^\alpha$, and, if k is in I, then α divides m_k. Let g_1 be an element of H such that $\overline{g_1} = (\overline{a_k})^\alpha$. Then the leading term of g_1 is a_k^α. By induction on $n-k$, there is a sequence $W = (g_2, \dots, g_s)$ in standard form such that $H \cap G_{k+1} = S(W)$. The sequence $U = (g_1, g_2, \dots, g_s)$ satisfies conditions (i) and (iii) of the definition of standard form, but it may not satisfy condition (ii) because the entries in row 1 of the matrix A associated with U which lie above corner entries of A may not be reduced modulo those corner entries. Thus it may be necessary to modify g_1. To do so, we execute the following instructions. At all times, $g_1 = a_1^{\alpha_1} \dots a_n^{\alpha_n}$.

For $i := 2$ to s do begin
  Let a_j^β be the leading term of g_i; $q := \alpha_j \operatorname{div} \beta$; $g_1 := g_1 g_i^{-q}$
End.

By Exercise 4.1, the entries in row 1 of A which lie above corner entries are now reduced modulo those entries. Hence U is in standard form. Every element of H has the form $g_1^\gamma u$, where u is in $H \cap G_{k+1}$ and $0 \le \gamma < m_k/\alpha$ if k is in I. Since $H \cap G_{k+1} = S(W)$, it follows that $H = S(U)$. Thus we have proved the existence part of the proposition.

To prove uniqueness, suppose that $H = S(U) = S(V)$, where both $U = (g_1, \dots, g_s)$ and $V = (h_1, \dots, h_t)$ are in standard form. The leading term of g_1 is a_k^α, where k is as defined earlier and $\alpha > 0$. If k is in I, then α divides m_k. The group $\overline{H}$ is generated by $\overline{g_1} = (\overline{a_k})^\alpha$, and $\overline{H}$ has order m_k/α if k is in I. These conditions uniquely determine α. By symmetry, the leading term of h_1 is also a_k^α. Now $H \cap G_{k+1} = S(g_2, \dots, g_s) = S(h_2, \dots, h_t)$. By induction on $n-k$, we have $s = t$ and $g_i = h_i$, $2 \le i \le s$. Let $u = g_1^{-1} h_1$. Then u is in $H \cap G_{k+1}$, so $u = g_2^{\beta_2} \dots g_s^{\beta_s}$, where $(\beta_2, \dots, \beta_s)$ is in $E(g_2, \dots, g_s)$. If $u = 1$, then $g_1 = h_1$ and we are done. Suppose that $u \ne 1$ and let i be minimal such that $\beta_i \ne 0$. Let a_j^δ be the leading term of g_i and let the collected forms of g_1 and h_1 be $a_1^{\mu_1} \dots a_n^{\mu_n}$ and $a_1^{\nu_1} \dots a_1^{\nu_1}$, respectively. Assume that j is not in I. Since $h_1 = g_1 u$, we have $\nu_j = \mu_j + \beta_i \delta$. But this is not possible, since both μ_j and ν_j are reduced modulo δ. A similar argument takes care of the case in which j is in I. □

Suppose $V = (h_1, \dots, h_r)$ is a sequence of elements of G and H is the subgroup generated by the h_i. If we knew the full sequence U in standard form such that $H = S(U)$, then we could decide membership in H using POLY_MEMBER. It is in fact possible to transform V into U using elementary operations. The procedure is a relatively straightforward generalization of the row reduction procedure of Section 8.1.

Initially set U equal to V. The first observation is that we may apply elementary operations in such a way that the matrix A associated with U is in row echelon form. Suppose that g_i and g_j have leading terms a_k^β and a_k^γ with $|\beta| \ge |\gamma|$. Let $q = \beta \operatorname{div} \gamma$. Replacing g_i by $g_i g_j^{-q}$ sets the exponent of a_k in g_i equal to $\beta \bmod \gamma$. Repeating this step until it is no longer possible produces a sequence in which no two leading terms involve the same generator. By permuting the elements in the sequence U and deleting elements which are the identity, we can assume that A is in row echelon form and has no zero rows. By replacing g_i with g_i^{-1} if necessary, we can make all corner entries positive.

Now suppose that there is a corner entry $\alpha = A_{ik}$ such that k is in I and A_{ik} does not divide m_k. Let $\beta = \gcd(\alpha, m_k) = p\alpha + qm_k$. The leading term of g_i^p is a_k^β. Add g_i^p as a new member of the sequence U and repeat the procedure in the previous paragraph to put A back into row echelon form. Since this iteration either introduces a new column containing a corner entry or replaces a corner entry with a proper divisor, the process stops eventually with $U = (g_1, \ldots, g_s)$ such that the associated matrix A is in row echelon form and all rows are nonzero, corner entries are positive, and any corner entry A_{ik} with k in I divides m_k. To get U into standard form, we have only to execute the following statements:

```
For i := 2 to s do begin
   Let a_k^α be the leading term of g_i;
   For j := 1 to i − 1 do begin
      Let β be the exponent on a_k in the collected word representing g_j;
      q := β div α;  g_j := g_j g_i^(−q)
   End
End.
```

Now that U is standard, we begin checking whether U is full. The test for fullness requires that various elements be in $S(U)$. Suppose that u is one of those elements and u is not in $S(U)$. Thus POLY_MEMBER($U; u$) returns false. Let v be the last value assigned to h within POLY_MEMBER. Add v as a new member of the sequence U and repeat the entire process. When U has again been put into standard form, either there will be a new column in A containing a corner entry or some corner entry will have been reduced. Thus this iteration must also stop. When it does, U will be full and $S(U)$ will be $H = \mathrm{Grp}\langle h_1, \ldots, h_r\rangle$.

Example 5.4. Let us continue Examples 5.2 and 5.3 and determine the subgroup of D generated by g_1, g_2, and g_3. The sequence U is already in standard form. However, $u = g_1^{-1}g_2g_1$ is not in $S(U)$. This becomes clear when we compute $v = g_2^{-1}u = a_5^3a_6^{-3}$. Thus we define g_4 to be v. Now A is

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

and U is still in standard form. Now let $u = g_1^{-1}g_3g_1 = a_4^2a_5a_6^{-1}$. To check membership of u in $S(U)$, we compute $v = g_3^{-1}u = a_6^{-2}$. Clearly v is not in $S(U)$. In order to have a positive corner entry, we define g_5 to be $v^{-1} = a_6^2$. To put U in standard form, we have only to replace g_4 by $g_4g_5^2 = a_5^3a_6$. This gives

$$A = \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

At this point $U = (g_1, g_2, g_3, g_4, g_5)$ is full.

Example 5.5. Let us now determine the subgroup generated by $h_1 = ad^3eg$ and $h_2 = bf$ in the group of order 1152 in Example 4.4. Initially set $g_1 = h_1$, $g_2 = h_2$, and $U = (g_1, g_2)$. The associated matrix A is

$$\begin{bmatrix} 1 & 0 & 0 & 3 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

and U is in standard form. If $u = g_1^{-1}g_2g_1$, then $u = bf^3$ and $g_2^{-1}u = f^2$. We define g_3 to be f^2. Now A is

$$\begin{bmatrix} 1 & 0 & 0 & 3 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 \end{bmatrix},$$

and U is still in standard form. The square of g_1 is g, and we define g_4 to be g. In order to get U into standard form, we replace g_1 by $g_1g_4^{-1} = ad^3e$. This gives

$$A = \begin{bmatrix} 1 & 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now U is full. The order of H can now be seen to be 16.

Let us give the name POLY_SUBGROUP to the procedure for determining a full sequence U of generators for a subgroup of a polycyclic group described by generating elements. In order to formalize POLY_SUBGROUP, we must choose an order in which to test the elements $g_i^{-1}g_jg_i$ and g_i^q in the definition of a full sequence. It is also useful to try to arrange the computation so that, when a new generator is added to U, we can avoid repeating all of the tests made previously. There is inadequate experience at this point on which to base a firm recommendation. The reader is encouraged to experiment with various ways of spelling out the details of POLY_SUBGROUP.

Given a finite subset Y of G, we can compute the normal closure N of Y in G using POLY_SUBGROUP. There are at least two ways to organize the computation. In the most straightforward approach, we first find the full sequence U in standard form such that $\mathrm{Grp}\langle Y\rangle = S(U)$. For each element y of Y and each generator x of G we compute $u = x^{-1}yx$ and check whether u is in $S(U)$. If u is not in $S(U)$, then we add u to Y and recompute U. Since the ascending chain condition holds for subgroups of G, this process eventually stops.

The second method of computing normal closures works with the group $G\times G$. This group is polycyclic and we can easily get a polycyclic presentation for it. Let $b_1, \ldots, b_n$ be a sequence of generators distinct from $a_1, \ldots, a_n$. For each relation in the standard polycyclic presentation for G on the a_i's add the corresponding relation on the b_i's, so $\mathrm{Grp}\langle b_1, \ldots, b_n\rangle$ is isomorphic to G. Now add relations $b_i^\alpha a_j^\beta = a_j^\beta b_i^\alpha$ for all i and j and all α and β in $\{1, -1\}$. The group M generated by $a_1, \ldots, a_n$, $b_1, \ldots, b_n$ subject to these relations is isomorphic to $G \times G$. The subgroup $D = \mathrm{Grp}\langle a_1b_1, \ldots, a_nb_n\rangle$ corresponds to the diagonal subgroup of $G \times G$. There are two obvious polycyclic generating sequences for M. They are $a_1, \ldots, a_n$, $b_1, \ldots, b_n$ and $b_1, \ldots, b_n$, $a_1, \ldots, a_n$. The only difference between the standard polycyclic presentations for these two sequences is the ordering of the left and right sides in the relations $b_i^\alpha a_j^\beta = a_j^\beta b_i^\alpha$.

According to Exercise 3.9 in Chapter 1, there is a one-to-one correspondence between the set of normal subgroups of G and the set of subgroups of M containing D. In this correspondence, a normal subgroup N of G corresponds to DN. Since $N = (DN) \cap G$, the following result is easily proved.

Proposition 5.4. *Let Y be a subset of G and let H be the subgroup of M generated by Y and $a_1b_1, \ldots, a_nb_n$. Then the normal closure N of Y in G is $H \cap G$.*

Suppose that we compute in M using the polycyclic generating sequence $b_1, \ldots, b_n$, $a_1, \ldots, a_n$ and determine the full sequence $U = (h_1, \ldots, h_s)$ such that $H = S(U)$. Let h_r be the first component of U which is in G, that is, whose collected form involves only a_i's. Then $N = S(h_r, \ldots, h_s)$. Conceptually, this approach to normal closures is very appealing. We do not have to write a new routine. However, the presentation for M is more than twice as large as the presentation for G and it is not clear that this method is any faster than the first method.

Once normal closures can be computed, it is possible to compute commutator subgroups. If X and Y are finite generating sets for subgroups H and K of G, then $[H,K]$ is the normal closure in $\mathrm{Grp}\langle X,Y\rangle$ of the set of commutators $[x,y]$ with x in X and y in Y. With the ability to find commutator subgroups, we can determine the derived series and the lower central series of G.

Exercises

5.1. Show that POLY_MEMBER works correctly even if U does not satisfy condition (ii) of the definition of standard form.

5.2. Determine the order of the subgroup of the group in Example 4.4 generated by *abdf* and *ceg*.

5.3. Let D and g_1, g_2, and g_3 be as in Examples 5.2 to 5.4 Determine the normal closure of $\{g_1, g_2, g_3\}$ in D.

9.6 Homomorphisms

In this section, various techniques are discussed for working with a homomorphism $f: G \to H$ from one polycyclic group to another. We shall assume that we know the standard polycyclic presentation for G relative to a polycyclic generating sequence $a_1, \ldots, a_n$. We shall want to do such things as describe the image of f, determine the kernel of f, and compute inverse images of elements and subgroups. As usual, G_i will be $\mathrm{Grp}\langle a_i, \ldots, a_n\rangle$.

Let us start with the case in which H is a quotient G/N, where N is a normal subgroup of G. Here $f: G \to H$ is the natural homomorphism. Suppose we have a full sequence $U = (g_1, \ldots, g_s)$ such that $N = S(U)$. For $1 \le i \le n$ let $b_i = f(a_i)$. Then $b_1, \ldots, b_n$ is a polycyclic generating sequence for H. Set $H_i = \mathrm{Grp}\langle b_i, \ldots, b_n\rangle$. An obvious problem is to find a polycyclic presentation for H in terms of $b_1, \ldots, b_n$. The commutation relations present no problems. We just replace each a_i by b_i in the commutation relations defining G. The only question concerns the power relations for H.

Suppose that $1 \le i \le n$. If no g_j has a leading term which is a power of a_i, then the order of b_i modulo H_{i+1} is the same as the order of a_i modulo G_{i+1}. If there is a power relation $a_i^{m_i} = W_i$ for G, then the power relation for b_i is obtained by replacing a's by the corresponding b's in this relation. If some $g_j = a_i^{\alpha_i} \ldots a_n^{\alpha_n}$ with $\alpha_i > 0$, then α_i is the order of b_i modulo H_{i+1} and $b_i^{\alpha_i} = b_n^{-\alpha_n} \ldots b_{i+1}^{-\alpha_{i+1}}$ holds in H.

Example 6.1. The presentation

$$c^\alpha a^\beta = a^\beta c^\alpha, \quad c^\alpha b^\beta = b^\beta c^\alpha, \quad b^\alpha a^\beta = a^\beta b^\alpha c^{\alpha\beta},$$

where α and β range over $\{1, -1\}$, is consistent. Let G be the group defined. If $g_1 = a^3b^3$, $g_2 = b^6c$, and $g_3 = c^3$, then $U = (g_1, g_2, g_3)$ is in standard form and is full. Moreover, $N = S(U)$ is normal in G. If $u = aN$, $v = bN$, and $w = cN$, then in $H = G/N$ we have

$$u^3 = v^{-3}, \quad v^6 = w^{-1}, \quad w^3 = 1.$$

Reworking these relations slightly and transferring the commutation relations from G to H yields the following consistent power-commutator presentation for H:

$$wu = uw, \quad wv = vw, \quad vu = uvw,$$
$$w^3 = 1, \quad v^6 = w^2, \quad u^3 = v^3w.$$

If our normal subgroup N is not given as $S(U)$ but as the normal closure of a finite subset T of G, then we have two possible courses of action. We could obtain a description of N as $S(U)$ using the techniques at the end of the previous section. We could also add the relations $t = 1$ for t in T as new defining relations. Here we are thinking of T as a set of collected words. The Knuth-Bendix procedure for strings can now be used to get the standard polycyclic presentation for H.

