Chapter 1

Generalities

1.1 Basic concepts

A *word* is an expression of the form

$$w(x_1, \ldots, x_k) = \prod_{j=1}^{s} x_{i_j}^{\varepsilon_j}$$

where $i_1, \ldots, i_s \in [1, k] := \{1, \ldots, k\}$ and each ε_j is ± 1. The *length* of w is s (this can be 0, when w is the 'empty word'). I will keep the symbol k for the number of variables, but this is not supposed to imply that every variable actually occurs in a given word (in practice, we may as well assume that k is a fixed, finite, but very large number). For any group G we have the *verbal mapping*

$$w : G^{(k)} \to G$$
$$\mathbf{g} \mapsto w(\mathbf{g}) = \prod_{i=1}^{s} g_{i_j}^{\varepsilon_j};$$

here and throughout I write

$$\mathbf{g} = (g_1, \ldots, g_k)$$

(and analogously with other letters of course); and $G^{(k)} = G \times \cdots \times G$ with k factors (to avoid confusion with the *subgroup* G^k generated by kth powers in G).

It is sometimes convenient to identify a word with an element of the free group F_k on $\{x_1, \ldots, x_k\}$. Different words may represent the same element of F_k, but of course they all induce the same verbal mapping. Indeed,

$$w(g_1, \ldots, g_k) = w\pi_{\mathbf{g}}$$

where $\pi_{\mathbf{g}}$ is the unique homomorphism $F_k \to G$ sending x_i to g_i for each i. This is discussed in a slightly more general setting in §1.3.

Let w be a word and G a group. We write

$$G_w = \left\{ w(\mathbf{g})^{\pm 1} \mid \mathbf{g} \in G^{(k)} \right\};$$

this is the (symmetrized) set of *w-values* in G.

For any subset S of G and $m \in \mathbb{N}$ we write

$$S^{*m} = \{ s_1 s_2 \dots s_m \mid s_1, s_2, \dots, s_m \in S \},$$

and denote by $\langle S \rangle$ the subgroup of G generated by S. The *verbal subgroup* corresponding to w is

$$w(G) = \langle G_w \rangle .$$

I will say that w has *width m* in G if

$$w(G) = G_w^{*m} .$$

Note: the word 'width' is here used in the wide sense, so 'width m' implies 'width n' for every $n \geq m$; we may define *the* width of w to be the least such m. One says that w has *infinite width* in G if it does not have finite width.

The group G is said to be *verbally elliptic* if every word has finite width in G.

Derived words

For $\mathbf{a}, \mathbf{g} \in G^{(k)}$ we set

$$w'_{\mathbf{g}}(\mathbf{a}) = w(\mathbf{a}.\mathbf{g}) w(\mathbf{g})^{-1}$$

where $\mathbf{a}.\mathbf{g} = (a_1 g_1, \dots, a_k g_k)$. For $H \leq G$ and $Y \subseteq G$ write

$$w'_Y(H) = \left\langle w'_{\mathbf{y}}(\mathbf{a}) \mid \mathbf{a} \in H^{(k)},\ \mathbf{y} \in Y^{(k)} \right\rangle .$$

Lemma 1.1.1 *If $H \lhd G = HY$ for some subset Y of G then*

$$w'_Y(H) = w'_G(H) \lhd G.$$

Proof. $w'_G(H)$ is normal in G because $w'_{\mathbf{g}}(\mathbf{a})^x = w'_{\mathbf{g}^x}(\mathbf{a}^x)$. Suppose $\mathbf{g} = \mathbf{b}.\mathbf{y}$ with $\mathbf{b} \in H^{(k)}$ and $\mathbf{y} \in Y^{(k)}$. Then for $\mathbf{a} \in H^{(k)}$ we have

$$\begin{aligned} w'_{\mathbf{g}}(\mathbf{a}) &= w(\mathbf{a}.\mathbf{b}.\mathbf{y}) w(\mathbf{b}.\mathbf{y})^{-1} \\ &= w'_{\mathbf{y}}(\mathbf{a}.\mathbf{b}) w(\mathbf{y}).w(\mathbf{y})^{-1} w'_{\mathbf{y}}(\mathbf{b})^{-1} \\ &= w'_{\mathbf{y}}(\mathbf{a}.\mathbf{b}) w'_{\mathbf{y}}(\mathbf{b})^{-1} \in w'_Y(H). \end{aligned}$$

