



## CONCISENESS ON NORMAL SUBGROUPS AND NEW CONCISE WORDS FROM LOWER CENTRAL AND DERIVED WORDS

G. A. FERNÁNDEZ-ALCOBER\* AND M. PINTONELLO

*Dedicated to Prof. D. J. S. Robinson*

ABSTRACT. Let  $w = w(x_1, \dots, x_r)$  be a lower central word or a derived word. We show that the word  $w(u_1, \dots, u_r)$  is concise, whenever  $u_1, \dots, u_r$  are non-commutator words in disjoint sets of variables. This proves a generalized version of a conjecture of Azevedo and Shumyatsky. This applies in particular to words of the form  $w(x_1^{n_1}, \dots, x_r^{n_r})$ , where the  $n_i$  are non-zero integers. Our approach is via the study of values of  $w$  on normal subgroups, and in this setting we obtain the following result: if  $N_1, \dots, N_r$  are normal subgroups of a group  $G$  and the set of all values  $w(g_1, \dots, g_r)$  with  $g_i \in N_i$  is finite, then also the subgroup generated by these values, i.e.  $w(N_1, \dots, N_r)$ , is finite.

### 1. Introduction

Let  $w = w(x_1, \dots, x_r)$  be a group word, which we identify with an element of the free group  $F = F(X)$  of countable rank on the set  $X = \{x_n\}_{n \in \mathbb{N}}$ . For every group  $G$  and every  $r$ -tuple  $\mathbf{g} = (g_1, \dots, g_r)$  of elements of  $G$ , we can consider the element  $w(\mathbf{g})$  resulting from replacing each  $x_i$  with  $g_i$ . We thus obtain the *set of word values* of  $w$  in  $G$ ,

$$w\{G\} = \{w(\mathbf{g}) \mid \mathbf{g} \in G \times \cdots \times G\}.$$

The subgroup generated by  $w\{G\}$  is called the *verbal subgroup* of  $w$  in  $G$ , and is denoted by  $w(G)$ . More generally, if  $\mathbf{S} = (S_1, \dots, S_r)$  is an  $r$ -tuple of subsets of  $G$ , we can consider the set of values

$$w\{\mathbf{S}\} = \{w(\mathbf{g}) \mid \mathbf{g} \in S_1 \times \cdots \times S_r\},$$

---

Communicated by Alireza Abdollahi

MSC(2020): Primary: 20F10; Secondary: 20F14.

Keywords: Word values, verbal subgroup, concise word.

Received: 12 April 2023, Accepted: 18 July 2023.

\*Corresponding author

and the corresponding verbal subgroup on  $\mathbf{S}$ , namely  $w(\mathbf{S}) = \langle w\{\mathbf{S}\} \rangle$ . Of special interest is the case when  $\mathbf{S}$  is a tuple  $\mathbf{N} = (N_1, \dots, N_r)$  of normal subgroups of  $G$ . We then say that  $w(\mathbf{N})$  is the  $\mathbf{N}$ -verbal subgroup of  $w$  and that it is a verbal subgroup on normal subgroups.

If a word  $w$  satisfies the property that, whenever the set of values  $w\{G\}$  is finite in a group  $G$ , then the verbal subgroup  $w(G)$  is also finite, we say that  $w$  is a *concise* word. Philip Hall conjectured that all words are concise. This is one of his three famous problems on verbal subgroups; see Section 4.2 in Robinson's book [12] or Section 1.4 in Segal's book [13]. In the attempt to settle this conjecture, the following types of words were proved to be concise:

- (i) Non-commutator words, that is, words lying outside  $F'$ . Equivalently, these are the words for which the exponent sum of some variable  $x_i$  is non-zero. (P. Hall, unpublished.)
- (ii) Lower commutator words  $\gamma_r$ , defined by

$$\gamma_r(x_1, \dots, x_r) = [x_1, \dots, x_r],$$

and whose corresponding verbal subgroup is the  $r$ th term of the lower central series. (P. Hall, unpublished.)

- (iii) Derived words  $\delta_k$ , defined recursively by  $\delta_0 = x_1$  and

$$\delta_k(x_1, \dots, x_{2^k}) = [\delta_{k-1}(x_1, \dots, x_{2^{k-1}}), \delta_{k-1}(x_{2^{k-1}+1}, \dots, x_{2^k})].$$

(Turner-Smith [14].) The verbal subgroup of  $\delta_k$  is the  $k$ th derived subgroup of the group.

- (iv) More generally, outer commutator words, also known as multilinear words. (Wilson [15].) These are words obtained by nesting commutators of variables arbitrarily, but never repeating a variable. For example,  $[[x_1, x_2, x_3], [[x_4, x_5], [x_6, x_7]]]$  is an outer commutator.

For convenience, a single variable is also considered to be an outer commutator word. Obviously, lower central words and derived words are outer commutator words.

However, Ivanov [10] refuted P. Hall's conjecture in 1989, showing that the word  $[[x^{pn}, y^{pn}]^n, y^{pn}]^n$  is not concise for big enough  $n$  and  $p$ , where  $n$  is odd and  $p$  is a prime. Another counterexample can be found in Theorem 39.7 of Olshanskii's book [11].

It was not until 2011, when some new examples of concise words were discovered. In two independent papers, Abdollahi and Russo [1], and Fernández-Alcober, Morigi and Traustason [7] proved that the Engel word

$$e_n(x, y) = [x, y, \overset{n}{.}, y]$$

is concise for  $n = 2, 3$  and  $4$ . To date, it is not known whether  $e_n$  is concise for  $n \geq 5$ . The topic of conciseness experienced a revival in the last years in the works of Shumyatsky and his collaborators. On the one hand, they have studied some variations of conciseness in profinite groups, and on the other hand, they have proved that the following families of words are concise in all groups:

- (v) If  $u_1$  and  $u_2$  are non-commutator words in disjoint sets of variables, then  $[u_1, u_2]$  is concise. (Delizia, Shumyatsky, Tortora, and Tota [4].) The same happens with commutators  $[u_1, u_2, u_3]$  of non-commutators. (Azevedo, Shumyatsky [2].)