Now let us assume that H is a second polycyclic group described by a consistent polycyclic presentation on generators $b_1, \ldots, b_m$. A homomorphism $f: G \to H$ is determined by the images $u_i = f(a_i)$, $1 \le i \le n$. A map $a_i \mapsto u_i$ of the generators of G into H defines a homomorphism if and only if the u_i satisfy the defining relations for G. This can be checked, since we can compute collected words in H. The image K of f is generated by the u_i and thus can be determined. We also need to compute the kernel of f, and, for an element k of K, we want to be able to find g in G such that $f(g) = k$.

Essentially all the information we need to compute with f can be obtained with one invocation of POLY_SUBGROUP in $M = H \times G$. This approach is similar to the second method for finding normal closures described at the end of Section 9.5. Assuming the generating sets for G and H are disjoint, we get a presentation for M on the a_i's and the b_j's by combining the relations for G and H and adding relations which say that each a_i commutes with each b_j. All computation in M will be done using the polycyclic generating sequence $b_1, \ldots, b_m, a_1, \ldots, a_n$. Let L be the subgroup of M generated by the elements $u_1a_1, \ldots, u_na_n$ and let $W = (w_1, \ldots, w_s)$ be the full sequence such that $L = S(W)$. Each w_i can be written uniquely as h_ig_i, where h_i is in H and g_i is in G. Let r be the largest index such that h_r is not trivial.

Proposition 6.1. *The group L consists of all elements of M of the form $f(g)g$ with g in G. The sequence $U = (h_1, \ldots, h_r)$ is full and $K = S(U)$. The sequence $V = (g_{r+1}, \ldots, g_s)$ is full and $S(V)$ is the kernel of f. If k is in K and $k = h_1^{\beta_1} \ldots h_r^{\beta_r}$, then $f(g) = k$, where $g = g_1^{\beta_1} \ldots g_r^{\beta_r}$.*

Proof. The set P of products $f(g)g$ with g in G is easily checked to be a subgroup of M and the elements $u_i a_i$ are all in P. Thus L is contained in P. Given $g = a_1^{\alpha_1} \ldots a_n^{\alpha_n}$ in G, the element $(u_1 a_1)^{\alpha_1} \cdots (u_n a_n)^{\alpha_n} = u_1^{\alpha_1} \ldots u_n^{\alpha_n} a_1^{\alpha_1} \ldots a_n^{\alpha_n} = f(g)g$. Therefore $L = P$. If $1 \le i \le r$, then the leading term of $h_i g_i$ is the leading term of h_i. If $r + 1 \le i \le s$, then the leading term of $h_i g_i$ is the leading term of g_i. Since W is in standard form, it is easy to check that U and V are in standard form and $E(W)$ is the set of s-tuples $(\beta_1, \ldots, \beta_r, \alpha_{r+1}, \ldots, \alpha_s)$, where $(\beta_1, \ldots, \beta_r)$ is in $E(U)$ and $(\alpha_{r+1}, \ldots, \alpha_s)$ is in $E(V)$. Clearly $S(U) \subseteq K$. However, for any g in G the element $f(g)g$ is in $S(W)$. Thus $f(g)g = (h_1 g_1)^{\beta_1} \cdots (h_r g_r)^{\beta_r} g_{r+1}^{\alpha_{r+1}} \ldots g_s^{\alpha_s}$, where $(\beta_1, \ldots, \beta_r, \alpha_{r+1}, \ldots, \alpha_s)$ is in $E(W)$. Therefore $f(g) = h_1^{\beta_1} \ldots h_r^{\beta_r}$ and $f(g)$ is in $S(U)$. Hence $K = S(U)$. If $f(g) = 1$, then $\beta_1 = \cdots = \beta_r = 0$. Thus g is in $S(V)$. Since each g_i with $i > r$ is in the kernel of f, the kernel is equal to $S(V)$. Finally, if $k = h_1^{\beta_1} \ldots h_r^{\beta_r}$, then $(h_1 g_1)^{\beta_r} \cdots (h_r g_r)^{\beta_r} = kg$ is in L, where $g = g_1^{\beta_1} \ldots g_r^{\beta_r}$. Therefore $f(g) = k$. □

Example 6.2. Let G be the group defined by the following power-conjugate presentation on the generators a, b, c:

$$c^8 = 1,$$
$$b^8 = 1, \quad cb = bc,$$
$$a^2 = b^2c^2, \quad ba = ac, \quad ca = ac.$$

Let H be the group on generators u and v defined by the presentation

$$v^8 = 1,$$
$$u^4 = 1, \quad vu = uv.$$

The map $a \mapsto u^2v^4$, $b \mapsto v^2$, $c \mapsto v^6$ defines a homomorphism f of G into H. To determine the kernel of f using Proposition 6.1, we apply POLY_SUBGROUP to $W = (u^2v^4a,\ v^2b,\ v^6c)$ in $M = H \times G$ using the polycyclic generating sequence u, v, a, b, c. The matrix associated with W is

$$\begin{bmatrix} 2 & 4 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 6 & 0 & 0 & 1 \end{bmatrix}.$$

The resulting full sequence has the following associated matrix:

$$\begin{bmatrix} 2 & 0 & 1 & 0 & 2 \\ 0 & 2 & 0 & 0 & 7 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}.$$

Thus the kernel of f is generated by bc and c^4.

Exercises

6.1. Use the ideas in the discussion following Example 6.1 to devise a third algorithm for computing normal closures in polycyclic groups based on the Knuth-Bendix procedure.

6.2. Let G be the group generated by a, b, c, d subject to the relations

$$ba = ac, \quad ca = ad, \quad da = ab, \quad cb = bc, \quad db = bd, \quad dc = cd,$$
$$a^3 = b^4 = c^4 = d^4 = 1,$$

and let H be a cyclic group of order 12 generated by an element u. The map

$$a \mapsto u^4, \quad b \mapsto u^3, \quad c \mapsto u^3, \quad d \mapsto u^3$$

extends to a homomorphism f of G onto H. Find the kernel of f.

9.7 Conjugacy in nilpotent groups

As noted at the end of Section 9.4, the conjugacy problem in polycyclic groups is solvable in principle, but the available algorithms are not practical for most infinite polycyclic groups. However, for finitely generated nilpotent groups the situation is much better. This section describes the computation of centralizers and the determination of conjugacy in nilpotent polycyclic groups.

Let G be a finitely generated nilpotent group. We shall assume that G is given by a standard polycyclic presentation on generators $a_1, \ldots, a_n$ and that this presentation is nilpotent as defined in Section 9.4. If the presentation is not nilpotent, then we can compute the lower central series of G, determine a polycyclic series which refines the lower central series, make a corresponding choice of a polycyclic generating sequence, and determine the standard polycyclic presentation of G with respect to the new generators.

The first problem to be considered is the computation of centralizers. Let g be an element of G. The following recursive procedure may be used to compute the centralizer $C_G(g)$ of g in G. The subgroup $N = \mathrm{Grp}\langle a_n \rangle$ is contained in the center of G and hence is normal. By induction on n, we can compute the centralizer K of gN in G/N and find the inverse image L

of K in G. The subgroup L is the set of all elements u of G such that $u^{-1}gu$ is in gN, or, equivalently, $f(u) = [g,u] = g^{-1}u^{-1}gu$ is in N. Restricted to L, the function f is a homomorphism into N, for if u_1 and u_2 are in L, then $f(u_1u_2) = [g,u_1u_2]$, which by Proposition 1.6(b) is $[g,u_2][g,u_1][g,u_1,u_2]$. But $[g,u_1]$ and $[g,u_2]$ are in N, so they commute. Also, $[g,u_1,u_2] = 1$. Thus $f(u_1u_2) = [g,u_1][g,u_2]$. The centralizer $C_G(g)$ is the kernel of f. Since N is a cyclic group, the computation of the kernel of f is an easy application of the methods of Section 9.6.

Example 7.1. Let us consider the group $D = D_4^{(1)}(\mathbb{Z})$ of Examples 2.1, 3.1, 4.1, 4.5, 5.2, 5.3, and 5.4. We shall find the centralizer of $g = a_1a_3^2$ in D. For $1 \le i \le 7$, let $N_i = \operatorname{Grp}\langle a_i, \dots, a_6\rangle$. Our approach is to compute the centralizer C_i of gN_i in D/N_i, $1 \le i \le 7$. The group we really want is C_7. For $1 \le i \le 4$, the group D/N_i is abelian, so $C_i = D/N_i$. The inverse image L_5 in D/N_5 of C_4 is generated by the images of a_1, a_2, a_3, and a_4. Working modulo N_5, we have

$$g^{-1}a_1^{-1}ga_1 \equiv 1, \quad g^{-1}a_2^{-1}ga_2 \equiv a_4, \quad g^{-1}a_3^{-1}ga_3 \equiv 1, \quad g^{-1}a_4^{-1}ga_4 \equiv 1.$$

The kernel C_5 of the homomorphism from L_5 into N_4/N_5 is generated by the images of a_1, a_3, and a_4. The inverse image L_6 in D/N_6 of C_5 is generated by the images of a_1, a_3, a_4, and a_5.

Modulo N_6, we have

$$g^{-1}a_1^{-1}ga_1 \equiv 1, \quad g^{-1}a_3^{-1}ga_3 \equiv 1, \quad g^{-1}a_4^{-1}ga_4 \equiv 1, \quad g^{-1}a_5^{-1}ga_5 \equiv 1.$$

Thus $C_6 = L_6$ and its inverse image L_7 in $D = D/N_7$ is generated by a_1, a_3, a_4, a_5, and a_6. In D,

$$g^{-1}a_1^{-1}ga_1 = 1, \quad g^{-1}a_3^{-1}ga_3 = 1, \quad g^{-1}a_4^{-1}ga_4 = a_6^{-2},$$
$$g^{-1}a_5^{-1}ga_5 = a_6, \quad g^{-1}a_6^{-1}ga_6 = 1.$$

The kernel C_7 of the homomorphism from L_7 to N_6 is $\operatorname{Grp}\langle a_1, a_3, a_4a_5^2, a_6\rangle$.

The determination of conjugacy involves only a slight extension of the algorithm for computing centralizers. Now we have two elements g and h of G and we want to decide whether g and h are conjugate. We consider gN and hN in G/N, where $N = \operatorname{Grp}\langle a_n\rangle$. If gN and hN are not conjugate, then g and h are not conjugate in G. If gN and hN are conjugate, then we can find an element u of G such that $u^{-1}(hN)u = u^{-1}huN$ is gN. Replacing h by $u^{-1}hu$, we may assume that $gN = hN$. In this case, if g and h are conjugate, the conjugating element lies in the inverse image L of the centralizer in G/N of gN. We know how to compute L. Let f be the homomorphism from L to N mapping v to $[g,v]$. The conjugates of g by elements of L are the elements of $gf(L)$. Let $w = g^{-1}h$. Then h is conjugate to g if and only if w is in $f(L)$. We can decide this and, if w is in $f(L)$, we can find an element v of L such that $f(v) = w$. In this case $h = v^{-1}gv$.

Exercise

7.1. In the group D of Example 7.1 determine the centralizers of the elements $a_1a_4a_5^3$ and $a_2^2a_4^4a_6$.

9.8 Cyclic extensions

A group G is a *cyclic extension* of a group N if N is a normal subgroup of G and G/N is cyclic. In this section we shall take the point of view that N is given and we wish to construct cyclic extensions of N. A polycyclic group is a group which can be built up by starting with the trivial group, constructing a cyclic extension, then making a cyclic extension of that group, and continuing a finite number of times. The theory of cyclic extensions is well understood. The exposition here is based in large part on Sections III.7 and III.8 of [Zassenhaus 1958].

We shall look first at the case in which G/N is infinite cyclic. Let x be an element of G such that xN generates G/N. Since N is normal, the map σ taking an element v of N to $x^{-1}vx$ is an automorphism of N. The image of v under σ will be written v^σ. The group G is determined up to isomorphism by N and σ, since any g in G can be represented uniquely as x^iv with v in N and

$$(x^iv)(x^jw) = x^ix^jx^{-j}vx^jw = x^{i+j}v^{\sigma^j}w.$$

Moreover, any automorphism of N can occur this way, for if σ is an automorphism of N, then the binary operation

$$(i,v)(j,w) = (i+j, v^{\sigma^j}w)$$

defines a group structure on $G = \mathbb{Z} \times N$. The identity element of G is $(0, 1_N)$. The inverse of (i,v) is $(-i, (v^{-1})^{\sigma^{-i}})$. The associative law is proved as follows.

$$(i,a)[(j,b)(k,c)] = (i,a)(j+k, b^{\sigma^k}c) = (i+j+k, a^{\sigma^{j+k}}b^{\sigma^k}c),$$
$$[(i,a)(j,b)](k,c) = (i+j, a^{\sigma^j}b)(k,c) = (i+j+k, \left(a^{\sigma^j}b\right)^{\sigma^k}c).$$

Since σ is an automorphism,

$$\left(a^{\sigma^j}b\right)^{\sigma^k} = \left(a^{\sigma^j}\right)^{\sigma^k}b^{\sigma_k} = a^{\sigma^{j+k}}b^{\sigma^k}.$$

If σ is the identity automorphism of N, then G is the ordinary direct product of $\mathbb{Z}$ and N. For any σ, if we identify $\{0\} \times N$ with N, then N is a normal subgroup of G and G/N is isomorphic to $\mathbb{Z}$.

There is an alternative way to see that any automorphism of N leads to a cyclic extension G of N such that G/N is infinite. Let x and x^{-1} be objects not in N and set $Y = N \cup \{x, x^{-1}\}$. Let $\mathcal{R}$ be the rewriting system on Y^* consisting of the following rules: the semigroup relations from the multiplication table of N as described in Section 2.3, the rules $xx^{-1} \to \varepsilon$ and $x^{-1}x \to \varepsilon$, and the rules $vx \to xv^\sigma$ and $vx^{-1} \to x^{-1}v^{\sigma^{-1}}$ for v in N. Without too much difficulty, it is possible to show that $\mathcal{R}$ is confluent. If S is the ideal of Y^* generated by N, then $(Y, \mathcal{R}, S)$ is a restricted presentation for the group G.