■

The marginal subgroup

This is

$$w^*(G) = \left\{ a \in G \mid w(\underline{a}^{(i)}.\mathbf{g}) = w(\mathbf{g}) \ \forall \mathbf{g} \in G^{(k)},\ i = 1, \ldots, k \right\}$$

where

$$\underline{a}^{(i)} = (1, \ldots, a, \ldots, 1)$$

with a in the ith place. It is easily seen that $w^*(G)$ is a characteristic subgroup of G. A subgroup H of G is *marginal* for w if $H \leq w^*(G)$.

It is clear that if $\mathbf{a} \in w^*(G)^{(k)}$ then $w(\mathbf{a}.\mathbf{g}) = w(\mathbf{g})$ for every $\mathbf{g}$. If $a \in w^*(G)$ then also $b_i = [a, g_i^{-1}] \in w^*(G)$ for $i = 1, \ldots, k$, and we have

$$w(\mathbf{g})^a = w(g_1^a, \ldots, g_k^a) = w(\mathbf{b}.\mathbf{g}) = w(\mathbf{g}),$$

whence the important observation:

$$[w(G), w^*(G)] = 1. \tag{1.1}$$

(Actually the argument shows that $[w(G), M] = 1$ where $M/w^*(G) = \mathrm{Z}(G/w^*(G))$. A similar argument shows that the left-right asymmetry in the definition of $w^*(G)$ is only apparent.)

The following basic lemma implies that every word has finite width in a finite group:

Lemma 1.1.2 *If $G = \langle S \rangle$, $1 \in S$ and $|G| = n > 1$ then $G = S^{*(n-1)}$.*

Proof. Since G is finite, every element of G can be written in the form $g = s_1 \ldots s_m$ with each $s_j \in S$. If $m \geq n$ then two of the 'initial segments'

$$1,\ s_1,\ s_1 s_2,\ \ldots,\ s_1 \ldots s_m$$

must be equal; replacing the longer by the shorter we get a shorter expression for g. ∎

Proposition 1.1.3 *Suppose that K is a finite normal subgroup of G. If w has finite width in G/K then w has finite width in G.*

Proof. Put $W = w(G)$. Then

$$WK/K = w(G/K) = (G/K)_w^{*m} = G_w^{*m} K/K$$

where w has width m in G/K, whence

$$W = G_w^{*m}(K \cap W).$$

As $K \cap W$ is finite, there exists $r < \infty$ such that each element of $K \cap W$ is a product of at most r elements of G_w. Then $W = G_w^{*(m+r)}$. ∎

Exercise: Show that $K \cap W$ has a generating set contained in $G_w^{*(2m+1)}$, and deduce that w has width

$$(2m+1)\,|K|$$

in G.

Applying the above proposition to a quotient we get

Corollary 1.1.4 *Let $T < K$ be normal subgroups of G with K/T finite and $T \subseteq G_w^{*n}$ for some n. If w has finite width in G/K then w has finite width in G.*

Concatenation

Let $w_1, \ldots, w_t$ be words in k variables, or functions $G^{(k)} \to G$. Then

$$w_1 \circledast \cdots \circledast w_t$$

is the word, or function, v in tk variables given by

$$v(x_1, \ldots, x_{tk}) = \prod_{i=0}^{t-1} w_{i+1}(x_{ik+1}, \ldots, x_{ik+k}).$$

Note that then $v(G) = \langle w_1(G), \ldots, w_t(G) \rangle$ and that

$$G_v \subseteq G_{w_1} \cdot G_{w_2} \cdot \ldots \cdot G_{w_t} \cup G_{w_t} \cdot G_{w_{t-1}} \cdot \ldots \cdot G_{w_1} \subseteq G_v^{*t}.$$

(When the w_i are words, the set $G_{w_1} \cdot G_{w_2} \cdot \ldots \cdot G_{w_t}$ is invariant under permutations of the factors, since each G_{w_i} is invariant under conjugation; but this is no longer true for the 'generalized word functions' we will meet in §1.3 below.)