- (vi) If  $u = u(x_1, \dots, x_k)$  is a non-commutator word and we produce copies of  $u$  in disjoint variables by setting  $u_i = u(x_{i1}, \dots, x_{ik})$  for  $i = 1, \dots, r$ , then  $[u_1, \dots, u_r]$  and  $[v, u_1, \dots, u_r]$  are concise, where  $v$  is another non-commutator in a set of variables disjoint from those of  $u_1, \dots, u_r$ . (Azevedo, Shumyatsky [2].)
- (vii) If  $u$  is an outer commutator word and  $v$  is a non-commutator in disjoint variables, then  $[u, v]$  is concise. (Azevedo, Shumyatsky [2].)

In the light of (v) and (vi) above, Azevedo and Shumyatsky made the following conjecture in [2].

**Conjecture** (Azevedo, Shumyatsky). *Let  $u_1, \dots, u_r$  be non-commutator words in disjoint sets of variables. Then the word*

$$\gamma_r(u_1, \dots, u_r) = [u_1, \dots, u_r]$$

*is concise.*

The main result in this paper is the proof of this conjecture, not only for lower central words, but also for derived words.

**Theorem A.** *Let  $w = w(x_1, \dots, x_r)$  be a lower central word or a derived word. If  $u_1, \dots, u_r$  are non-commutator words in disjoint sets of variables, then the word  $w(u_1, \dots, u_r)$  is concise. In particular, the word  $w(x_1^{n_1}, \dots, x_r^{n_r})$  is concise whenever  $n_1, \dots, n_r \in \mathbb{Z} \setminus \{0\}$ .*

In [6] Fernández-Alcober and Morigi showed that if a word  $w$  is concise then it is actually *boundedly concise*, that is, there is a function  $f_w : \mathbb{N} \rightarrow \mathbb{N}$  such that whenever  $|w\{G\}| = m$  in a group  $G$ , we have  $|w(G)| \leq f_w(m)$ . Their argument requires the use of non-principal ultrafilters over infinite sets, whose existence is independent of Zermelo-Fraenkel set theory. We remark that our proof of Theorem A shows that the words therein are boundedly concise, without relying on the result in [6].

Our approach to Theorem A is via the study of  $w\{\mathbf{N}\}$  and  $w(\mathbf{N})$  for an  $r$ -tuple  $\mathbf{N} = (N_1, \dots, N_r)$  of normal subgroups of a group  $G$ . More precisely, for every  $i = 1, \dots, r$  we consider a generating set  $S_i$  of  $N_i$ , that is a normal subset of  $G$ , and relate  $w(\mathbf{N})$  to  $w\{\mathbf{S}\}$ , where  $\mathbf{S} = (S_1, \dots, S_r)$ . Under an additional assumption on the sets  $S_i$ , we are able to prove that the finiteness of  $w\{\mathbf{S}\}$  implies that of  $w(\mathbf{N})$  (see Theorem 2.11 and Theorem 3.6). Theorem A is a corollary of this general result, and another consequence of it is the theorem below, which we think has an independent interest.

**Theorem B.** *Let  $w = w(x_1, \dots, x_r)$  be a lower central word or a derived word. Assume that  $\mathbf{N} = (N_1, \dots, N_r)$  is a tuple of normal subgroups of a group  $G$ , such that  $w\{\mathbf{N}\}$  is finite. Then the subgroup  $w(\mathbf{N})$  is also finite.*

If a word  $w$  satisfies the property stated in Theorem B, we say that  $w$  is *concise on normal subgroups*. Note that Theorem B was known for the usual commutator word  $w = [x, y]$ , as a consequence of Theorem 4.14 in Robinson’s book [12]. Actually, it suffices to require that either of the subgroups  $H$  or  $K$  are ascendant in  $\langle H, K \rangle$  in order to conclude that  $[H, K]$  is finite whenever the set  $\{[h, k] \mid h \in H, k \in K\}$  is finite. Baer [3] had previously proved this result under the assumption that one of  $H$  or  $K$  is normal in  $\langle H, K \rangle$ . We impose the stronger condition that all subgroups should be normal, but

in return , we generalize the result to all lower central and derived words. Theorem B actually holds with bounds: if  $w\{\mathbf{N}\}$  is of order  $m$ , then the order of  $w(\mathbf{N})$  can be bounded by a function of  $m$  and  $w$ . We conclude this introduction with an extension of the conjecture of Azevedo and Shumyatsky.

**Conjecture.** *Theorems A and B hold for an arbitrary outer commutator word  $w$ .*

## 2. Preliminaries and lower central words

We start with a few results regarding word values and verbal subgroups on normal subgroups or on normal subsets, in the case of outer commutator words. If  $w = w(x_1, \dots, x_r)$  is an outer commutator word that is not a variable, then we can write  $w = [\alpha, \beta]$ , where  $\alpha$  and  $\beta$  are again outer commutator words. Without loss of generality, after renaming variables if necessary, we may assume that  $\alpha = \alpha(x_1, \dots, x_q)$  and  $\beta = \beta(x_{q+1}, \dots, x_r)$ , with  $1 \leq q < r$ .