Now let us turn to the case in which G is a cyclic extension of N and N has finite index n in G. Again choose an element x such that xN generates G/N and let σ be the automorphism of N induced by x. Since the coset xN has order n in G/N, it follows that $(xN)^n = x^nN = N$. Therefore $u = x^n$ is an element of N. Knowing N, σ, n, and u determines G up to isomorphism. Every element of G can be expressed uniquely as x^iv, where v is in N and $0 \le i < n$. Also

$$(x^iv)(x^jw) = \begin{cases} x^{i+j}v^{\sigma^j}w, & i+j<n, \\ x^{i+j-n}uv^{\sigma^j}w, & i+j \ge n. \end{cases}$$

Not all pairs (σ, u) can occur. Since $u = x^n$, it follows that $u^\sigma = x^{-1}x^nx = x^n = u$. Thus σ fixes u. Also, for any v in N we have $v^{\sigma^n} = x^{-n}vx^n = u^{-1}vu$. Therefore σ^n is the inner automorphism of N induced by u. These two conditions are sufficient for the pair (σ, u) to arise in a cyclic extension G of N with G/N of order n. Again there are two ways to see this. We can define a binary operation on $\{0, 1, \dots, n-1\} \times N$ by the formula

$$(i, v)(j, w) = \begin{cases} (i+j, v^{\sigma^j}w), & i+j<n, \\ (i+j-n, uv^{\sigma^j}w), & i+j \ge n, \end{cases}$$

and prove the structure defined is a group. On the other hand, we can choose an object x not in N, define $Y = N \cup \{x\}$, and form the rewriting system $\mathcal{R}$ consisting of the multiplication-table rules for N, the rule $x^n \to u$, and the rules $vx \to xv^\sigma$ for v in N. We then must prove that $\mathcal{R}$ is confluent and that $(Y, \mathcal{R}, S)$ is a restricted presentation for a group, where S is the ideal of Y^* generated by N. The amounts of work involved in the two approaches are roughly the same, and the choice reduces to a matter of taste.

The point of view of cyclic extensions gives us a new perspective on the test for consistency of a presentation with the form of a standard polycyclic presentation. In such a presentation we have generators $a_1, \ldots, a_n$, a subset I of $\{1, \ldots, n\}$, and for each i in I an integer m_i greater than 1. In this discussion we shall use the monoid version of the standard polycyclic presentation in which there is a generator a_i^{-1} only when i is not in I. The set $\mathcal{R}$ of monoid relations has the form

$$\begin{aligned}
a_i a_1^{-1} &= 1, \quad a_1^{-1} a_i = 1, \quad i \notin I, \\
a_j a_i &= a_i a_{i+1}^{\alpha_{iji+1}} \ldots a_n^{\alpha_{ijn}}, \quad j > i, \\
a_j^{-1} a_i &= a_i a_{i+1}^{\beta_{iji+1}} \ldots a_n^{\beta_{ijn}}, \quad j > i,\ j \notin I, \\
a_j a_i^{-1} &= a_i^{-1} a_{i+1}^{\gamma_{iji+1}} \ldots a_n^{\gamma_{ijn}}, \quad j > i,\ j \notin I, \\
a_j^{-1} a_i^{-1} &= a_i^{-1} a_{i+1}^{\delta_{iji+1}} \ldots a_n^{\delta_{ijn}}, \quad j > i,\ i, j \notin I, \\
a_i^{m_i} &= a_{i+1}^{\mu_{ij+1}} \ldots a_n^{\mu_{in}}, \quad i \in I,
\end{aligned}$$

where the right sides are collected in the sense that the integers α_{ijk}, β_{ijk}, γ_{ijk}, δ_{ijk}, and μ_{ik} are between 0 and $m_k - 1$ when k is in I. Let Y be the set of generators occurring in $\mathcal{R}$. The monoid G defined by $(Y, \mathcal{R})$ is a group. If $(Y, \mathcal{R})$ is consistent, then many of the relations in $\mathcal{R}$ are redundant. To prove this, we shall use a criterion equivalent to consistency. Let $Y_i = Y \cap \{a_i, \ldots, a_n\}^{\pm}$, let $\mathcal{R}_i$ consist of the relations in $\mathcal{R}$ which involve only generators in Y_i, and let G_i be the subgroup of G generated by Y_i.

Proposition 8.1. *The presentation $(Y, \mathcal{R})$ is consistent if and only if for $1 \leq i \leq n$ any relation in G of the form $U = V$ with U and V in Y_i^* is a consequence of the relations in $\mathcal{R}_i$.*

Proof. Suppose that $(Y, \mathcal{R})$ is consistent. Then any word in Y_i^* can be rewritten into collected form using the relations in $\mathcal{R}_i$, interpreted as rewriting rules. If the relation $U = V$ holds in G and both U and V are in Y_i^*, then U and V have the same collected form W, and the relations $U = W$ and $V = W$ are consequences of $\mathcal{R}_i$. Thus $U = V$ is a consequence of $\mathcal{R}_i$.

Now suppose that any relation $U = V$ in G with U and V in Y_i^* is a consequence of $\mathcal{R}_i$. If $(Y, \mathcal{R})$ is not consistent, then for some i, $1 \leq i \leq n$, there is an integer $m > 0$ such that a_i^m is in G_{i+1} and either i is not in I or i is in I and $m < m_i$. The relations in $\mathcal{R}_i$ are satisfied if we set a_j equal to 1 for $i < j \leq n$. If this is done, then the group defined is infinite cyclic if i is not in I or cyclic of order m_i if i is in I. Therefore in the group generated by Y_i and defined by $\mathcal{R}_i$ no relation $a_i^m = W$ with W in Y_{i+1}^* can hold. Therefore $(Y, \mathcal{R})$ is consistent. □

Proposition 8.2. *If $(Y, \mathcal{R})$ is consistent, then the relations in $\mathcal{R}$ with positive left sides define G as a group.*

Proof. Assume that $(Y, \mathcal{R})$ is consistent and let $\mathcal{S}_i$ denote the set of relations in $\mathcal{R}$ with left sides $a_k a_j$ or $a_j^{m_j}$, where $i \le j$. Thus $\mathcal{S}_n$ is empty if n is not in I and $\mathcal{S}_n$ consists of the single relation $a_n^{m_n} = 1$ if n is in I. In either case, $\mathcal{S}_n$ defines G_n as a group. Now suppose that we know $\mathcal{S}_{i+1}$ defines G_{i+1}. To show that $\mathcal{S}_i$ defines G_i, we must prove that all the relations in $\mathcal{R}_i$ are consequences of $\mathcal{S}_i$. (Since we are considering $\mathcal{S}_i$ to be a set of group relations, the relations $a_i a_i^{-1} = a_i^{-1} a_i = 1$ come for free.) By assumption, the relations in $\mathcal{R}_{i+1}$ are consequences of $\mathcal{S}_{i+1}$ and hence of $\mathcal{S}_i$. Suppose that $j > i$ and j is not in I. The relation

$$a_j^{-1} a_i = a_i a_{i+1}^{\beta_{iji+1}} \dots a_n^{\beta_{ijn}}$$

is equivalent to

$$a_i = a_j a_i a_{i+1}^{\beta_{iji+1}} \dots a_n^{\beta_{ijn}},$$

which is equivalent modulo the relations in $\mathcal{S}_i$ to

$$a_i = a_i a_{i+1}^{\alpha_{iji+1}} \dots a_n^{\alpha_{ijn}} a_{i+1}^{\beta_{iji+1}} \dots a_n^{\beta_{ijn}},$$

which in turn is equivalent to

$$1 = a_{i+1}^{\alpha_{iji+1}} \dots a_n^{\alpha_{ijn}} a_{i+1}^{\beta_{iji+1}} \dots a_n^{\beta_{ijn}}.$$

This is a relation in G_{i+1} and hence is a consequence of $\mathcal{R}_{i+1}$ and thus of $\mathcal{S}_i$.

Suppose that i is not in I and $j > i$. The relation

$$a_j a_i^{-1} = a_i^{-1} a_{i+1}^{\gamma_{iji+1}} \dots a_n^{\gamma_{ijn}}$$

is equivalent to

$$a_j = a_i^{-1} a_{i+1}^{\gamma_{iji+1}} \dots a_n^{\gamma_{ijn}} a_i.$$

Using only relations with left sides $a_k^{\eta} a_i$, $\eta = \pm 1$, and $a_i^{-1} a_i = 1$, one can rewrite the right side of this relation into a word W in Y_{i+1}^*. The relation $a_j = W$ is a consequence of $\mathcal{R}_{i+1}$ and hence of $\mathcal{S}_i$. Therefore the relation

$$a_j a_i^{-1} = a_i^{-1} a_{i+1}^{\gamma_{iji+1}} \dots a_n^{\gamma_{ijn}}$$

is a consequence of $\mathcal{S}_i$. The relations with left sides $a_j^{-1} a_i^{-1}$ are handled like those with left sides $a_j^{-1} a_i$. By induction on $n - i$, the relations in $\mathcal{S}_1$ define $G_1 = G$. □

Suppose now that we have a presentation of the form $(Y,\mathcal{R})$ as before and we want to decide whether it is consistent. If $n=1$, then either there are the two relations $a_1a_1^{-1}=a_1^{-1}a_1=1$ or there is the single relation $a_1^{m_1}=1$. In either case the presentation is consistent. Let us assume that $n>1$. The group G defined by $(Y,\mathcal{R})$ is a cyclic extension of the subgroup G_2 generated by $a_2,\dots,a_n$. By induction on n, we may assume that we have already checked that the relations in $(Y,\mathcal{R})$ which involve only $a_2,\dots,a_n$ and their inverses form a consistent presentation of a group K. We can compute products and test equality of elements in K. There is an obvious homomorphism from K onto G_2. By Proposition 8.1, $(Y,\mathcal{R})$ is consistent if and only if this homomorphism is an isomorphism. This can be decided by checking that the conditions for a cyclic extension are satisfied.

We first test whether the map of generators

$$a_j \mapsto a_2^{\alpha_{1j2}}\dots a_n^{\alpha_{1jn}}, \quad 2\le j\le n,$$

extends to a homomorphism σ of K into itself. This is done by checking whether the images of $a_2,\dots,a_n$ satisfy a set of defining relations for K. Proposition 8.2 can be used to reduce the number of relations which must be checked. Assuming that σ is defined, we next need to decide whether σ is surjective. This could be done by showing that $a_2^\sigma,\dots,a_n^\sigma$ generate K. However, if 1 is not in I, then it is quicker to test whether

$$(a_2^{\gamma_{1j2}}\dots a_n^{\gamma_{1jn}})^\sigma = a_j, \quad 2\le j\le n.$$

If 1 is in I, then we shall have to test whether σ^{m_1} is the inner automorphism of K induced by $a_1^{m_1}$, and a positive result implies that σ maps K onto itself. If σ is surjective, then σ is an automorphism of K, since K is hopfian by Corollary 3.11. To complete the first phase of our consistency check, we test whether

$$a_2^{\beta_{1j2}}\dots a_n^{\beta_{1jn}} = (a_2^{\alpha_{1j2}}\dots a_n^{\alpha_{1jn}})^{-1}$$

in K if j is not in I and whether

$$a_2^{\delta_{1j2}}\dots a_n^{\delta_{1jn}} = (a_2^{\gamma_{1j2}}\dots a_n^{\gamma_{1jn}})^{-1}$$

if 1 and j are not in I.

Suppose that all the tests so far have been successful. Then the conditions for an infinite cyclic extension of K are satisfied. If 1 is not in I, then G_2 is isomorphic to K and $(Y,\mathcal{R})$ is consistent. However, if 1 is in I, then there is more work to do. Let u be the element $a_2^{\mu_{12}}\dots a_n^{\mu_{1n}}$ of K. We must check whether $u^\sigma=u$ and whether $a_j^{\sigma^{m_1}}=u^{-1}a_ju$, $2\le j\le n$. If these conditions are satisfied, then the pair (σ, u) defines a cyclic extension of K with quotient of order m_1.