Some notation

$$H \leq_f G, \; H \lhd_f G$$

means 'H (is) a subgroup (respectively normal subgroup) of finite index in G'.

$$F_k = F(x_1, \ldots, x_k), \; F_\infty = F(x_1, x_2, \ldots)$$

denote the free group on k, respectively countably infinitely many, free generators $x_1, x_2, \ldots$

Certain important words have a standard notation: the *commutator* of x and y is

$$\gamma_2(x, y) = [x, y] = x^{-1} y^{-1} x y.$$

For $n > 2$, the *left-normed repeated commutator* in n variables is

$$\gamma_n(x_1, \ldots, x_n) = [x_1, \ldots, x_n] = [\gamma_{n-1}(x_1, \ldots, x_{n-1}), x_n].$$

For subgroups $A_1,\ A_2, \ldots,\ A,\ B$ of a group G,

$$[A, B] = [A,_1 B] = \langle [a, b] \mid a \in A,\ b \in B \rangle,$$
$$[A_1, A_2, \ldots, A_n] = [[A_1, A_2, \ldots, A_{n-1}], A_n],$$
$$[A,_n B] = [[A,_{n-1} B], B].$$

But for $g \in G$ and $S \subseteq G$ we write

$$[S, g] = \{[s, g] \mid s \in S\}$$
$$[g, S] = \{[g, s] \mid s \in S\}$$

(each is a *set* of values, not the subgroup they generate).

The verbal subgroups corresponding to the commutator words are

$$G' := \gamma_2(G) = [G, G],$$

the derived group, and for $n > 2$

$$\gamma_n(G) = \left\langle [g_1, \ldots, g_n] \mid \mathbf{g} \in G^{(n)} \right\rangle$$

(*note*: this is not the usual definition! But don't worry, see Exercise 1.2.1 below).

The term **rank** is always used in the sense of *Prüfer rank*; that is

$$\mathrm{rk}(G) = \sup\{\mathrm{d}(H) \mid H \text{ is a finitely generated subgroup of } G\}$$

and $\mathrm{d}(H) \in \mathbb{N} \cup \{\infty\}$ denotes the minimal size of a generating set for H.

1.2 Commutators

The commutator

$$\gamma_2(x, y) = [x, y] = x^{-1}y^{-1}xy$$

is the simplest really interesting word. A good understanding of how it behaves is both the first step to understanding more complicated words, and an essential tool in many arguments in group theory. The key feature of γ_2 as a word mapping is that it is 'bilinear to a first approximation'; this simple observation has a surprising amount of mileage in it (it may be seen as the foundation of Lie theory, for example). Here we record some of its more or less direct consequences. Most of the basic results will be taken for granted in later sections.

Let's start with

Exercise 1.2.1. Prove that if $A_1,\ A_2, \ldots,\ A_n$ are normal subgroups of G then

$$[A_1, A_2, \ldots, A_n] = \langle [a_1, \ldots, a_n] \mid a_j \in A_j\ (j = 1, \ldots, n) \rangle.$$

Deduce that the verbal subgroup for the word γ_n is indeed the nth term of the lower central series of G, namely

$$\gamma_n(G) = [G,_{n-1} G]$$

(so our verbal subgroup notation is consistent with the usual one for the lower central series).

This exercise is a typical application of the following basic identities. These express the bilinearity of γ_2 'modulo higher commutators'; they will frequently be used without special mention.

$$\begin{aligned}
[x,y]^{-1} &= [y,x] = [x^y, y^{-1}] = [x,y^{-1}][x,y^{-1},y] \\
[x^{-1},y] &= [x,y^{x^{-1}}]^{-1} = [x,y,x^{-1}]^{-1}[x,y]^{-1} \\
[xy,z] &= [x,z]^y[y,z] = [x^y,z^y][y,z] = [x,z][x,z,y][y,z] \\
[x,yz] &= [x,z][x,y]^z = [x,z][x^z,y^z] = [x,z][x,y][x,y,z].
\end{aligned}$$

Another useful identity is the *Hall-Witt identity*, the group-theoretic analogue of the Jacobi identity:

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

From this one deduces the **Three-subgroup Lemma**: if X, Y and Z are normal subgroups of G then

$$[X,Y,Z] \leq [Y,Z,X][Z,X,Y];$$

more generally, if X, Y and Z are any subgroups of G then $[X,Y,Z]$ is contained in the normal closure of $[Y,Z,X][Z,X,Y]$.