**Lemma 2.1.** *Let  $w = w(x_1, \dots, x_r)$  be an outer commutator word, and let  $\mathbf{N} = (N_1, \dots, N_r)$  be an  $r$ -tuple of normal subgroups of a group  $G$ .*

- (i) *Assume that  $w = [\alpha, \beta]$ , with  $\alpha = \alpha(x_1, \dots, x_q)$  and  $\beta = \beta(x_{q+1}, \dots, x_r)$ . If we set  $\mathbf{N}_1 = (N_1, \dots, N_q)$  and  $\mathbf{N}_2 = (N_{q+1}, \dots, N_r)$ , then  $w(\mathbf{N}) = [\alpha(\mathbf{N}_1), \beta(\mathbf{N}_2)]$ .*
- (ii) *Assume that  $N_i = \langle S_i \rangle$  for every  $i = 1, \dots, r$ , where each  $S_i$  is a normal subset of  $G$ . If we set  $\mathbf{S} = (S_1, \dots, S_r)$ , then the subgroup  $w(\mathbf{N})$  is generated by  $w\{\mathbf{S}\}$ .*

*Proof.* Both (i) and (ii) follow immediately from the simple fact that if  $S$  and  $T$  are two normal subsets of a group  $G$ , then

$$[\langle S \rangle, \langle T \rangle] = \langle [s, t] \mid s \in S, t \in T \rangle,$$

where for part (ii) we use (i) and induction on the number of variables. □

We are interested in words of the form  $w(u_1, \dots, u_r)$ , where  $u_1, \dots, u_r$  are non-commutator words that involve different variables. Let us introduce the following concept.

**Definition 2.2.** *Let  $u_1, \dots, u_r$  be group words. We say that these words are disjoint if the sets of variables that they involve, are pairwise disjoint.*

If  $w$  is a word in  $r$  variables and  $u_1, \dots, u_r$  are disjoint words, then the set of values of the word  $w^* = w(u_1, \dots, u_r)$  in a group  $G$  can be written as  $w\{\mathbf{S}\}$ , where

$$\mathbf{S} = (u_1\{G\}, \dots, u_r\{G\}).$$

Since every  $u_i\{G\}$  is a normal subset of  $G$ , we get the following consequence of the previous lemma.

**Corollary 2.3.** *Let  $w = w(x_1, \dots, x_r)$  be an outer commutator word, and let  $u_1, \dots, u_r$  be arbitrary disjoint words. If  $w^* = w(u_1, \dots, u_r)$  then for every group  $G$  we have*

$$w^*(G) = w(u_1(G), \dots, u_r(G)).$$

Now we want to make part (ii) of Lemma 2.1 quantitative. If we take a standard generator  $w(n_1, \dots, n_r)$  of  $w(\mathbf{N})$ , with  $n_i \in N_i$ , how can we estimate the number of factors from  $w\{\mathbf{S}\}^{\pm 1}$ , that are needed to write it? We need to introduce the following notation.

**Definition 2.4.** Let  $G$  be a group, and let  $S$  be a subset of  $G$ . For every  $n \in \mathbb{N}$ , we define  $S^{*n}$  to be the set of all products of elements of  $S \cup S^{-1}$  of length at most  $n$ .

In other references,  $S^{*n}$  is defined as the set of products of exactly  $n$  elements of  $S$ . Since we can always replace  $S$  with  $S \cup S^{-1} \cup \{1\}$ , both definitions are basically equivalent. We prefer the definition above because it suits better the description of the elements of the subgroup  $\langle S \rangle$ . Also, with this definition, we have  $S^{*k} \subseteq S^{*n}$  whenever  $n \geq k$ . Note that if  $|S| \leq m$  then  $|S^{*n}| \leq (2m + 1)^n$  for every  $n \in \mathbb{N}$ . Let us connect this concept with the values of outer commutator words.

**Lemma 2.5.** Let  $w = w(x_1, \dots, x_r)$  be an outer commutator word, and let  $S$  be a normal subset of a group  $G$ . Suppose that  $\mathbf{t} = (t_1, \dots, t_r)$  is a tuple of elements of  $G$ , one of whose components belongs to  $S$ . Then  $w(\mathbf{t}) \in S^{*2^{r-1}}$ .

*Proof.* The result is obvious for  $r = 1$ , so we assume  $r > 1$ . Then we can write  $w(\mathbf{t}) = [\alpha(\mathbf{t}'), \beta(\mathbf{t}'')]$ , where  $\alpha$  and  $\beta$  are outer commutator words, and the tuples  $\mathbf{t}'$  and  $\mathbf{t}''$  form a partition of  $\mathbf{t}$ . Assume without loss of generality that  $\mathbf{t}'$  contains an entry from  $S$ . By induction on  $r$ , we have  $\alpha(\mathbf{t}') \in S^{*2^{r-2}}$ . Consequently,

$$w(\mathbf{t}) = \alpha(\mathbf{t}')^{-1} \alpha(\mathbf{t}')^{\beta(\mathbf{t}'')} \in S^{*2^{r-1}},$$

since  $S$  is a normal subset of  $G$ . □

On the other hand, by [9, Lemma 2.8], if  $w = w(x_1, \dots, x_r)$  is an outer commutator and  $g_1, \dots, g_r, h$  are elements of a group  $G$ , then for every  $i = 1, \dots, r$  we have

$$w(g_1, \dots, g_{i-1}, g_i h, g_{i+1}, \dots, g_r) = w(g_1^*, \dots, g_{i-1}^*, g_i^*, g_{i+1}^*, \dots, g_r^*) \cdot w(g_1, \dots, g_{i-1}, h, g_{i+1}, \dots, g_r),$$

where  $g_j^*$  is a conjugate of  $g_j$  in  $G$  for every  $j = 1, \dots, r$ . The following lemma follows easily from this result by induction on  $m_1 \cdots m_r$ .