We can rephrase the previous discussion in terms of rewriting rules. The following approach is inspired in part by (Vaughan-Lee 1984, 1985). Let us interpret the relations in $\mathcal{R}$ as rewriting rules. Among the overlaps of left sides in $\mathcal{R}$ are the following:

$$\begin{array}{rll} a_k a_j a_i, & k > j > i, & \\ a_j^{m_j} a_i, & j \in I, & j > i, \\ a_j a_i^{m_i}, & i \in I, & j > i, \\ a_j a_i^{-1} a_i, & i \notin I, & j > i, \\ a_i^{m_i+1}, & i \in I, & \\ a_j^{-1} a_j a_i, & j \notin I, & j > i, \\ a_j^{-1} a_j a_i^{-1}, & i, j \notin I, & j > i. \end{array} \tag{$*$}$$

Proposition 8.3. *If local confluence holds at the overlaps $(*)$, then $\mathcal{R}$ is confluent.*

Proof. Let $\prec$ be the basic wreath-product ordering of Y^* with $a_n \prec \cdots \prec a_1$, $a_1 \prec a_i^{-1}$ if 1 is not in I, and $a_i \prec a_i^{-1} \prec a_{i-1}$ if $i > 1$ and i is not in I. To simplify the exposition, let us introduce the following notation:

$$\begin{aligned} S_{ij} &= a_{i+1}^{\alpha_{iji+1}} \dots a_n^{\alpha_{ijn}}, \\ T_{ij} &= a_{i+1}^{\beta_{iji+1}} \dots a_n^{\beta_{ijn}}, \\ U_{ij} &= a_{i+1}^{\gamma_{iji+1}} \dots a_n^{\gamma_{ijn}}, \\ V_{ij} &= a_{i+1}^{\delta_{iji+1}} \dots a_n^{\delta_{ijn}}, \\ W_i &= a_{i+1}^{\mu_{ij+1}} \dots a_n^{\mu_{in}}. \end{aligned}$$

Suppose that $\mathcal{R}$ is not consistent. By induction on n, we may assume that the relations not involving a_1 or a_1^{-1} form a consistent presentation for a group K. This means that, if $k > j > 1$ and the indicated words are defined, then S_{jk} and T_{jk} represent inverse elements in K, as do U_{jk} and V_{jk}. Also, $U_{jk} a_j$ and $a_j a_k$ represent the same element of K. The notation $Q \to R$ will be used to indicate that R is obtained from Q by the application of one rewriting rule in $\mathcal{R}$, while $Q \xrightarrow{*} R$ will signal that zero or more rules have been used. $\square$

Lemma 8.4. *If $j > 1$ and j is not in I, then S_{1j} and T_{1j} represent inverse elements of K.*

Proof. By assumption, there is local confluence at the word $a_j^{-1}a_ja_1$. The first few steps in processing this overlap are uniquely determined. They are

$$a_j^{-1}a_ja_1 \to a_1$$

and

$$a_j^{-1}a_ja_1 \to a_j^{-1}a_1S_{1j} \to a_1T_{1j}S_{1j}.$$

The word a_1 is irreducible with respect to $\mathcal{R}$. Since local confluence holds, there is a reduction of $T_{1j}S_{1j}$ to the empty word using $\mathcal{R}$. This means that S_{1j} and T_{1j} represent inverse elements of K. $\square$

Let P be the first word with respect to $\prec$ at which confluence fails. By Proposition 7.1 in Chapter 2, P is an overlap containing exactly two left sides. Local confluence fails at P and P contains a_1 or a_1^{-1}. Since P is not one of the overlaps $(*)$, P must have one of the following forms:

(1) $a_k^{-1}a_ja_1$,	(9) $a_ka_j^{-1}a_1^{-1}$,
(2) $a_ka_j^{-1}a_1$,	(10) $a_k^{-1}a_j^{-1}a_1^{-1}$,
(3) $a_k^{-1}a_j^{-1}a_1$,	(11) $a_j^{-1}a_1^{-1}a_1$,
(4) $a_ja_j^{-1}a_1$,	(12) $a_ja_j^{-1}a_1^{-1}$,
(5) $a_j^{-1}a_1^{m_1}$,	(13) $a_ja_1a_1^{-1}$,
(6) $a_ka_ja_1^{-1}$,	(14) $a_j^{-1}a_1a_1^{-1}$,
(7) $a_k^{-1}a_ja_1^{-1}$,	(15) $a_1a_1^{-1}a_1$,
(8) $a_j^{m_j}a_1^{-1}$,	(16) $a_1^{-1}a_1a_1^{-1}$.

For each of these forms there are assumptions that one or more of the indices involved does or does not belong to I. For example, in (5) j is not in I and 1 is in I. Local confluence clearly holds in cases (15) and (16). There are many similarities in the consideration of the other 14 cases. Only a few cases will be discussed in detail. The remaining ones are left as exercises.

Case (1). Suppose that P has the form (1) for some indices j and k with $1 < j < k \le n$ and k not in I. Let y and z be the elements of K defined by S_{1j} and S_{1k}, respectively. Then T_{1k} defines z^{-1}. The word $Q = S_{jk}T_{jk}$ is in Y_{j+1}^*, so Qa_1 precedes P with respect to $\prec$. Therefore confluence holds at Qa_1. Let M and N be any words such that $S_{jk}a_1 \xrightarrow{*} a_1M$ and $T_{jk}a_1 \xrightarrow{*} a_1N$. Then we have the reductions

$$S_{jk}T_{jk}a_1 \xrightarrow{*} S_{jk}a_1N \xrightarrow{*} a_1MN$$

and

$$S_{jk}T_{jk}a_1 \xrightarrow{*} a_1,$$

since S_{jk} and T_{jk} represent inverse elements of K. Therefore MN must reduce to the empty word. Hence, if M represents the element u of K, then N represents u^{-1}.

By assumption, local confluence holds at $a_k a_j a_1$. The initial reductions here are

$$a_k a_j a_1 \to a_k a_1 S_{1j} \to a_1 S_{1k} S_{1j}$$

and

$$a_k a_j a_1 \to a_j S_{jk} a_1 \xrightarrow{*} a_j a_1 M \to a_1 S_{1j} M.$$

Hence $S_{1k}S_{1j}$ and $S_{1j}M$ define the same element of K. Therefore $zy = yu$, so $yu^{-1} = z^{-1}y$.

The initial reductions in processing the overlap $a_k^{-1}a_j a_1$ are

$$a_k^{-1}a_j a_1 \to a_k^{-1}a_1 S_{1j} \to a_1 T_{1k} S_{1j}$$

and

$$a_k^{-1}a_j a_1 \to a_j T_{jk} a_1 \xrightarrow{*} a_j a_1 N \to a_1 S_{1j} N.$$

Now $T_{1k}S_{1j}$ defines $z^{-1}y$ and $S_{1j}N$ defines yu^{-1}. By the previous remark, if R is the reduced word representing $z^{-1}y$, then $a_1 R$ is derivable from both $a_1 T_{1k} S_{1j}$ and $a_1 S_{1j} N$. Thus local confluence holds at $a_k^{-1}a_j a_1$ after all.

Case (2). Suppose that P has the form (2) for some j and k with $1 < j < k \le n$ and j not in I. Let y and z be as in case (1). Then T_{1j} represents y^{-1}. Let L be any word such that $U_{jk}a_1 \xrightarrow{*} a_1 L$ and let v be the element of K represented by L. Since $U_{jk}a_j$ and $a_j a_k$ represent the same element of K and $a_j a_k$ is reduced, we have $U_{jk}a_j \xrightarrow{*} a_j a_k$. The word $U_{jk}a_j a_1$ precedes P with respect to $\prec$, so confluence holds at $U_{jk}a_j a_1$. Now

$$U_{jk}a_j a_1 \to U_{jk}a_1 S_{1j} \xrightarrow{*} a_1 L S_{1j}$$

and

$$U_{jk}a_j a_1 \xrightarrow{*} a_j a_k a_1 \to a_j a_1 S_{1k} \to a_1 S_{1j} S_{1k}.$$

By confluence, $vy = yz$, so $zy^{-1} = y^{-1}v$.

Processing the overlap P, we have

$$a_k a_j^{-1} a_1 \to a_k a_1 T_{1j} \to a_1 S_{1k} T_{1j}$$

and

$$a_k a_j^{-1} a_1 \to a_j^{-1} U_{jk} a_1 \xrightarrow{*} a_j^{-1} a_1 L \to a_1 T_{1j} L.$$

If R is the reduced word defining zy^{-1}, then $a_1 R$ is derivable from both $a_1 S_{1k} T_{1j}$ and $a_1 T_{1j} L$. Therefore local confluence holds at P.

Case (5). Suppose that P has the form (5) with j not in I and 1 in I. Let M and N be words such that $S_{1j} a_1^{m_1-1} \xrightarrow{*} a_1^{m_1-1} M$ and $T_{1j} a_1^{m_1-1} \xrightarrow{*} a_1^{m_1-1} N$, and let u be the element of K represented by M. The word $S_{1j} T_{1j} a_1^{m_1-1}$ precedes P with respect to $\prec$. By essentially the same argument as used in case (1), N represents u^{-1}.

By assumption, local confluence holds at $a_j a_1^{m_1-1}$. The initial reductions at that word are

$$a_j a_1^{m_1} \to a_j W_1$$

and

$$a_j a_1^{m_1} \to a_1 S_{ij} a_1^{m_1-1} \xrightarrow{*} a_1^{m_1} M \to W_1 M.$$

Therefore $a_j W_1$ and $W_1 M$ represent the same element of K. From this it follows that $a_j^{-1} W_1$ and $W_1 N$ represent the same element.

The initial reductions at P are

$$a_j^{-1} a_1^{m_1} \to a_j^{-1} W_1$$

and

$$a_j^{-1} a_1^{m_1} \to a_1 T_{1j} a_1^{m_1-1} \xrightarrow{*} a_1^{m_1} N \to W_1 N.$$

By the remark above, local confluence holds.

If P is not of the first five forms, then confluence holds at all words not involving a_1^{-1}. Thus we may assume that 1 is not in I. We can define a map $^{-}$ from Y_2^* to itself as follows: Given Q in Y_2^*, define $\overline{Q}$ by $Qa_1 \xrightarrow{*} a_1 \overline{Q}$ and $a_1 \overline{Q}$ is irreducible with respect to $\mathcal{R}$. Confluence at Qa_1 implies that $\overline{Q}$ is well defined. If $Q \xrightarrow{*} R$, then one of the ways to reduce Qa_1 is

$$Qa_1 \xrightarrow{*} Ra_1 \xrightarrow{*} a_1 \overline{R},$$

so $\overline{Q} = \overline{R}$. This means that we can define a map $\sigma: K \to K$ which takes the element represented by Q to the element represented by $\overline{Q}$.

It is easy to check that σ is a homomorphism. If $j > 1$, then we have local confluence at $a_j a_1^{-1} a_1$. The two reductions of this word are

$$a_j a_1^{-1} a_1 \to a_j$$

and

$$a_j a_1^{-1} a_1 \to a_1^{-1} U_{jk} a_1 \xrightarrow{*} a_1^{-1} a_1 M \to M \xrightarrow{*} \overline{U_{jk}},$$

where M is some word such that $U_{jk} a_1 \xrightarrow{*} a_1 M$. Thus $\overline{U_{jk}} = a_j$. Since the image of σ is a subgroup of K, that image is all of K. Therefore, by Corollary 3.11, σ is an automorphism of K. This mean that if Q and R are in Y_2^* and $\overline{Q} = \overline{R}$, then Q and R represent the same element of K.

Case (6). Suppose that P has the form (6) for some j and k with $1 < j < k \le n$. Let L be any word such that $S_{jk} a_1^{-1} \xrightarrow{*} a_1^{-1} L$. Consider the reductions

$$S_{jk} a_1^{-1} a_1 \to S_{jk}$$

and

$$S_{jk} a_1^{-1} a_1 \xrightarrow{*} a_1^{-1} L a_1 \xrightarrow{*} a_1^{-1} a_1 \overline{L} \to \overline{L}.$$

Since $S_{jk} a_1^{-1} a_1$ precedes P, confluence holds and $\overline{L} = S_{jk}$.

Processing the overlap P leads to the following reductions:

$$a_k a_j a_1^{-1} \to a_k a_1^{-1} U_{1j} \to a_1^{-1} U_{1k} U_{1j}$$

and

$$a_k a_j a_1^{-1} \to a_j S_{jk} a_1^{-1} \xrightarrow{*} a_j a_1^{-1} L \to a_1^{-1} U_{1j} L.$$

Now

$$U_{1k} U_{1j} a_1 \xrightarrow{*} U_{1k} a_1 a_j \xrightarrow{*} a_1 a_k a_j \to a_1 a_j S_{jk}$$

and

$$U_{1j} L a_1 \xrightarrow{*} U_{1j} a_1 S_{jk} \xrightarrow{*} a_1 a_j S_{jk}.$$

Hence

$$\overline{U_{1j} L} = a_j S_{jk} = \overline{U_{1k} U_{1j}},$$

so $U_{1j}L$ and $U_{1k}U_{1j}$ represent the same element of K. This means that local confluence holds at P.

Corollary 8.5. *If $(Y, \mathcal{R})$ is not consistent, then the first word with respect to $\prec$ at which confluence fails is one of the overlaps $(*)$.*

If $I = \{1, \ldots, n\}$, Proposition 8.3 says nothing new. However, if $I = \emptyset$, then for large n Proposition 8.3 reduces the amount of work needed to check consistency by a factor of 8. The number of overlaps between left sides of commutation relations is cubic in n, while the number of all other overlaps is quadratic in n. Without Proposition 8.3, we would have to check local consistency at all overlaps $a_k^\epsilon a_j^\nu a_i^\chi$, where $k > j > i$ and ϵ, η, and χ range independently over $\{1, -1\}$. With Proposition 8.3, we need only consider the case $\epsilon = \eta = \chi = 1$.

It is possible to convert the consistency test of Proposition 8.3 into a procedure for producing a consistent presentation for the group given by a presentation $(Y, \mathcal{R})$ which is not consistent. The following examples illustrate the ideas. The details are left as an exercise to the reader.

Example 8.1. Let G be the group defined by the following power-conjugate presentation on generators x, y, z, in that order:

$$zy = yz, \quad yx = xz, \quad zx = xy, \quad z^{15} = 1, \quad y^9 = 1, \quad x^2 = 1.$$

The relations involving only y and z form a consistent presentation for $\mathbb{Z}_9 \times \mathbb{Z}_{15}$. Conjugation by x takes y to z and z to y. In order for this map to define an automorphism of $\mathrm{Grp}\langle y, z\rangle$, y^{15} and z^9 must be trivial. This implies the power relations $y^3 = 1$ and $z^3 = 1$. With these new power relations the presentation is confluent.

Example 8.2. Let H be the group defined by the nilpotent presentation on generators a, b, c, d, e, f, g, h, in that order, in which the nontrivial commutation relations are

$$ba = abc, \quad ca = acd, \quad cb = bce,$$

and the power relations are

$$a^2 = fh, \quad b^2 = gh, \quad c^2 = h^2.$$

The relations on c, d, e, f, g, and h form a consistent presentation of an abelian group. Conjugation by b maps c to ce and fixes d, e, f, g, and h. In order for this map to extend to an automorphism σ, the relation $c^2 = h$ must be preserved. That is, $(ce)^2$ must equal h, so $e^2 = 1$. Let us add this relation to the presentation. Since b^2 is in Grp $\langle c, d, e, f, g, h\rangle$, which is abelian, σ^2 must be the identity. Now $\sigma^2(c) = \sigma(ce) = ce^2 = c$, so σ^2 is 1. The condition that σ must preserve $b^2 = gh$ is also satisfied. Thus with the addition of $e^2 = 1$ we have a consistent presentation on b, c, d, e, f, g, and h. Now let τ be the automorphism induced on Grp $\langle b, c, d, e, f, g, h\rangle$ by a. The condition that τ preserves the relations $c^2 = h^2$ and $b^2 = gh$ leads to the relations $d^2 = 1$ and $e = h^2$, respectively. The existing relation $e^2 = 1$ now implies that $h^4 = 1$. The condition that τ^2 is the identity and hence fixes a produces $d = h^2$. The original presentation together with the relations of $d = h^2$, $e = h^2$, and $h^4 = 1$ is consistent.