Arguing by induction on n, it is easy to deduce a fundamental property of the lower central series:

$$[\gamma_m(G), \gamma_n(G)] \leq \gamma_{m+n}(G). \tag{1.2}$$

The multilinear nature of the commutator mapping is expressed in

Proposition 1.2.1 *Let $n \geq 2$ and let $1 \leq i \leq n$. Fix $x_1, \ldots, x_n \in G$ and define $f_i = f_i^n : G \to \gamma_n(G)$ by*

$$f_i(g) = [x_1, \ldots, x_{i-1}, g, x_{i+1}, \ldots, x_n].$$

If $x \in G$ and $y \in \gamma_m(G)$ then

$$\begin{aligned}
f_i(xy) &\equiv f_i(x)f_i(y) \mod \gamma_{n+m}(G) && (1.3)\\
&\equiv f_i(x) \mod \gamma_{n+m-1}(G). && (1.4)
\end{aligned}$$

Proof. Write $G_s = \gamma_s(G)$ for each s. Note that (1.4) follows from (1.3) and (1.2). We first prove (1.3) for $i = 1$, by induction on n. When $n = 2$ the claim is

$$[xy, z] \equiv [x, z][y, z] \mod G_{2+m};$$

this is clear from the basic identities. Suppose $n > 2$, and assume inductively that

$$f_1^{n-1}(xy) = f_1^{n-1}(x) f_1^{n-1}(y) z$$

with $z \in G_{n-1+m}$. Then

$$\begin{aligned} f_1^n(xy) &= [f_1^{n-1}(xy), x_n] \\ &= [f_1^{n-1}(x) f_1^{n-1}(y) z, x_n] \\ &\equiv f_1^n(x) f_1^n(y) \mod G_{n+m}, \end{aligned}$$

again using some basic identities.

Now suppose that $i > 1$, and put $u = [x_1, \ldots, x_{i-1}]$. Then

$$\begin{aligned} f_i^n(xy) &= [u, xy, x_{i+1}, \ldots, x_n] \\ &= [[u, x][u, y]z, x_{i+1}, \ldots, x_n] \\ &= \widetilde{f}_1^{n-i+1}([u, x][u, y]z) \end{aligned}$$

where $z = [[u, y], [u, x]][u, x, y] \in G_{i+m}$, and $\widetilde{f}_1^{n-i+1}$ is obtained from f_1^{n-i+1} on replacing each x_j by x_{i+j-1}. Since $[u, y]z \in G_{i-1+m}$, two applications of the first part with $\widetilde{f}_1$ in place of f_1 show that

$$f_i^n(xy) \equiv \widetilde{f}_1^{n-i+1}([u, x]) \widetilde{f}_1^{n-i+1}([u, y]) \widetilde{f}_1^{n-i+1}(z) \mod G_{n+m}.$$

The result follows since the first two factors are equal, respectively, to $f_i^n(x)$ and $f_i^n(y)$, and the third factor lies in G_{n+m}. (For a slicker and more general proof see Exercise 1.2.2 below.) ∎

Corollary 1.2.2 *Let G be a nilpotent group of class $c \geq n$. Then $\gamma_{c-n+2}(G)$ is marginal for γ_n in G.*

Corollary 1.2.3 *Let G be a nilpotent group of class $c > 1$. Then γ_c induces a multilinear mapping*

$$(G/G')^{(c)} \to \gamma_c(G).$$

Corollary 1.2.4 *Let $w \in \gamma_n(\eta(F_k))$ where η is any word. Let H be any group. Then $\gamma_m(\eta(H))$ is marginal for w in H modulo $\gamma_{n+m-1}(\eta(H))$.*