**Lemma 2.6.** Let  $w = w(x_1, \dots, x_r)$  be an outer commutator word, and let  $\mathbf{S} = (S_1, \dots, S_r)$  be a tuple of normal subsets of a group  $G$ . If  $\mathbf{t} = (t_1, \dots, t_r)$  with  $t_i \in S_i^{*m_i}$  for every  $i = 1, \dots, r$ , then

$$w(\mathbf{t}) \in w\{\mathbf{S}\}^{*m_1 \cdots m_r}.$$

In an abelian group  $G$ , the word map  $(g_1, \dots, g_r) \mapsto w(g_1, \dots, g_r)$  is a group homomorphism for every word  $w$ , and consequently  $w(G) = w\{G\}$ . Of course, this is far from being true in arbitrary groups. Outer commutator words, although they are also called multilinear words because the same type of commutator arrangements yields multilinear words in Lie rings, are not multilinear in groups. However, our approach to proving Theorems A and B relies on showing that, in suitable sections that cover the section  $w(\mathbf{N})/w(\mathbf{N})'$ , outer commutator words are linear in one specific variable (which depends on the section). To this purpose, we give the following definition.

**Definition 2.7.** Let  $w = w(x_1, \dots, x_r)$  be a word and let  $\mathbf{N} = (N_1, \dots, N_r)$  be an  $r$ -tuple of normal subgroups of a group  $G$ . We say that  $w$  is linear in position  $i$  of the tuple  $\mathbf{N}$  provided that, for all  $g_j \in N_j$  for  $j = 1, \dots, r$  and  $h_i \in N_i$ , we have

$$(2.1) \quad w(g_1, \dots, g_{i-1}, g_i h_i, g_{i+1}, \dots, g_r) = w(g_1, \dots, g_{i-1}, g_i, g_{i+1}, \dots, g_r) \cdot w(g_1, \dots, g_{i-1}, h_i, g_{i+1}, \dots, g_r).$$

Typically, we will search for linearity in a normal section  $K/L$  of the ambient group  $G$ , which is generated by the image of  $w\{\mathbf{N}\}$ , so that condition (2.1) above is required to hold modulo  $L$ . Obviously, this type of linearity is inherited by sections of the form  $KN/LN$  for a given  $N \trianglelefteq G$ . Next we show that, it is also preserved under taking suitable commutators.

**Lemma 2.8.** Let  $w = [\alpha, \beta]$  be an outer commutator word, where  $\alpha = \alpha(x_1, \dots, x_q)$  and  $\beta = \beta(x_{q+1}, \dots, x_r)$ . Assume that  $\mathbf{N} = (N_1, \dots, N_r)$  is a tuple of normal subgroups of a group  $G$ , and set  $\mathbf{N}_1 = (N_1, \dots, N_q)$  and  $\mathbf{N}_2 = (N_{q+1}, \dots, N_r)$ . Then the following hold:

- (i) If  $K/L$  is a normal section of  $G$  generated by the image of  $\alpha\{\mathbf{N}_1\}$  and  $\alpha$  is linear in component  $i$  of  $\mathbf{N}_1$  modulo  $L$ , then the section  $U/V$ , where  $U = [K, \beta(\mathbf{N}_2)]$  and  $V = [w(\mathbf{N}), \alpha(\mathbf{N}_1)][L, \beta(\mathbf{N}_2)]$ , is generated by the image of  $w\{\mathbf{N}\}$  and  $w$  is linear in component  $i$  of  $\mathbf{N}$  modulo  $V$ .
- (ii) If  $K/L$  is a normal section of  $G$  generated by the image of  $\beta\{\mathbf{N}_2\}$  and  $\beta$  is linear in component  $i$  of  $\mathbf{N}_2$  modulo  $L$ , then the section  $U/V$ , with  $U = [\alpha(\mathbf{N}_1), K]$  and  $V = [w(\mathbf{N}), \beta(\mathbf{N}_2)][\alpha(\mathbf{N}_1), L]$ , is generated by the image of  $w\{\mathbf{N}\}$  and  $w$  is linear in component  $q + i$  of  $\mathbf{N}$  modulo  $V$ .

*Proof.* We only prove part (i). To start with, we have

$$[K, \beta(\mathbf{N}_2)] = [\alpha(\mathbf{N}_1)L, \beta(\mathbf{N}_2)] = [\alpha(\mathbf{N}_1), \beta(\mathbf{N}_2)] [L, \beta(\mathbf{N}_2)] = w(\mathbf{N}) [L, \beta(\mathbf{N}_2)],$$

where the last equality follows from (i) of Lemma 2.1. Thus the section  $U/V$  is generated by the image of  $w\{\mathbf{N}\}$ .

As for the assertion about linearity, let us consider the general congruence stating that  $\alpha$  is linear in component  $i$  of  $\mathbf{N}_1$  modulo  $L$ . This can be written in the form  $x \equiv yz \pmod{L}$ , where  $x, y$ , and  $z$  are like the three elements appearing in (2.1) (with  $\alpha$  playing the role of  $w$ ). In particular,  $x, y, z \in \alpha\{\mathbf{N}_1\}$ . Standard commutator identities then yield that, for every  $n \in \beta\{\mathbf{N}_2\}$ , we have

$$[x, n] \equiv [y, n][z, n] \pmod{[\alpha(\mathbf{N}_1), \beta(\mathbf{N}_2), \alpha(\mathbf{N}_1)] [L, \beta(\mathbf{N}_2)]}.$$

This proves the result. □

We can use the previous lemma to determine, for every lower central word  $\gamma_r$  and every  $r$ -tuple  $\mathbf{N}$  of normal subgroups, a series from  $[\gamma_r(\mathbf{N}), \gamma_r(\mathbf{N})]$  to  $\gamma_r(\mathbf{N})$ , that is linear for  $\gamma_r$  at every section.