Exercises

8.1. Use the ideas of this section to determine a consistent polycyclic presentation for the group generated by a, b, c subject to the relations

$$ba = ab^{-1}, \quad ba^{-1} = a^{-1}b^{-1}, \quad ca = abc, \quad ca^{-1} = a^{-1}c^{-1}, \quad cb = bc^{-1}, \quad cb^{-1} = b^{-1}c^{-1}.$$

8.2. Complete the case analysis in the proof of Proposition 8.3.

9.9 Consistency, the nilpotent case

This section continues the discussion of Section 9.8. The goal remains to reduce as much as possible the amount of work needed to check the consistency of a presentation $(Y, \mathcal{R})$ which has the form of a standard monoid polycyclic presentation. The notation established in the previous section is still in effect.

The consistency criterion of Proposition 8.3 can be strengthened if $\mathcal{R}$ is a γ-weighted presentation as defined in Section 9.4. In this case the generators $a_1, \dots, a_n$ have positive integer weights $1 = w_1 \le \cdots \le w_n$. (We assign weight w_i to a_1^{-1} if i is not in I.) All generators in the word W_i have weight at least $w_i + 1$ and S_{ij} has the form $a_j A_{ij}$, where every generator in A_{ij} has weight at least $w_i + w_j$. For each index k with $w_k > 1$ there are indices i and j with $i < j$, $w_i = 1$, $w_j = w_k - 1$, and $A_{ij} = a_k$. We choose one such pair (i, j) and call the relation $a_j a_i = a_i a_j a_k$ the *definition* of a_k. For each $e \ge 1$, the subgroup $G(e)$ generated by the a_k with $w_k \ge e$ is normal in G. Thus a_i and a_j commute modulo $G(w_i + w_j)$. Therefore, if $(Y, \mathcal{R})$ is consistent, then

$$\begin{aligned} S_{ij} &= a_j A_{ij}, \\ T_{ij} &= a_j^{-1} B_{ij}, \\ U_{ij} &= a_j C_{ij}, \\ V_{ij} &= a_j^{-1} D_{ij}, \end{aligned}$$

where generators occurring in A_{ij}, B_{ij}, C_{ij}, or D_{ij} have weight at least $w_i + w_j$. We shall make this additional assumption.

Let $c = w_n$ and consider the following set of overlaps:

$$\begin{array}{lll}
a_k a_j a_i, & k > j > i, \; w_i + w_j + w_k \le c, & \\
a_j^{m_j} a_i, & j \in I, \quad j > i, \; w_i + w_j < c, & \\
a_j a_i^{m_i}, & i \in I, \quad j > i, \; w_i + w_j < c, & \\
a_i^{m_i+1}, & i \in I, \qquad\qquad 2w_i < c, & (**) \\
a_j a_i^{-1} a_i, & i \notin I, \quad j > i, \; w_i + w_j \le c, & \\
a_j^{-1} a_j a_i, & i \notin I, \quad j > i, \; w_i + w_j \le c, & \\
a_j^{-1} a_j a_i^{-1}, & i, j \notin I, \; j > i, \; w_i + w_j \le c. &
\end{array}$$

Proposition 9.1. *Suppose that $(Y, \mathcal{R})$ is γ-weighted and satisfies the additional assumption. If $(Y, \mathcal{R})$ is not confluent, then the first word P at which confluence fails is one of the overlaps $(**)$.*

Proof. By Corollary 8.5, P is one of the overlaps $(*)$. Suppose first that $P = a_k a_j a_i$ with $k > j > i$ and $w_i + w_j + w_k > c$. The initial reductions at P are

$$a_k a_j a_i \to a_k a_i a_j A_{ij} \to a_i a_k A_{ik} a_j A_{ij}$$

and

$$a_k a_j a_i \to a_j a_k A_{jk} a_i.$$

Now any a_r^α, $\alpha = \pm 1$, occurring in A_{jk} satisfies $r > i$ and $w_r \ge w_j + w_k$. Therefore $w_i + w_r > c$ and $a_r^\alpha a_i \to a_i a_r^\alpha$ is in $\mathcal{R}$. Thus

$$a_j a_k A_{jk} a_i \xrightarrow{*} a_j a_k a_i A_{jk} \to a_j a_i a_k A_{ik} A_{jk} \to a_i a_j A_{ij} a_k A_{ik} A_{jk}.$$

By the minimality of P, $(Y_{i+1}, \mathcal{R}_{i+1})$ is consistent and defines a group K. Any monoid generator a_r^α occurring in A_{ij} has weight at least $w_i + w_j$. Hence a_k and A_{ij} define commuting elements of K. By the same argument, a_j and A_{ik} define commuting elements, as do A_{ij}, A_{ik}, and A_{jk}. Therefore, if u is the element of K defined by $a_k A_{ik} a_j A_{ij}$, then u is also defined by the following words:

$$a_k a_j A_{ik} A_{ij}, \quad a_j a_k A_{jk} A_{ik} A_{ij}, \quad a_j A_{ij} a_k A_{ik} A_{jk}.$$

Therefore, if R is the collected word in Y_{i+1}^* defining u, then $a_i R$ is derivable from both $a_i a_k A_{ik} a_j A_{ij}$ and $a_i a_j A_{ij} a_k A_{ik} A_{jk}$. Therefore local confluence holds at P.

Now suppose that $P = a_j^{m_j} a_i$ with $j > i$ and $w_i + w_j > c$. Since A_{ij} is empty, the initial reductions at P are

$$a_j^{m_j} a_i \to W_j a_i$$

and

$$a_j^{m_j} a_i \to a_j^{m_j - 1} a_i a_j \xrightarrow{*} a_i a_j^{m_j} \to a_i W_j.$$

Every generator in W_j has weight at least $w_j + 1$. Therefore $W_j a_i \xrightarrow{*} a_i W_j$, so local confluence holds at P.

If $P = a_j a_i^{m_i}$ with $j > i$ and $w_i + w_j > c$, then the initial reductions at P are

$$a_j a_i^{m_i} \to a_j W_i$$

and

$$a_j a_i^{m_i} \to a_i a_j a_i^{m_i - 1} \xrightarrow{*} a_i^{m_i} a_j \to W_i a_j.$$

Each generator occurring in W_i commutes with a_j, so $a_j W_i$ and $W_i a_j$ define the same element of the group $K = \mathrm{Mon}\langle Y_{i+1} \mid \mathcal{R}_{i+1} \rangle$. Thus confluence holds at P.

Suppose $P = a_i^{m_i + 1}$ with $2w_i \geq c$. The reductions at P are

$$a_i^{m_i + 1} \to a_i W_i$$

and

$$a_i^{m_i + 1} \to W_i a_i \xrightarrow{*} a_i W_i,$$

since each a_r^{α}, $\alpha = \pm 1$, which occurs in W_i has weight $w_r \geq w_i + 1$. Therefore $w_r + w_i > c$ and $a_r^{\alpha} a_i \to a_i a_r^{\alpha}$ is in $\mathcal{R}$.

Finally, if $j > i$ and $w_i + w_j > c$, then it is easy to check that local confluence holds at $a_j a_i^{-1} a_i$, $a_i^{-1} a_j a_i$, and $a_i^{-1} a_j a_i^{-1}$. □

One can improve Proposition 9.1 using ideas from (Vaughan-Lee 1984). For any i and j with $1 \leq i < j \leq n$ the set $\mathcal{W} = Y_j^* a_i Y_j^*$ is closed under rewriting. Suppose that confluence holds on $\mathcal{W}$. If P is in Y_j^*, then $P \xrightarrow{*} Q$ if and only if $a_i P \xrightarrow{*} a_i Q$. Therefore confluence holds on Y_j^* and $(Y_j, \mathcal{R}_j)$ is a consistent presentation for a group K_j. Let u be an element of K_j, let P in Y_j^* represent u, and let $P a_i \xrightarrow{*} a_i Q$. (Such a Q always exists.) An argument in the proof of Proposition 8.3 shows that the element v of K_j defined by Q depends only on u and that the map $u \mapsto v$ is a homomorphism

σ_i of K_j into itself. Since $a_k a_i \to a_i a_k A_{ik}$ for $j \le k \le n$, it is easy to see that σ_i is surjective. Therefore σ_i is an automorphism of K_j. (In the proof of Proposition 8.3, we could be sure that $\mathcal{W}$ was closed under rewriting only when $j = i+1$ and we needed local confluence at the overlaps $a_k a_i^{-1} a_i$ or $a_k a_i^{m_i}$ to know that σ_i was surjective.)

Now suppose that $1 < j \le n$ and let $L = b_1 \dots b_r$ be a word in $\{a_1, \dots, a_{j-1}\}^*$ which is irreducible with respect to $\mathcal{R}$. Set $\mathcal{W}(j,L) = Y_j^* b_1 Y_j^* \dots Y_j^* b_r Y_j^*$.

Proposition 9.2. *Assume that j and L are as described and that confluence holds on $\mathcal{W}(j,b_i) = Y_j^* b_i Y_j^*$, $1 \le i \le r$. Then confluence holds on $\mathcal{W}(j,L)$. If P in Y_j^* represents an element u of K_j and $PL \xrightarrow{*} LQ$, then Q represents u^τ, where $\tau = \tau_1 \cdots \tau_r$ and τ_i is the automorphism of K_j induced by b_i.*

Proof. If $r \le 1$, then there is nothing to prove, so we may assume that $r \ge 2$ and that confluence does not hold on $\mathcal{W}(j,L)$. Let $\mathcal{V}$ be the union over all subwords M of L of the sets $\mathcal{W}(j,M)$. The set $\mathcal{V}$ is closed under rewriting and under taking subwords. Let R be the first word in $\mathcal{V}$ with respect to $\prec$ at which confluence fails. By Exercise 7.2 in Chapter 2, R is an overlap of precisely two left sides. By assumption, R must involve at least two consecutive factors from L. Since L is irreducible, R must contain elements of Y_j as well. All elements of Y_j come before the factors of L in the ordering $\prec$. A simple case analysis shows that no such R can exist.

Now let P be any word in Y_j^* and let P represent u in K_j. We can rewrite PL as follows:

$$\begin{aligned} PL = Pb_1 \dots b_r &\xrightarrow{*} b_1 P_1 b_2 \dots b_r \xrightarrow{*} b_1 b_2 P_2 b_3 \dots b_r \xrightarrow{*} \cdots \xrightarrow{*} b_1 \dots b_r P_r \\ &= LP_r, \end{aligned}$$

where each P_i is in K_j and represents the image u_i of u under $\tau_1 \cdots \tau_i$. We may assume that LP_r is irreducible. Now let Q be any word in K_j such that $PL \xrightarrow{*} LQ$. Since we have confluence on $\mathcal{W}(j,L)$, it follows that $LQ \xrightarrow{*} LP_r$, which means that $Q \xrightarrow{*} P_r$. Therefore Q represents u_r too. □

Proposition 9.3. *Assume that $(Y, \mathcal{R})$ is as in Proposition 9.1. In testing local confluence at the overlaps $(**)$, we may assume that $w_i = 1$ when considering $a_k a_j a_i$ with $k > j > i$ and $a_j^{m_j} a_i$ with $j > i$.*

Proof. Suppose that local confluence holds at all the overlaps $(**)$, except perhaps at some of the words $a_k a_j a_i$ or $a_j^{m_j} a_i$ with $w_i > 1$. The presentation $(Y_n, \mathcal{R}_n)$ is consistent and confluence holds on $\mathcal{W}(n, a_k)$ for all $k < n$. The automorphisms of K_n induced by $a_1, \dots, a_{n-1}$ are all trivial. Suppose r is such that $(Y_r, \mathcal{R}_r)$ is consistent and confluence holds on $\mathcal{W}(r, a_k)$ for all k

with $1 \leq k < r$. Each of the generators a_k induces an automorphism σ_k on K_r. Choose r minimal such that if $k < r$ and $w_k > 1$, then $\sigma_j\sigma_i = \sigma_i\sigma_j\sigma_k$, where $a_ja_i = a_ia_ja_k$ is the definition of a_k. As noted, $r \leq n$.

If $r = 1$, then $(Y, \mathcal{R})$ is consistent. Suppose $r > 1$.

Lemma 9.4. *The presentation $(Y_{r-1}, \mathcal{R}_{r-1})$ is consistent.*

Proof. By assumption, a_{r-1} induces an automorphism σ_{r-1} on K_r. If $r-1$ is not in I, then σ_{r-1} defines a cyclic extension H of K_r with H/K_r infinite. If $r-1$ is in I, then local confluence at the words $a_s a_{r-1}^{m_{r-1}}$ with $s \geq r-1$ shows that σ_{r-1} and W_{r-1} define a cyclic extension H of K_r with H/K_r of order m_{r-1}. In either case, local confluence at the last three types of overlaps in $(**)$ with $i = r-1$ implies that H is a homomorphic image of $\mathrm{Mon}\langle Y_{r-1}, \mathcal{R}_{r-1}\rangle$. Therefore $(Y_{r-1}, \mathcal{R}_{r-1})$ is consistent. $\square$

Suppose that confluence holds on all $\mathcal{W}(r-1, a_k)$ with $k < r-1$. Then each a_k defines an automorphism σ_k of K_{r-1}. Let $a_ja_i = a_ia_ja_k$ be the definition of some a_k with $k < r-1$ and $w_k > 1$. The automorphisms $\sigma_j\sigma_i$ and $\sigma_i\sigma_j\sigma_k$ agree on K_r. To see if they agree on K_{r-1}, it suffices to check whether they agree on the element u represented by a_{r-1}. Since $w_i = 1$, we know that local confluence holds at $a_{r-1}a_ja_i$. The reductions confirming local confluence begin as follows:

$$a_{r-1}a_ja_i \to a_ja_{r-1}A_{jr-1}a_i,$$
$$a_{r-1}a_ja_i \to a_{r-1}a_ia_ja_k.$$

Any word derivable from both $a_ja_{r-1}A_{jr-1}a_i$ and $a_{r-1}a_ia_ja_k$ must have a_i occurring earlier in the word than a_j. Thus the first reduction must continue with

$$a_ja_{r-1}A_{jr-1}a_i \xrightarrow{*} a_ja_iM \to a_ia_ja_kM.$$

This means that the second reduction must involve

$$a_{r-1}a_ia_ja_k \xrightarrow{*} a_ia_ja_kN.$$

Local confluence means that M and N define the same element of K_{r-1}. But M represents the image of u under $\sigma_j\sigma_i$ and, by Proposition 9.2, N represents the image of u under $\sigma_i\sigma_j\sigma_k$. This means that $\sigma_j\sigma_i$ and $\sigma_i\sigma_j\sigma_k$ agree on u and therefore they agree on K_{r-1}.