Proof. Write $F = F_k$. We have

$$w = \prod_j \gamma_n(u_{j1}, \ldots, u_{jn})^{\varepsilon_j}$$

where each $u_{jl} \in \eta(F)$ is a word in $x_1, \ldots, x_k$ and $\varepsilon_j = \pm 1$. Fix $i \in \{1, \ldots, n\}$, let $y \in \gamma_m(\eta(F))$ and put

$$v_{jl} = u_{jl}(x_1, \ldots, x_{i-1}, x_i y, x_{i+1}, \ldots, x_n).$$

Then

$$v_{jl} \equiv u_{ji} \mod \gamma_m(\eta(F))$$

for each j and l. Hence by Proposition 1.2.1 (applied with $G = \eta(F)$) we have

$$\gamma_n(v_{j1}, \ldots, v_{jn}) \equiv \gamma_n(u_{j1}, \ldots, u_{jn}) \mod \gamma_{n+m-1}(\eta(F)).$$

Therefore

$$w(x_1, \ldots, x_{i-1}, x_i y, x_{i+1}, \ldots, x_n) \equiv w(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) = w$$

modulo $\gamma_{n+m-1}(\eta(F))$, and the result follows on evaluating this in H. ■

A different kind of application is:

Proposition 1.2.5 *Let G be a group, H a normal subgroup and suppose that $G = G' \langle x_1, \ldots, x_m \rangle$. Then*

$$[H, G] = [H, x_1] \ldots [H, x_m][H,_n G]$$

for every $n \geq 1$.

Proof. Suppose this holds for a certain value of $n \geq 1$. To deduce that it holds with $n+1$ in place of n we may as well assume that $[H,_{n+1} G] = 1$. Now put $A = [H,_{n-1} G]$. Then $(a, g) \mapsto [a, g]$ induces a bilinear mapping

$$A \times G/G' \to [H,_n G] = B \leq \mathrm{Z}(G).$$

It follows that

$$B = [A, x_1] \ldots [A, x_m].$$

With the inductive hypothesis this implies that

$$\begin{aligned} [H, G] &= [H, x_1] \ldots [H, x_m] B \\ &= [H, x_1] \ldots [H, x_m][A, x_1] \ldots [A, x_m] \\ &= \prod_{i=1}^{m} [H, x_i][A, x_i] \\ &= \prod_{i=1}^{m} [H, x_i]; \end{aligned}$$

the third equality holds because each $[A, x_i] \subseteq \mathrm{Z}(G)$, and the final equality holds because

$$[h, x][a, x] \equiv [ha, x] \mod [H, G, A] = 1$$

for $h \in H$, $a \in A$ and $x \in G$. ■

If G is nilpotent of class c, we can take $n \geq c$ and $H = G$ to infer

Corollary 1.2.6 *Suppose that $G = G' \langle x_1, \ldots, x_m \rangle$ is nilpotent. Then*

$$\gamma_2(G) = [G, x_1] \ldots [G, x_m].$$

It follows that *the word γ_2 has width m in every m-generator nilpotent group.* A similar result holds when G is a *finite* soluble group (Exercise 4.7.4), but the following is open:

Problem 1.2.1 Does γ_2 have finite width in every finitely generated soluble group?

In [S11] Stroud speculates that there may be a counterexample G satisfying $[G''', G] = 1$.

More generally, we have

Proposition 1.2.7 *If $G = G'\langle x_1, \ldots, x_m \rangle$ is nilpotent and $t, r \in \mathbb{N}$ then*

$$\gamma_{t+r}(G) = \prod [\gamma_r(G), x_{i_1}, \ldots, x_{i_t}],$$

the product ranging over $\mathbf{i} = (i_1, \ldots, i_t) \in [1, m]^{(t)}$, in any chosen order.

Proof. Write $G_n = \gamma_n(G)$ for each n. First we prove by induction on t that for $s \in \mathbb{N}$,

$$G_{t+s} = \prod_{\mathbf{i}} [G_s, x_{i_1}, \ldots, x_{i_t}] \cdot G_{t+s+1}. \tag{1.5}$$

If $t = 1$ this follows from Proposition 1.2.5 with $H = G_s$. Now let $t > 1$ and take $H = G_{t-1+s}$. Then

$$\begin{aligned} [H, G] &= \prod_{j=1}^{m} [H, x_j] \cdot G_{t+s+1} \\ &= \prod_{j=1}^{m} \left[\prod\nolimits_{\mathbf{i}} [G_s, x_{i_1}, \ldots, x_{i_{t-1}}] \cdot G_{t+s}, x_j \right] \cdot G_{t+s+1} \\ &= \prod_{\mathbf{j}} \left[[G_s, x_{i_1}, \ldots, x_{i_{t-1}}], x_j \right] \cdot G_{t+s+1}, \end{aligned}$$

using the inductive hypothesis at the second step and Proposition 1.2.1 at the third step. This establishes (1.5).