**Theorem 2.9.** Let  $r \in \mathbb{N}$ . Assume that  $\mathbf{N} = (N_1, \dots, N_r)$  is a tuple of normal subgroups of a group  $G$ , and define

$$\mathbf{N}_i = (N_1, \dots, N_{i-1}, \gamma_i(N_1, \dots, N_i), N_{i+1}, \dots, N_r)$$

DOI: <https://dx.doi.org/10.30504/JIMS.2023.392862.1105>

for every  $i = 1, \dots, r$ . Then there is a series

$$[\gamma_r(\mathbf{N}), \gamma_r(\mathbf{N})] = P_{r+1}^r \leq P_r^r \leq \dots \leq P_i^r \leq \dots \leq P_1^r = \gamma_r(\mathbf{N})$$

such that, for every  $i = 1, \dots, r$ , the section  $P_i^r/P_{i+1}^r$  is generated by the image of  $\gamma_r\{\mathbf{N}_i\}$  and the word  $\gamma_r$  is linear in component  $i$  of  $\mathbf{N}_i$  modulo  $P_{i+1}^r$ .

*Proof.* Set  $Q_i^r = \gamma_r(\mathbf{N}_i)$  for every  $i = 1, \dots, r$ , and  $Q_{r+1}^r = [\gamma_r(\mathbf{N}), \gamma_r(\mathbf{N})]$ . Obviously, the conditions that  $P_{r+1}^r = [\gamma_r(\mathbf{N}), \gamma_r(\mathbf{N})]$  and that  $P_i^r/P_{i+1}^r$  is generated by the image of  $Q_i^r$  mean that we need to choose  $P_i^r = Q_i^r Q_{i+1}^r \dots Q_{r+1}^r$ , for  $i = 1, \dots, r + 1$ .

Let us then prove the linearity of  $\gamma_r$  in component  $i$  of  $\gamma_r(\mathbf{N}_i)$  modulo  $P_{i+1}^r$ . We argue by induction on  $r - i$ . The basis of the induction,  $i = r$ , follows from the congruence  $[g, x_r y_r] \equiv [g, x_r][g, y_r] \pmod{P_{r+1}^r}$  for all  $g \in G$  and  $x_r, y_r \in \gamma_r(\mathbf{N})$ , which holds because  $P_{r+1}^r = [\gamma_r(\mathbf{N}), \gamma_r(\mathbf{N})]$ .

Let us now assume that  $1 \leq i < r$  and that the result is true for differences less than  $r - i$ . For every  $i = 1, \dots, r$ , let  $Q_i^{r-1}$  and  $P_i^{r-1}$  be defined from the tuple  $\mathbf{N}^* = (N_1, \dots, N_{r-1})$  in the same way as we defined  $P_i^r$  and  $Q_i^r$  from  $\mathbf{N}$ . Then linearity holds in position  $i$  of

$$\mathbf{N}_i^* = (N_1, \dots, N_{i-1}, \gamma_i(N_1, \dots, N_i), N_{i+1}, \dots, N_{r-1})$$

modulo  $P_{i+1}^{r-1}$ . Now we apply Lemma 2.8 by taking  $K = P_i^{r-1}$ ,  $L = P_{i+1}^{r-1}$ ,  $\alpha = \gamma_{r-1}$  and  $\beta = x_r$ . Thus  $\gamma_r$  is linear in component  $i$  of  $\mathbf{N}_i$  modulo the subgroup

$$(2.2) \quad [\gamma_r(\mathbf{N}), \gamma_{r-1}(\mathbf{N}^*)] [P_{i+1}^{r-1}, N_r].$$

Observe that

$$[\gamma_r(\mathbf{N}), \gamma_{r-1}(\mathbf{N}^*)] = [N_1, \dots, N_{r-1}, \gamma_r(N_1, \dots, N_r)] = Q_r^r \leq P_{i+1}^r,$$

since  $r - i \geq 1$ . On the other hand,

$$[P_{i+1}^{r-1}, N_r] = \left( \prod_{j=i+1}^{r-1} [Q_j^{r-1}, N_r] \right) \cdot [Q_r^{r-1}, N_r] = \left( \prod_{j=i+1}^{r-1} Q_j^r \right) \cdot [Q_r^{r-1}, N_r],$$

and

$$[Q_r^{r-1}, N_r] = [\gamma_{r-1}(N_1, \dots, N_{r-1}), \gamma_{r-1}(N_1, \dots, N_{r-1}), N_r] \leq [N_1, \dots, N_{r-1}, \gamma_r(N_1, \dots, N_r)] = Q_r^r,$$

where the inclusion follows from P. Hall’s Three Subgroup Lemma. Hence the subgroup in (2.2) is contained in  $P_{i+1}^r$ , and the result follows.  $\square$

We are now in a position to prove the key theorem that will provide both Theorem A and Theorem B for lower central words. We need the following version for normal subgroups of a well-known lemma in the theory of concise words (see, for example, [5, Lemma 4]). The proof is exactly the same, based on Schur’s Theorem, and taking into account also part (ii) of Lemma 2.1 in this case, so we omit it. In the remainder of the paper, for a tuple  $S$  of parameters, we use the expression  $S$ -bounded to mean “bounded by a function of  $S$ ”.