By the choice of r, we conclude that confluence must fail on some $\mathcal{W}(r-1, a_k)$ with $k < r-1$. Choose k minimal. By Exercise 7.2 in Chapter 2, the first word R in $\mathcal{W}(r-1, a_k)$ at which confluence fails is the overlap of precisely two left sides, local confluence fails at R, and R

involves a_k. By arguments similar to those in Proposition 8.3, R has the form $a_t a_s a_k$ with $t > s > k$ or $a_s^{m_s} a_k$. Since confluence holds on $\mathcal{W}(r, a_k)$, in either case $s = r-1$. If $w_k = 1$, then local confluence is known to hold at R. Therefore $w_k > 1$. Let $a_j a_i = a_i a_j a_k$ be the definition of a_k. Then a_i and a_j define automorphisms σ_i and σ_j, respectively, of K_{r-1} and a_k defines an automorphism σ_k of K_r. On K_r, we have $\sigma_j \sigma_i = \sigma_i \sigma_j \sigma_k$.

As before, local confluence holds at $a_{r-1} a_j a_i$ and one of the reductions confirming this starts out

$$a_{r-1} a_j a_i \to a_j a_{r-1} A_{jr-1} a_i \xrightarrow{*} a_j a_i M \to a_i a_j a_k M.$$

We must analyze the other reduction a little more closely. It begins

$$a_{r-1} a_j a_i \to a_{r-1} a_i a_j a_k \to a_i a_{r-1} A_{ir-1} a_j a_k \xrightarrow{*} a_i a_{r-1} a_j P a_k Q,$$

where it is possible to write A_{ir-1} as EF and $F a_j a_k \xrightarrow{*} a_j a_k Q$ and $E a_j \xrightarrow{*} a_j P$. The reduction continues

$$a_i a_{r-1} a_j P a_k Q \to a_i a_j a_{r-1} A_{jr-1} P a_k Q \xrightarrow{*} a_i a_j a_{r-1} a_k L \to a_i a_j a_k a_{r-1} A_{kr-1} L.$$

Let u and v be the elements of K_{r-1} represented by a_{r-1} and $a_{r-1} A_{kr-1}$, respectively. Then A_{ir-1} and A_{jr-1} represent $u^{-1} u^{\sigma_i}$ and $u^{-1} u^{\sigma_j}$, respectively. The word L represents

$$(u^{-1} u^{\sigma_j})^{\sigma_k} (u^{-1} u^{\sigma_i})^{\sigma_j \sigma_k} = [u^{-1} u^{\sigma_j} (u^{-1} u^{\sigma_i})^{\sigma_j}]^{\sigma_k} = [u^{-1} u^{\sigma_i \sigma_j}]^{\sigma_k}.$$

The element $u^{-1} u^{\sigma_i \sigma_j}$ is in K_r and on K_r we know that σ_k is the same as $\sigma_j^{-1} \sigma_i^{-1} \sigma_j \sigma_i$. Therefore L represents

$$(u^{\sigma_j^{-1} \sigma_i^{-1} \sigma_j \sigma_i})^{-1} u^{\sigma_j \sigma_i}.$$

As before, M represents $u^{\sigma_j \sigma_i}$. Local confluence at $a_{r-1} a_j a_i$ implies that

$$v (u^{\sigma_j^{-1} \sigma_i^{-1} \sigma_j \sigma_i})^{-1} u^{\sigma_j \sigma_i} = u^{\sigma_j \sigma_i}$$

or

$$v = u^{\sigma_j^{-1} \sigma_i^{-1} \sigma_j \sigma_i}.$$

This means that all relations in $\mathcal{R}$ which can be used to rewrite elements of $\mathcal{W}(r-1, a_k)$ are satisfied in the cyclic extension of K_{r-1} defined by $\sigma_j^{-1} \sigma_j^{-1} \sigma_j \sigma_i$. Therefore confluence must hold on $\mathcal{W}(r-1, a_k)$ after all. This contradiction completes the proof of Proposition 9.3. □

9.10 Free nilpotent groups

Let F be the free group on a finite set X with $|X| = r$ and let e be a positive integer. Any group isomorphic to $G = F/\gamma_{e+1}(F)$ is called a *free nilpotent group* of rank r and class e. The term "collection" was first used in (P. Hall 1934) in connection with computing in free nilpotent groups. In this section we shall describe polycyclic generating sequences for G. Consistent polycyclic presentations are given for G in terms of some of these sequences, but the proof of consistency requires considerably more space than is available here. Details can be found in [Hall 1959] and [Magnus et al. 1976].

Since G is finitely generated and nilpotent, it is polycyclic by Proposition 3.4. Suppose $X = \{x_1, \ldots, x_r\}$. By Propositions 2.5 and 2.6, for $k \geq 2$ the commutators $[x_{i_1}, \ldots, x_{i_k}]$ with $i_1 > i_2$ generate $\gamma_k(F)$ modulo $\gamma_{k+1}(F)$. Thus the images in G of these elements with $k \leq e$ form a polycyclic generating sequence for G when arranged in increasing order of k. A presentation for G is obtained by setting all commutators $[x_{i_1}, \ldots, x_{i_{k+1}}]$ with $i_1 > i_2$ equal to 1. The Knuth-Bendix procedure can be used to find the corresponding standard polycyclic presentation.

Example 10.1. Let $X = \{x_1, x_2\}$ and take $e = 5$. One polycyclic generating sequence for G is $x_1, \ldots, x_{17}$, where $x_3, \ldots, x_{17}$ are defined as follows:

$$\begin{aligned}
x_3 &= [x_2, x_1],\\
x_4 &= [x_3, x_1] = [x_2, x_1, x_1],\\
x_5 &= [x_3, x_2] = [x_2, x_1, x_2],\\
x_6 &= [x_4, x_1] = [x_2, x_1, x_1, x_1],\\
x_7 &= [x_4, x_2] = [x_2, x_1, x_1, x_2],\\
x_8 &= [x_5, x_1] = [x_2, x_1, x_2, x_1],\\
x_9 &= [x_5, x_2] = [x_2, x_1, x_2, x_2],\\
x_{10} &= [x_6, x_1] = [x_2, x_1, x_1, x_1, x_1],\\
x_{11} &= [x_6, x_2] = [x_2, x_1, x_1, x_1, x_2],\\
x_{12} &= [x_7, x_1] = [x_2, x_1, x_1, x_2, x_1],\\
x_{13} &= [x_7, x_2] = [x_2, x_1, x_1, x_2, x_2],\\
x_{14} &= [x_8, x_1] = [x_2, x_1, x_2, x_1, x_1],\\
x_{15} &= [x_8, x_2] = [x_2, x_1, x_2, x_1, x_2],\\
x_{16} &= [x_9, x_1] = [x_2, x_1, x_2, x_2, x_1],\\
x_{17} &= [x_9, x_2] = [x_2, x_1, x_2, x_2, x_2].
\end{aligned}$$

A presentation for G on these generators is $[x_j, x_i] = 1$, $10 \le j \le 17$, $1 \le i \le 2$. The Knuth-Bendix procedure using the basic wreath-product ordering with $x_{17} \prec x_{17}^{-1} \prec x_{16} \prec \cdots \prec x_1 \prec x_1^{-1}$ will produce the standard polycyclic presentation for G on the x_i. The Knuth-Bendix procedure can be helped along by including some redundant relations $[x_j, x_i] = 1$, where $[x_j, x_i]$ is clearly in $\gamma_6(G) = 1$. Examples of such relations are $[x_6, x_3] = 1$ and $[x_j, x_i] = 1$ for $4 \le i < j \le 17$.

The standard presentation obtained in this case is too large to list here, but it is instructive to look at the power relations which occur. They are

$$\begin{aligned} x_7 &= x_8 x_{11} x_{14}^{-1} x_{15} x_{16}^{-1}, \\ x_{12} &= x_{14}, \\ x_{13} &= x_{15}. \end{aligned}$$

Thus $\gamma_4(F)/\gamma_5(F) \cong \gamma_4(G)/\gamma_5(G)$ is the abelian group generated by x_6, x_7, x_8, and x_9 subject to the single relation $x_7 = x_8$. Therefore $\gamma_4(G)/\gamma_5(G)$ is free abelian of rank 3. Similarly, $\gamma_5(F)/\gamma_6(F) \cong \gamma_5(G)$ is free abelian of rank 6.

Computations like those in Example 10.1 suggest that the quotients $\gamma_k(F)/\gamma_{k+1}(F)$ are always free abelian groups and that for large k the ranks of these groups are substantially smaller than the upper bound of

$$\frac{r^k - r^{k-1}}{2}$$

given by Propositions 2.5 and 2.6. The rank of $\gamma_k(F)/\gamma_{k+1}(F)$ is known and bases for these groups have been determined. To describe these bases we need to introduce the concept of a basic sequence of commutators.

A *basic sequence of commutators* in F is an infinite sequence $c_1, c_2, \ldots$ of elements of F, where each c_i has associated with it a positive integer w_i called its *weight*. The c_i must satisfy several conditions, which will now be described. The c_i are ordered by weight. That is, if $j > i$, then $w_j \ge w_i$. The commutators of weight 1 are $c_1, \ldots, c_r$, which are the elements of X arranged in some order. If $w_k > 1$, then c_k is described explicitly as the commutator $[c_j, c_i]$, where $j > i$ and $w_k = w_i + w_j$. If $w_j > 1$, so that c_j is described as $[c_q, c_p]$ with $q > p$, then $p \le i$. Finally, for each $j > i$ such that either $w_j = 1$ or $w_j > 1$ and c_j is described as $[c_q, c_p]$ with $p \le i$, there is a unique index k such that c_k is described as $[c_j, c_i]$. The phrase "sequence of basic commutators" is used by most authors, but being basic is a property of the sequence, not of the individual terms in the sequence. We shall say that a commutator u is basic only when a basic sequence of commutators has previously been specified and u is a term in that sequence.

Example 10.2. Suppose that $X = \{a, b\}$. The terms of weight at most 6 in one basic sequence are

$$\begin{aligned}
c_1 &= a, & c_{13} &= [c_7, c_2], \\
c_2 &= b, & c_{14} &= [c_8, c_2], \\
c_3 &= [c_2, c_1], & c_{15} &= [c_5, c_4], \\
c_4 &= [c_3, c_1], & c_{16} &= [c_6, c_3], \\
c_5 &= [c_3, c_2], & c_{17} &= [c_7, c_3], \\
c_6 &= [c_4, c_1], & c_{18} &= [c_8, c_3], \\
c_7 &= [c_4, c_2], & c_{19} &= [c_{11}, c_1], \\
c_8 &= [c_5, c_2], & c_{20} &= [c_{11}, c_2], \\
c_9 &= [c_4, c_3], & c_{21} &= [c_{12}, c_2], \\
c_{10} &= [c_5, c_3], & c_{22} &= [c_{13}, c_2], \\
c_{11} &= [c_6, c_1], & c_{23} &= [c_{14}, c_2]. \\
c_{12} &= [c_6, c_2], & &
\end{aligned}$$

Here c_1 and c_2 have weight 1, c_3 has weight 2, c_4 and c_5 have weight 3, c_6, c_7, and c_8 have weight 4, $c_9, \ldots, c_{14}$ have weight 5, and $c_{15}, \ldots, c_{23}$ have weight 6.

As defined here, a basic sequence of commutators includes the description of each term of weight greater than 1 as the commutator of earlier terms in the sequence. The reason for this is at this stage it is conceivable that $[c_j, c_i] = [c_q, c_p]$ even though $(j, i) \neq (q, p)$ and that some of the c_k could be trivial. This cannot occur, but the proof is somewhat involved. From now on we shall write $c_k = [c_j, c_i]$ for the assertion that c_k is described as $[c_j, c_i]$.

Let us fix a basic sequence c_1, $c_2, \ldots$ of commutators in F. The commutators of weight 1 are in $F = \gamma_1(F)$. By Proposition 1.10 and a simple induction argument, it is easy to prove that c_i is in $\gamma_{w_i}(F)$ for all i. Suppose that $c_1, \ldots, c_t$ are the commutators in the sequence with weight at most e, and for $1 \leq i \leq t$ let a_i be the image of c_i in $G = F/\gamma_{e+1}(F)$. We shall prove that the a_i form a polycyclic generating sequence for G. In the proof, the following proposition will be needed.

Proposition 10.1. *Let u and v be elements of a group. Then*

$$[v, u^{-1}] = [v, u, u^{-1}]^{-1}[v, u]^{-1}$$

and

$$[v^{-1}, u] = [v, u, v^{-1}]^{-1}[v, u]^{-1}.$$

Set $u_1 = v_1 = [v, u]$ *and for* $i \geq 1$ *let* $u_{i+1} = [u_i, u]$ *and* $v_{i+1} = [v_i, v]$. *Then, for any odd positive integer* s,

$$[v, u^{-1}] = u_2 u_4 \ldots u_{s-1} [u_s, u^{-1}]^{-1} u_s^{-1} u_{s-2}^{-1} \ldots u_3^{-1} u_1^{-1},$$
$$[v^{-1}, u] = v_2 v_4 \ldots v_{s-1} [v_s, v^{-1}]^{-1} v_s^{-1} v_{s-2}^{-1} \ldots v_3^{-1} v_1^{-1}.$$

Proof. The first two identities are proved by direct computation in the free group generated by u and v. They correspond to the case $s = 1$ of the second pair of identities. The second pair is proved by induction on s, using the following applications of the first identity:

$$\begin{aligned}
[u_s, u^{-1}]^{-1} &= ([u_s, u, u^{-1}]^{-1} [u_s, u]^{-1})^{-1} \\
&= u_{s+1} [u_{s+1}, u^{-1}] \\
&= u_{s+1} [u_{s+1}, u, u^{-1}]^{-1} [u_{s+1}, u]^{-1} \\
&= u_{s+1} [u_{s+2}, u^{-1}]^{-1} u_{s+2}^{-1}, \\
[v_s, v^{-1}]^{-1} &= ([v_s, v, v^{-1}]^{-1} [v_s, v]^{-1})^{-1} \\
&= v_{s+1} [v_{s+1}, v^{-1}] \\
&= v_{s+1} [v_{s+1}, v, v^{-1}]^{-1} [v_{s+1}, v]^{-1} \\
&= v_{s+1} [v_{s+2}, v^{-1}]^{-1} v_{s+2}^{-1}. \qquad \square
\end{aligned}$$

Now we are ready to prove that a basic sequence of commutators defines a polycyclic generating sequence for G. For $1 \leq i \leq t+1$ define $G_i = \mathrm{Grp}\langle a_i, \ldots, a_t \rangle$.