Now let $n \geq 1$ and suppose that

$$G_{t+s} = \prod_{\mathbf{i}} [G_s, x_{i_1}, \ldots, x_{i_t}] \cdot G_{t+s+n} \tag{1.6}$$

for every $s \in \mathbb{N}$. Taking $s = r$ and $s = r + n$ we get

$$\begin{aligned} G_{t+r} &= \prod_{\mathbf{i}} [G_r, x_{i_1}, \ldots, x_{i_t}] \cdot \prod_{\mathbf{i}} [G_{r+n}, x_{i_1}, \ldots, x_{i_t}] \cdot G_{t+r+n+1} \\ &= \prod_{\mathbf{i}} [G_r, x_{i_1}, \ldots, x_{i_t}] \cdot G_{t+r+n+1} \end{aligned}$$

by Proposition 1.2.1. Thus by induction (1.6) holds for every n, and the result follows since $G_{t+r+n} = 1$ for sufficiently large n. ∎

Corollary 1.2.8 *If $G = G' \langle x_1, \ldots, x_m \rangle$ is nilpotent and $t \geq 1$ then*

$$\gamma_{t+1}(G) = \prod_{\mathbf{i} \in [1,m]^t} [G, x_{i_1}, \ldots, x_{i_t}].$$

The word γ_{t+1} has width m^t in G.

Remark. Recall that if G is a nilpotent group, we have $G = G' \langle x_1, \ldots, x_m \rangle$ if and only if $G = \langle x_1, \ldots, x_m \rangle$ – e.g. because the subgroup $\langle x_1, \ldots, x_m \rangle$ is in any case subnormal in G, so if it is proper it is contained in some proper normal subgroup N, and then $G > G'N \geq G' \langle x_1, \ldots, x_m \rangle$.

Exercise 1.2.2. (i) Let U, A and B be normal subgroups of a group G and for $n \geq 0$ set

$$Q_n = \prod_{r+s=n} [A,_r G, B,_s G], \quad P_n = \prod_{r+s=n} [U, A,_r G, B,_s G]$$

(reading $[X,_0 Y] = X$). Prove that if $u \in U$, $a \in A$, $b \in B$ and $y_1, \ldots, y_n \in G$ then

$$[ab, y_1, \ldots, y_n] \equiv [a, y_1, \ldots, y_n][b, y_1, \ldots, y_n] \mod Q_n$$

and

$$[u, ab, y_1, \ldots, y_n] \equiv [u, a, y_1, \ldots, y_n][u, b, y_1, \ldots, y_n] \mod P_n.$$

[*Hint*: Both parts are proved the same way, using induction on n.]

(ii) Deduce Proposition 1.2.1.

Exercise 1.2.3. Let G be a group, H a normal subgroup and suppose that $G = G' \langle x_1, \ldots, x_m \rangle$. Prove that

$$[H,_t G] = \prod [H, x_{i_1}, \ldots, x_{i_t}] \cdot [H,_l G]$$

for every $l \geq t$, the product ranging over $\mathbf{i} = (i_1, \ldots, i_t) \in [1, m]^{(t)}$, in any chosen order. [*Hint*: assume without loss of generality that $[H,_l G] = 1$, and argue by induction on $l - t$, using the preceding exercise.]

1.3 Generalized words

It is usual to consider words as elements of a free group. We shall also need to deal with 'generalized words', that is, words twisted by group automorphisms. These are conveniently identified with elements of a free 'group with operators'. Let $X = \{x_1, \ldots, x_k\}$ and $\Phi = \{\phi_1, \ldots, \phi_s\}$ be disjoint finite alphabets, and set

$$F_\Phi(X) = \left\langle x^\phi \mid x \in X,\ \phi \in \Phi \right\rangle < F(X \cup \Phi)$$

where $F(X \cup \Phi)$ is the free group on $X \cup \Phi$. It is easy to see that $F_\Phi(X)$ is actually free on the exhibited generating set (*Exercise*: consider its image under the natural homomorphism from $F(X \cup \Phi)$ onto the wreath product