**Lemma 2.10.** *Let  $w = w(x_1, \dots, x_r)$  be an arbitrary word, and let  $\mathbf{N} = (N_1, \dots, N_r)$  be an  $r$ -tuple of normal subgroups of a group  $G$ . Suppose that  $N_i = \langle S_i \rangle$ , where  $S_i$  is a normal subset of  $G$  for every  $i = 1, \dots, r$ , and set  $\mathbf{S} = (S_1, \dots, S_r)$ . If  $w\{\mathbf{S}\}$  is finite of order  $m$  then  $w(\mathbf{N})'$  is finite of  $m$ -bounded order.*

**Theorem 2.11.** *Let  $r \in \mathbb{N}$  and let  $\mathbf{N} = (N_1, \dots, N_r)$  be a tuple of normal subgroups of a group  $G$ . Assume that  $N_i = \langle S_i \rangle$  for every  $i = 1, \dots, r$ , where:*

- (i)  $S_i$  is a normal subset of  $G$ .
- (ii) There exists  $n_i \in \mathbb{N}$  such that all  $n_i$ th powers of elements of  $N_i$  are contained in  $S_i$ .

*If for the tuple  $\mathbf{S} = (S_1, \dots, S_r)$  the set of values  $\gamma_r\{\mathbf{S}\}$  is finite of order  $m$ , then the subgroup  $\gamma_r(\mathbf{N})$  is also finite, of  $(m, r, n_1, \dots, n_r)$ -bounded order.*

*Proof.* We follow the notation  $\mathbf{N}_i$  and  $P_i^r$ , introduced in the statement of Theorem 2.9. We are going to prove that  $P_i^r$  is finite of bounded order for  $i = 1, \dots, r + 1$  by reverse induction on  $i$ . Since  $P_1^r = \gamma_r(\mathbf{N})$ , this proves the result.

The basis of the induction follows from Lemma 2.10, since  $P_{r+1}^r = [\gamma_r(\mathbf{N}), \gamma_r(\mathbf{N})]$ . Let us assume that  $P_{i+1}^r$  is finite of bounded order and prove that the same holds for  $P_i^r$ . Recall that the quotient  $P_i^r/P_{i+1}^r$  is the image of  $\gamma_r(\mathbf{N}_i)$ , and then, by a suitable application of Lemma 2.1, it can be generated by the images of the set  $\mathbf{T}$  of commutators

$$[s_1, \dots, s_{i-1}, x_i, s_{i+1}, \dots, s_r],$$

with  $s_j \in S_j$  for  $1 \leq j \leq r$ ,  $j \neq i$ , and  $x_i \in \gamma_i\{\mathbf{S}_i\}$ , where  $\mathbf{S}_i = (S_1, \dots, S_i)$ . By Lemma 2.5, we have  $\gamma_i\{\mathbf{S}_i\} \subseteq S_i^{*2^{i-1}}$ , and then Lemma 2.6 implies that

$$[s_1, \dots, s_{i-1}, x_i, s_{i+1}, \dots, s_r] \in \gamma_r\{\mathbf{S}\}^{*2^{i-1}} \subseteq \gamma_r\{\mathbf{S}\}^{*2^{r-1}}.$$

From the assumption that  $|\gamma_r\{\mathbf{S}\}| = m$ , we get

$$|\mathbf{T}| \leq (2m + 1)^{2^{r-1}},$$

and consequently  $P_i^r/P_{i+1}^r$  can be generated by an  $(m, r)$ -bounded number of elements. Since  $P_i^r/P_{i+1}^r$  is abelian, the proof will be complete once we show that all elements in  $\mathbf{T}$  have bounded finite order modulo  $P_{i+1}^r$ .

By Theorem 2.9, the word  $\gamma_r$  is linear in position  $i$  of the tuple  $\mathbf{N}_i$  modulo  $P_{i+1}^r$ . In particular,

$$(2.3) \quad [s_1, \dots, s_{i-1}, x_i, s_{i+1}, \dots, s_r]^{\lambda n_i} \equiv [s_1, \dots, s_{i-1}, x_i^{\lambda n_i}, s_{i+1}, \dots, s_r] \pmod{P_{i+1}^r},$$

for every  $\lambda \in \mathbb{Z}$ . Since  $x_i \in \gamma_i(N_1, \dots, N_i) \leq N_i$ , it follows from (ii) in the statement of the theorem that  $x_i^{\lambda n_i} \in S_i$  for all  $\lambda \in \mathbb{Z}$ . Thus we get

$$[s_1, \dots, s_{i-1}, x_i^{\lambda n_i}, s_{i+1}, \dots, s_r] \in \gamma_r\{\mathbf{S}\}.$$

Since  $\gamma_r\{\mathbf{S}\}$  is finite of order  $m$ , it follows that there exist  $\lambda, \mu \in \{0, \dots, m\}$ ,  $\lambda \neq \mu$ , such that

$$[s_1, \dots, s_{i-1}, x_i, s_{i+1}, \dots, s_r]^{\lambda n_i} \equiv [s_1, \dots, s_{i-1}, x_i, s_{i+1}, \dots, s_r]^{\mu n_i} \pmod{P_{i+1}^r}.$$



This implies that  $[s_1, \dots, s_{i-1}, x_i, s_{i+1}, \dots, s_r]$  has  $(m, n_i)$ -bounded finite order modulo  $P_{i+1}^r$ , as desired.  $\square$

If we take  $S_i = N_i$ , we get Theorem B for the lower central words.

**Corollary 2.12.** *Let  $r \in \mathbb{N}$  and let  $\mathbf{N} = (N_1, \dots, N_r)$  be a tuple of normal subgroups of a group  $G$ . If  $\gamma_r\{\mathbf{N}\}$  is finite of order  $m$ , then the subgroup  $\gamma_r(\mathbf{N})$  is also finite, of  $(m, r)$ -bounded order.*

Now we deduce Theorem A for lower central words.