Proposition 10.2. *The sequence* $a_1, \ldots, a_t$ *is a polycyclic generating sequence for* G.

Proof. We must show that for $1 \leq i < j \leq t$ and any α and β in $\{1, -1\}$ there is an element z of G_{i+1} such that $a_j^\alpha a_i^\beta = a_i^\beta z$. If $w_i + w_j > e$, then $[c_j, c_i]$ is in $\gamma_{e+1}(F)$ and so a_i and a_j commute. Therefore we may take $z = a_j^\alpha$ in this case. Let us assume that $w_i + w_j \leq e$. Suppose first that $[c_j, c_i] = c_k$ is basic. In Proposition 10.1, let $u = c_i$ and $v = c_j$. Each of the commutators u_m and v_m is basic. If m is large, then both u_m and v_m are in $\gamma_{e+1}(F)$ and can be ignored when we pass to the quotient group G. Let x_m and y_m denote the images of u_m and v_m, respectively, in G. Then x_m and y_m lie in G_{i+1} and

$$a_j a_i = a_i a_j a_k,$$
$$a_j a_i^{-1} = a_i^{-1} a_j x_2 x_4 \ldots x_3^{-1} x_1^{-1},$$
$$a_j^{-1} a_i = a_i a_j^{-1} y_2 y_4 \ldots y_3^{-1} y_1^{-1}.$$

The products indicated with dots are finite products. Finally,

$$a_j^{-1}a_i^{-1} = a_i^{-1}(a_ia_ja_i^{-1})^{-1} = a_i^{-1}x_1x_3\ldots x_4^{-1}x_2^{-1}a_j^{-1}.$$

Now suppose that $[c_j, c_i]$ is not basic. Then $c_j = [c_q, c_p]$ with $j > q > p > i$. In particular, $j > i+1$, so we may proceed by induction on j. The conjugates $a_i^{-1}a_qa_i$ and $a_i^{-1}a_pa_i$ have already been shown to be in G_{i+1}. Therefore $a_i^{-1}a_ja_i = a_i^{-1}[a_q, a_p]a_i$ is also in G_{i+1}, or, equivalently, $a_ja_i = a_iz$, where z is in G_{i+1}. The other cases are handled in the same way. □

The proof of Proposition 10.2 allows us to construct a polycyclic presentation for a group H of which G is a homomorphic image. It is in fact the case that the presentation obtained is consistent and H is isomorphic to G, although we cannot give the proof here. It is also true that the series $G = G_1 \supseteq \cdots \supseteq G_{t+1}$ refines the lower central series of G. Moreover, if $1 \le m \le e$, then the quotient $Q = \gamma_m(G)/\gamma_{m+1}(G) \cong \gamma_m(F)/\gamma_{m+1}(F)$ is a free abelian group. If $c_p, \ldots, c_q$ are the basic commutators of weight m, then the images of $a_p, \ldots, a_q$ in Q form a basis of Q. The rank of Q is

$$\frac{1}{m}\sum_{d|m}\mu(d)r^{m/d},$$

where μ is the Möbius function, which is defined as follows: $\mu(n) = (-1)^s$ if n is the product of s distinct primes and $\mu(n) = 0$ if n is divisible by the square of a prime. In particular, $\mu(1) = 1$.

Example 10.3. Let us again look at the case $r = 2$ and $e = 5$, using the basic sequence of commutators in Example 10.2. The generators of G are $a_1, \ldots, a_{14}$, which are the images of $c_1, \ldots, c_{14}$. There are 91 commutators $[a_j, a_i]$ with $i < j$. For 76 of these, $w_i + w_j \ge 6$, and hence the commutator is trivial. Of the 15 remaining commutators, 12 are basic. This leaves only three to be determined. They are $[a_5, a_1]$, $[a_7, a_1]$, and $[a_8, a_1]$.

$$\begin{aligned}
a_1^{-1}a_5a_1 &= a_1^{-1}[a_3, a_2]a_1 = [a_1^{-1}a_3a_1, a_1^{-1}a_2a_1] = [a_3a_4, a_2a_3] \\
&= a_4^{-1}a_3^{-1}a_3^{-1}a_2^{-1}a_3a_4a_2a_3 = a_4^{-1}a_3^{-1}a_3^{-1}a_2^{-1}a_3a_2a_4a_7a_3 \\
&= a_4^{-1}a_3^{-1}a_3^{-1}a_2^{-1}a_2a_3a_5a_4a_7a_3 = a_4^{-1}a_3^{-1}a_5a_4a_7a_3 \\
&= a_4^{-1}a_3^{-1}a_5a_4a_3a_7 = a_4^{-1}a_3^{-1}a_5a_3a_4a_9a_7 = a_4^{-1}a_3^{-1}a_3a_5a_{10}a_4a_9a_7 \\
&= a_4^{-1}a_5a_{10}a_4a_9a_7 = a_4^{-1}a_4a_5a_{10}a_9a_7 = a_5a_7a_9a_{10}.
\end{aligned}$$

Therefore $a_5a_1 = a_1a_5a_7a_9a_{10}$. Now

$$a_1^{-1}a_7a_1 = [a_4a_6, a_2a_3] = a_6^{-1}a_4^{-1}a_3^{-1}a_2^{-1}a_4a_6a_2a_3,$$

which simplifies to $a_7a_9a_{12}$. Finally,

$$a_1^{-1}a_8a_1 = [a_5a_7a_9a_{10}, a_2a_3],$$

which turns out to be $a_8a_{10}a_{13}$.

Once we have a consistent polycyclic presentation for our free nilpotent group G, we can compute in G using the techniques developed in Sections 9.4 to 9.7. However, if many calculations with elements of G are to be carried out, then collection is not the most efficient way to compute products. The approach to be described applies not only to free nilpotent groups but to any group defined by a consistent nilpotent presentation without power relations.

Let $a_1, \dots, a_n$ be a polycyclic generating sequence for a group H such that the corresponding standard polycyclic presentation for H is nilpotent and contains no power relations. We can identify H and $\mathbb{Z}^n$ as sets by letting $(\alpha_1, \dots, \alpha_n)$ in $\mathbb{Z}^n$ correspond to $a_1^{\alpha_1} \dots a_n^{\alpha_n}$. Suppose that g and h are elements of H and that g, h, and gh correspond to $(\alpha_1, \dots, \alpha_n)$, $(\beta_1, \dots, \beta_n)$, and $(\sigma_1, \dots, \sigma_n)$, respectively. We may consider $\sigma_1, \dots, \sigma_n$ to be functions of $\alpha_1, \dots, \alpha_n$ and $\beta_1, \dots, \beta_n$. In fact, $\sigma_i = \sigma_i(\alpha_1, \dots, \alpha_i, \beta_1, \dots, \beta_i)$ depends only on $\alpha_1, \dots, \alpha_i$ and $\beta_1, \dots, \beta_i$. In (P. Hall 1957) it is shown that each σ_i is a polynomial function of the α's and the β's. If these polynomials can be determined, they can be used to compute products much more rapidly than can be accomplished with collection. The polynomials defining multiplication in H take on integer values whenever the α's and β's are integers. Any such polynomial is an integer linear combination of products of binomial coefficients of the form

$$\binom{\alpha_1}{s_1} \cdots \binom{\alpha_n}{s_n} \binom{\beta_1}{t_1} \cdots \binom{\beta_n}{t_n}.$$

Notice that these polynomials do not necessarily have integer coefficients when expressed as linear combinations of ordinary monomials. Not only is multiplication in H defined by polynomials, so is inversion. That is, if $a_1^{\delta_1} \dots a_n^{\delta_n}$ is the inverse of $a_1^{\alpha_1} \dots a_n^{\alpha_n}$, then the δ's are polynomials in the α's.

Example 10.4. Suppose that H is $D_3^{(1)}$ as defined in Example 2.1 and

$$a_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then $a_1^{\alpha_1}a_2^{\alpha_2}a_3^{\alpha_3}$ is

$$\begin{bmatrix} 1 & \alpha_2 & \alpha_3 \\ 0 & 1 & \alpha_1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Moreover,

$$\begin{bmatrix} 1 & \alpha_2 & \alpha_3 \\ 0 & 1 & \alpha_1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \beta_2 & \beta_3 \\ 0 & 1 & \beta_1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \sigma_2 & \sigma_3 \\ 0 & 1 & \sigma_1 \\ 0 & 0 & 1 \end{bmatrix},$$

where $\sigma_1 = \alpha_1 + \beta_1$, $\sigma_2 = \alpha_2 + \beta_2$, and $\sigma_3 = \alpha_3 + \beta_3 + \alpha_2\beta_1$.

A crude upper bound for the total degree of σ_i is i. However, if H is our free nilpotent group G of class e and $a_1, \ldots, a_n$ are defined by a basic sequence of commutators $c_1, c_2, \ldots$ with associated weights $w_1, w_2, \ldots$, then each of the products

$$\binom{\alpha_1}{s_1} \cdots \binom{\alpha_i}{s_i} \binom{\beta_1}{t_1} \cdots \binom{\beta_i}{t_i}$$

which occurs with nonzero coefficient in σ_i satisfies

$$\sum_{j=1}^{i} (s_j + t_j) w_j \le w_i.$$

If the class e of G is not very large, then the polynomials σ_i can be computed fairly easily. For the moment, let us assume that we know the bivariate polynomials η_{ijk}, $n \ge k > j > i$, such that

$$a_j^\alpha a_i^\beta = a_i^\beta a_j^\alpha a_{j+1}^{\eta_{ijj+1}(\alpha,\beta)} \ldots a_n^{\eta_{ijn}(\alpha,\beta)}.$$

Notice that this formula remains valid even if the exponents α and β are themselves polynomials in one or more indeterminates. Given the η_{ijk}, we can use any of the classical computer algebra systems to construct a "symbolic collector", which collects products of powers of the generators in which the exponents are polynomials. To compute the σ_i, one has only to collect the single "symbolic word" $a_1^{\alpha_1} \ldots a_n^{\alpha_n} a_1^{\beta_1} \ldots a_n^{\beta_n}$, where the α's and the β's are indeterminates.

Thus it suffices to determine the polynomials η_{ijk}. By the remark above, if $\alpha^r\beta^s$ occurs with nonzero coefficient in η_{ijk}, then $rw_j + sw_i \le w_k \le e$. Thus we may compute $a_j^\alpha a_i^\beta$ for those nonnegative values of α and β for which $\alpha w_j + \beta w_i \le e$ and then use interpolation to find the η_{ijk}. Exercise 10.4 describes one convenient way to carry out the interpolation. Although a basic collection procedure would be adequate in principle to compute the necessary values of $a_j^\alpha a_i^\beta$, we can accomplish the task much more quickly by using the symbolic collector whenever possible. The values of i are considered in the order $n-1, n-2, \ldots, 1$. For a given i, the values of j are considered in the order $n, n-1, \ldots, i+1$. For a given pair (j,i), we take $\alpha = 0, 1, \ldots$ and for a given α we let $\beta = 0, 1, \ldots$. The first collection step is $a_j^\alpha a_i^\beta = a_j^{\alpha-1} a_i a_j [a_j, a_i] a_i^{\beta-1}$. After this, if we need to replace $a_q^\mu a_p^\nu$ by $a_p^\nu a_q^\mu [a_q^\mu, a_p^\nu]$, then either we have already determined the polynomials η_{pqk} or we have $(q,p) = (j,i)$ and $a_j^\mu a_i^\nu$ has already been computed.

Example 10.5. Let us continue with the case $r=2$ and $e=5$ of Example 10.3. Here are the polynomials σ_i:

$$\sigma_1 = \alpha_1 + \beta_1,$$
$$\sigma_2 = \alpha_2 + \beta_2,$$
$$\sigma_3 = \alpha_3 + \beta_3 + \alpha_2\beta_1,$$
$$\sigma_4 = \alpha_4 + \beta_4 + \alpha_3\beta_1 + \alpha_2\binom{\beta_1}{2},$$
$$\sigma_5 = \alpha_5 + \beta_5 + \alpha_3\beta_2 + \binom{\alpha_2}{2}\beta_1 + \alpha_2\beta_1\beta_2,$$
$$\sigma_6 = \alpha_6 + \beta_6 + \alpha_4\beta_1 + \alpha_3\binom{\beta_1}{2} + \alpha_2\binom{\beta_1}{3},$$
$$\sigma_7 = \alpha_7 + \beta_7 + \alpha_5\beta_1 + \alpha_4\beta_2 + \alpha_3\beta_1\beta_2 + \binom{\alpha_2}{2}\binom{\beta_1}{2} + \alpha_2\binom{\beta_1}{2}\beta_2,$$
$$\sigma_8 = \alpha_8 + \beta_8 + \alpha_5\beta_2 + \alpha_3\binom{\beta_2}{2} + \binom{\alpha_2}{3}\beta_1 + \binom{\alpha_2}{2}\beta_1\beta_2 + \alpha_2\beta_1\binom{\beta_2}{2},$$
$$\sigma_9 = \alpha_9 + \beta_9 + \alpha_4\beta_3 + \alpha_5\beta_1 + \alpha_7\beta_1 + \binom{\alpha_3}{2}\beta_1 + \alpha_5\binom{\beta_1}{2} + 2\binom{\alpha_2}{2}\binom{\beta_1}{2}$$
$$+ \alpha_2\alpha_3\binom{\beta_1}{2} + \alpha_2\binom{\beta_1}{3} + \alpha_2\binom{\beta_1}{2}\beta_3 + \alpha_3\beta_1\beta_3 + 3\binom{\alpha_2}{2}\binom{\beta_1}{3},$$
$$\sigma_{10} = \alpha_{10} + \beta_{10} + \alpha_5\beta_1 + \alpha_5\beta_3 + \alpha_8\beta_1 + \binom{\alpha_2}{2}\beta_1 + \binom{\alpha_3}{2}\beta_2 + \alpha_3\beta_2\beta_3$$
$$+ 2\binom{\alpha_2}{3}\beta_1 + \binom{\alpha_2}{2}\alpha_3\beta_1 + 3\binom{\alpha_2}{2}\binom{\beta_1}{2} + \binom{\alpha_2}{2}\beta_1\beta_3 + \alpha_2\alpha_3\beta_1\beta_2$$
$$+ \alpha_2\binom{\beta_1}{2}\beta_2 + \alpha_2\beta_1\beta_2\beta_3 + 4\binom{\alpha_2}{3}\binom{\beta_1}{2} + 2\binom{\alpha_2}{2}\binom{\beta_1}{2}\beta_2$$

$$
\begin{aligned}
&+\binom{\alpha_2}{2}\beta_1\beta_2,\\
\sigma_{11} &= \alpha_{11}+\beta_{11}+\alpha_6\beta_1+\alpha_4\binom{\beta_1}{2}+\alpha_3\binom{\beta_1}{3}+\alpha_2\binom{\beta_1}{4},\\
\sigma_{12} &= \alpha_{12}+\beta_{12}+\alpha_6\beta_2+\alpha_7\beta_1+\alpha_4\beta_1\beta_2+\alpha_5\binom{\beta_1}{2}+\binom{\alpha_2}{2}\binom{\beta_1}{3}\\
&+\alpha_2\binom{\beta_1}{3}\beta_2+\alpha_3\binom{\beta_1}{2}\beta_2,\\
\sigma_{13} &= \alpha_{13}+\beta_{13}+\alpha_7\beta_2+\alpha_8\beta_1+\alpha_4\binom{\beta_2}{2}+\alpha_5\beta_1\beta_2+\binom{\alpha_2}{3}\binom{\beta_1}{2}+\\
&+\binom{\alpha_2}{2}\binom{\beta_1}{2}\beta_2+\alpha_2\binom{\beta_1}{2}\binom{\beta_2}{2}+\alpha_3\beta_1\binom{\beta_2}{2},\\
\sigma_{14} &= \alpha_{14}+\beta_{14}+\alpha_8\beta_2+\alpha_5\binom{\beta_2}{2}+\alpha_3\binom{\beta_2}{3}+\binom{\alpha_2}{4}\beta_1+\binom{\alpha_2}{3}\beta_1\beta_2\\
&+\binom{\alpha_2}{2}\beta_1\binom{\beta_2}{2}+\alpha_2\beta_1\binom{\beta_2}{3}.
\end{aligned}
$$

As e increases, the number of terms in the σ_i grows very rapidly and it becomes difficult to store these polynomials. A compromise is to store only the polynomials

$$\sigma_i^{(j)} = \sigma_i(\alpha_1,\ldots,\alpha_i,0,\ldots,0,\beta_j,0,\ldots,0),$$

where $1 \le j \le i$. The product

$$(a_1^{\alpha_1}\ldots a_n^{\alpha_n})(a_1^{\alpha_1}\ldots a_n^{\alpha_n})$$

is now evaluated as

$$(\cdots((a_1^{\alpha_1}\ldots a_n^{\alpha_n})a_1^{\beta_1})a_2^{\beta_2}\cdots)a_n^{\beta_n}$$

using the polynomials $\sigma_j^{(j)}$, $\sigma_{j+1}^{(j)},\ldots,\sigma_n^{(j)}$ in the computation of the j-th product.

Exercises

10.1. Show that the polycyclic presentation

$$a_2a_1 = a_1a_2^{-1},\quad a_1^{-1}a_1 = a_1a_2,\quad a_2a_1^{-1} = a_1^{-1}a_2^{-1},\quad a_2^{-1}a_1^{-1} = a_1^{-1}a_2$$

on generators a_1 and a_2 is consistent. Determine a formula for $(a_1^{\alpha_1}a_2^{\alpha_2})(a_1^{\beta_1}a_2^{\beta_2})$. Conclude that multiplication in this group is not defined by polynomials.

10.2. Each polynomial f in $\mathbb{Q}[X_1, \ldots, X_m]$ defines a function $\overline{f}$ from $\mathbb{Q}^m$ to $\mathbb{Q}$. Show that the map $f \mapsto \overline{f}$ is an injective ring homomorphism from $\mathbb{Q}[X_1, \ldots, X_m]$ to the ring of functions from $\mathbb{Q}^m$ to $\mathbb{Q}$ under pointwise addition and multiplication.

10.3. If $s_1, \ldots, s_m$ are nonnegative integers, define $g_{s_1 s_2 \ldots s_m}$ to be the product

$$\binom{X_1}{s_1} \cdots \binom{X_m}{s_m}$$

of binomial coefficients. Prove that the polynomials $g_{s_1 s_2 \ldots s_m}$ are a vector space basis for $\mathbb{Q}[X_1, \ldots, X_m]$ over $\mathbb{Q}$. Suppose f is in $\mathbb{Q}[X_1, \ldots, X_m]$ and $\overline{f}(\alpha_1, \ldots, \alpha_m)$ is an integer for all $(\alpha_1, \ldots, \alpha_m)$ in $\mathbb{Z}^m$. Show that f is an integer linear combination of polynomials $g_{s_1 s_2 \ldots s_m}$.

10.4. Suppose that $s_1, \ldots, s_m$ and $t_1, \ldots, t_m$ are sequences of nonnegative integers. Let us say that $g_{t_1 t_2 \ldots t_m}$ precedes $g_{s_1 s_2 \ldots s_m}$ if $t_i \leq s_i$ for all i and $t_i < s_i$ for some i. Let f be in $\mathbb{Q}[X_1, \ldots, X_m]$ and let h be the sum of the terms in f involving polynomials $g_{t_1 t_2 \ldots t_m}$ which precede $g_{s_1 s_2 \ldots s_m}$. Prove that the coefficient of $g_{s_1 s_2 \ldots s_m}$ in f is $f(s_1, \ldots, s_m) - h(s_1, \ldots, s_m)$. Assume that values of f can be computed and that a bound on the degree of f is known. Show how to express f as a linear combination of the $g_{s_1 s_2 \ldots s_m}$.

9.11 p-Groups

Let p be a prime. A *p-group* is a group in which every element has finite order and these orders are powers of p. A finite p-group has order p^n for some nonnegative integer n. An *elementary abelian* p-group is a finite abelian p-group in which the p-th power of every element is 1. Such a group is a direct product of cyclic groups of order p and may be considered to be a vector space over the field $\mathbb{Z}_p$ of integers modulo p. (Additive notation should be used when this is done.) In Section 11.7 we shall describe an algorithm for determining finite p-groups which are quotients of a given finitely presented group. This section summarizes the facts about finite p-groups which we shall need.

Let P be a nontrivial finite p-group. The following statements are proved in most basic texts on finite groups:

(1) The center of P is nontrivial.
(2) Every proper subgroup of P is contained in a subgroup of index p and all subgroups of index p are normal.
(3) P is nilpotent.

By (3), the lower central series of P eventually reaches the trivial subgroup. However, there is another central series which is particularly useful in studying P. For any group G, define G^p to be the subgroup generated by $\{g^p \mid g \in G\}$. The terms $\varphi_i(G)$ of the *lower exponent-p central series* of G are defined as follows: $\varphi_1(G) = G$ and $\varphi_{i+1}(G) = [\varphi_i(G), G]\varphi_i(G)^p$. By induction, $\varphi_i(G)$ is normal in G and hence so are $[\varphi_i(G), G]$ and G^p. Therefore the product $[\varphi_i(G), G]\varphi_i(G)^p$ is a subgroup. We could also define $\varphi_{i+1}(G)$ as the smallest normal subgroup N of G contained in $\varphi_i(G)$ such that $\varphi_i(G)/N$ is in the center of G/N and is an elementary abelian p-group. Thus the quotients $\varphi_i(G)/\varphi_{i+1}(G)$ are vector spaces over $\mathbb{Z}_p$.

In the case of our finite p-group P, eventually $\varphi_i(P)$ is trivial. For if $\varphi_i(P)$ is not trivial, then $Q = [\varphi_i(P), P]$ is a proper subgroup of $\varphi_i(P)$, since P is nilpotent. The quotient $R = \varphi_i(P)/Q$ is a nontrivial finite abelian p-group and hence is a direct sum of finite cyclic groups of orders which are powers of p. It is easy to show that R^p is a proper subgroup of R. Then $\varphi_{i+1}(P)$, which is the inverse image of R^p in $\varphi_i(P)$, is a proper subgroup of $\varphi_i(P)$. The smallest nonnegative integer c such that $\varphi_{c+1}(P)$ is trivial is called the *exponent-p class* of P.

The subgroup $\varphi_2(P)$ is the intersection of the subgroups of index p in P and is called the Frattini subgroup of P. The quotient $P/\varphi_2(P)$ is the largest elementary abelian quotient of P. In general, the *Frattini subgroup* of a group G is the intersection of the maximal subgroups of G if G has maximal subgroups. If there are no maximal subgroups, then the Frattini subgroup is G. The Frattini subgroup is the set of "nongenerators", elements x in G with the property that, if a subset X of G generates G, then $X - \{x\}$ generates G. Thus a subset X of our finite p-group P generates P if and only if the image of X in $V = P/\varphi_2(P)$ generates V. If V has order p^d, then V is a vector space of dimension d over $\mathbb{Z}_p$ and minimal generating sets of V all have d elements. Thus minimal generating sets of P also have d elements.

Let us fix a prime p. Results analogous to Propositions 1.10, 2.5, and 2.6 hold for the terms of the exponent-p central series.

Proposition 11.1. *For any group G and for any positive integers i and j, the subgroup $[\varphi_i(G), \varphi_j(G)]$ is contained in $\varphi_{i+j}(G)$.*

Proposition 11.2. *Suppose that $G/\varphi_2(G)$ is generated by the images of $x_1, \dots, x_d$. Then $\varphi_2(G)/\varphi_3(G)$ is generated by the images of $x_i^p, 1 \le i \le d$, and $[x_j, x_i]$, $1 \le i < j \le d$.*

Proposition 11.3. *Suppose that $s > 1$, that X is a subset of G which generates G modulo $\varphi_2(G)$, and U is a subset of $\varphi_s(G)$ which generates $\varphi_s(G)$ modulo $\varphi_{s+1}(G)$. Then $\varphi_{s+1}(G)$ is generated modulo $\varphi_{s+2}(G)$ by the elements u^p with u in U and the elements $[u, x]$ with u in U and x in X.*

Let F be a free group of rank r. Propositions 11.2 and 11.3 give upper bounds for the dimensions of the quotients $\varphi_s(F)/\varphi_{s+1}(F)$. In fact, these dimensions are known exactly. Let c_1, $c_2, \dots$ be a basic sequence of commutators in F. Then a basis for $\varphi_s(F)/\varphi_{s+1}(F)$ is given by the elements $c_i^{p^{s-w_i}}, 1 \le i \le t$, where w_i is the weight of c_i and c_t is the last commutator of weight s. See [Huppert & Blackburn 1986].

Let us return to our finite p-group P. It follows from Propositions 11.2 and 11.3 that P has a *φ-weighted presentation* relative to the prime p. This is a power-commutator presentation on generators $a_1, \ldots, a_n$ with the following properties:

(a) Each a_i has associated with it a positive integer weight w_i such that $w_1 = 1$ and $w_i \leq w_{i+1}$, $1 \leq i < n$.
(b) For $1 \leq i \leq n$ there is a power relation $a_i^p = W_i$ and the generators which occur in W_i all have weight at least $w_i + 1$.
(c) For each commutation relation $a_j a_i = a_i a_j A_{ij}$, the generators which occur in A_{ij} all have weight at least $w_i + w_j$.
(d) If $w_k > 1$, then one of the following holds:
  (1) There are indices i and j with $w_i = 1$ and $w_j = w_k - 1$ such that $A_{ij} = a_k$.
  (2) There is an index i with $w_i = w_k - 1$ such that $W_i = a_k$.

  We choose one relation $a_j a_i = a_i a_j a_k$ or $a_i^p = a_k$ as the definition of a_k.

The differences between a φ-weighted presentation and a γ-weighted presentation are the following: In a φ-weighted presentation there is a power relation for each generator, so the group P defined by the presentation is finite; the exponents in the power relations are all equal to a fixed prime p; and generators of weight greater than 1 are defined either as commutators or as p-th powers of earlier generators. Let d be the largest index for which $w_d = 1$. Then P is generated by $x_1, \ldots, x_d$ and d is the minimum number of generators in any generating set. The subgroup $\varphi_s(P)$ is generated modulo $\varphi_{s+1}(P)$ by the generators of weight s.

A φ-weighted presentation is a monoid presentation on generators $a_1, \ldots, a_n$. The techniques of Sections 9.8 and 9.9 can be extended to obtain results which reduce the amount of work needed to test the consistency of a presentation which has the form of a φ-weighted presentation. In fact, (Vaughan-Lee 1984), which provided the ideas for the proof of Proposition 9.3, actually deals with φ-weighted presentations. Let $c = w_n$. If $w_i + w_j = c$, then A_{ij}^p represents the identity in the group defined by the presentation, since all generators of weight c commute and have order dividing p. This implies that a few of the overlaps in Proposition 9.3 do not need to be considered here. To determine consistency we need only test for local confluence at the following overlaps:

$$\begin{array}{rl} a_k a_j a_i, & k > j > i,\ w_k + w_j + w_i \leq c,\ w_i = 1, \\ a_j^p a_i \quad \text{and} \quad a_j a_i^p, & j > i,\ w_j + w_i < c, \\ a_i^{p+1}, & 2w_i < c. \end{array}$$