**Corollary 2.13.** *Let  $r \in \mathbb{N}$  and let  $u_1, \dots, u_r$  be disjoint non-commutator words. Then the word  $\gamma_r(u_1, \dots, u_r)$  is boundedly concise. In particular,  $\gamma_r(x_1^{n_1}, \dots, x_r^{n_r})$  is boundedly concise for all  $n_i \in \mathbb{Z} \setminus \{0\}$ .*

*Proof.* Let us consider the word  $w = \gamma_r(u_1, \dots, u_r)$ , and let  $G$  be a group in which  $w$  takes finitely many values, say  $|w\{G\}| = m$ . By Corollary 2.3, we have  $w(G) = \gamma_r(u_1(G), \dots, u_r(G))$ . Note that  $u_i(G) = \langle S_i \rangle$ , where  $S_i = u_i\{G\}$ , and that  $w\{G\} = \gamma_r\{\mathbf{S}\}$ , where  $\mathbf{S} = (S_1, \dots, S_r)$ . Now observe that  $S_i$  is a normal subset of  $G$  and that, since  $u_i$  is a non-commutator word, for some  $n_i \in \mathbb{Z} \setminus \{0\}$  we have  $\{g^{n_i} \mid g \in G\} \subseteq u_i\{G\}$ . Hence  $w(G)$  is finite of  $(m, r, n_1, \dots, n_r)$ -bounded order by Theorem 2.11.  $\square$

### 3. Derived words

In this section, we prove Theorems A and B for derived words. The general strategy is still the same as for lower central words: we are going to obtain a suitable series of normal subgroups of  $G$ , going from  $[\delta_k(\mathbf{N}), \delta_k(\mathbf{N})]$  to  $\delta_k(\mathbf{N})$ , with the property that each of the factors of the series can be generated by a verbal subgroup on a tuple of normal subgroups that is closely related to  $\delta_k(\mathbf{N})$  and linear in one component. This is basically Theorem 3.5 below. For simplicity, let us refer to such a series as a linear series. The argument needed to obtain a linear series for derived words presents difficulties and subtleties that did not arise with lower central words, and is also significantly more technical. For the convenience of the reader and in order to make the procedure for a general  $\delta_k$  more understandable, first of all, we are going to provide a sketch of it in the particular case of  $\delta_2$ .

Of course,  $\delta_1 = \gamma_2$  and, according to Theorem 2.9, we have the following linear series for  $\delta_1(N_1, N_2)$ :

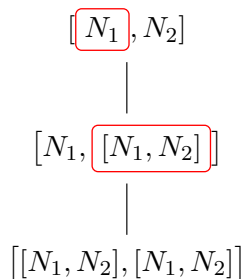


FIGURE 1. Series of  $[N_1, N_2]$

In this and in the next diagrams, a red box indicates the component in which we have linearity.

Let us see how we can construct a linear series for  $\delta_2(N_1, N_2, N_3, N_4)$  from the series above for  $\delta_1$ . To this purpose, we will use Lemma 2.8, which ensures that linearity is preserved after taking suitable

commutators, and also the remark made before that lemma, saying that linearity is preserved after multiplying by a normal subgroup. To start with, we take the commutator of the terms of the previous series with  $[N_3, N_4]$ , obtaining the series

$$\begin{array}{c} [[N_1, N_2], [N_3, N_4]] \\ | \\ [[N_1, [N_1, N_2]], [N_3, N_4]] \\ | \\ [[[N_1, N_2], [N_1, N_2]], [N_3, N_4]] \end{array}$$

Now we multiply this series with  $[[N_1, N_2], [[N_1, N_2], [N_3, N_4]]]$ , which contains  $[[N_1, N_2], [N_1, N_2], [N_3, N_4]]$  by P. Hall's Three Subgroup Lemma, and we get the following diagram:

$$\begin{array}{c} [[N_1, N_2], [N_3, N_4]] \\ | \\ [[N_1, [N_1, N_2]], [N_3, N_4]] \\ | \\ [[N_1, N_2], [[N_1, N_2], [N_3, N_4]]] \end{array}$$

FIGURE 2. First diagram for  $[[N_1, N_2], [N_3, N_4]]$

Here, and in the remaining diagrams, instead of the subgroups of the series, we show verbal subgroups on normal subgroups whose images coincide with the corresponding factors of the series. After all, it is in these subgroups where we are going to obtain the linearity conditions. Be aware then that vertical lines in the diagrams do not denote inclusions from now on.

By swapping the roles of  $(N_1, N_2)$  and  $(N_3, N_4)$ , we can obtain this other diagram:

$$\begin{array}{c} [[N_1, N_2], [N_3, N_4]] \\ | \\ [[N_1, N_2], [N_3, [N_3, N_4]]] \\ | \\ [[[N_1, N_2], [N_3, N_4]], [N_3, N_4]] \end{array}$$

FIGURE 3. Second diagram for  $[[N_1, N_2], [N_3, N_4]]$

Now we take the commutator of  $[N_1, N_2]$  with the terms of this last diagram, and we add the extra term  $\delta_2(N_1, N_2, N_3, N_4)'$  at the bottom:

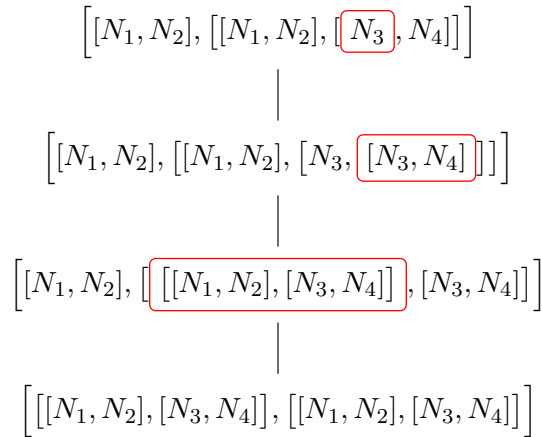


FIGURE 4. Series of  $\left[ [N_1, N_2], [ [N_1, N_2], [N_3, N_4] ] \right]$

Finally, by gluing the diagrams in Figures 2 and 4 together, we obtain a linear series for the subgroup  $\delta_2(N_1, N_2, N_3, N_4)$ .

Of course, this is simply a sketch without proofs, but we follow the same procedure in the proof of Theorem 3.5, to get a linear series for  $\delta_k$  from a series for  $\delta_{k-1}$ . At this point, it is worth noting an important difference with the situation for a lower central word  $\gamma_r$ . In that case, every factor of the linear series is of the following form (again we show the linear component in red):

$$[N_1, \dots, N_{i-1}, [N_1, \dots, N_i], N_{i+1}, \dots, N_r].$$

We observe that this subgroup is of the form  $\gamma_r(\mathbf{M})$ , where the  $j$ th component  $M_j$  of  $\mathbf{M}$  is either  $N_j$  or a commutator of the terms of  $\mathbf{N}$  that involves  $N_j$ , and the linearity happens in  $M_i$ . However, if we look at the series for  $\delta_2$  obtained above, the first two subgroups in Figure 4 are

$$(3.1) \quad \delta_2(N_1; N_2; [N_1, N_2]; [N_3, N_4])$$

and

$$(3.2) \quad \delta_2(N_1; N_2; [N_1, N_2]; [N_3, [N_3, N_4]]),$$

which are not of the form  $\delta_2(\mathbf{M})$  with every  $M_j$  a commutator from  $\mathbf{N}$  involving  $N_j$ , as we can see by looking at the third component of  $\delta_2$ . Also, the linearity does not happen in a component of  $\delta_2$ , but in a more interior position. Nevertheless, we can write these subgroups as verbal subgroups on normal subgroups for outer commutator words different from  $\delta_2$ . More specifically, if

$$v(x_1, x_2, x_3, x_4, y_1, y_2) = [x_1, x_2], [[y_1, y_2], [x_3, x_4]]$$

then the subgroups in (3.1) and (3.2) are  $v(\mathbf{M}_1)$  and  $v(\mathbf{M}_2)$ , where

$$(3.3) \quad \mathbf{M}_1 = (N_1, N_2, [N_3, N_4], N_4, N_1, N_2) \quad \text{and} \quad \mathbf{M}_2 = (N_1, N_2, N_3, [N_3, N_4], N_1, N_2),$$

where we have marked the linear components in red, again.

After illustrating the procedure with the case of  $\delta_2$ , let us proceed to systematically develop the tools that are necessary for the proof of Theorem 3.5.

We start by introducing a special type of words that we can derive from a given outer commutator word  $w$ , which we call extended words of  $w$ . Before giving the definition, we show the idea behind extended words with an example. Consider the word  $\delta_2 = [[x_1, x_2], [x_3, x_4]]$ . This is formed by taking the commutator of  $x_1$  and  $x_2$ , taking the commutator of  $x_3$  and  $x_4$ , and then taking the commutator of these two commutators. Now suppose that on some occasions, before performing one of these commutators, we introduce a change by taking first the commutator of one (or both) of the components with an outer commutator word not involving the variables  $x_1, \dots, x_4$  appearing in  $\delta_2$ . For example, before producing  $[x_1, x_2]$ , we take the commutator  $[[y_1, y_2], x_1]$  and now we follow as in  $\delta_2$  taking the commutator with  $x_2$ , obtaining  $[[[y_1, y_2], x_1], x_2]$ . We could continue with the process of taking commutators without making any other changes, so we get

$$[[[y_1, y_2], x_1], x_2], [x_3, x_4],$$

but we could also make some similar changes in the process, as in the words

$$[[[y_1, y_2], x_1], x_2], [x_3, [y_3, x_4]]$$

and

$$[[[y_1, y_2], x_1], x_2], [[x_3, [y_3, x_4]], y_4].$$

Another possibility is to make a commutator at the very end, after having completed  $\delta_2$ , as in

$$[y_1, [[x_1, x_2], [x_3, x_4]]].$$

Observe that all these extended words are again outer commutator words, because we never repeat a variable when we make changes in the construction of  $\delta_2$ .

Let us now give the formal definition of extended words. Notice that this definition differs from the one of extensions of outer commutator words given in [8, Definition 3.1].

**Definition 3.1.** Let  $w = w(x_1, \dots, x_r)$  be an outer commutator, and let  $Y = \{y_n\}_{n \in \mathbb{N}}$  be a set of variables that are disjoint from  $X$ . For every  $k \in \mathbb{N} \cup \{0\}$ , we define recursively the set  $\text{ext}_k(w)$  of  $k$ th extended words of  $w$  as follows:

- (i)  $\text{ext}_0(w) = \{w\}$ .
- (ii) For  $k \geq 1$ ,  $\text{ext}_k(w)$  consists of the set

$$\begin{aligned} & \{[p, q], [q, p] \mid p \text{ is an outer commutator in } Y, q \in \text{ext}_{k-1}(w), p \text{ and } q \text{ disjoint}\} \\ & = \{[p, q] \mid p \text{ is an outer commutator in } Y, q \in \text{ext}_{k-1}(w), p \text{ and } q \text{ disjoint}\}^{\pm 1}, \end{aligned}$$

and, if  $w = [\alpha, \beta]$ , also of the set

$$\bigcup_{\ell+m=k} \{[p, q] \mid p \in \text{ext}_\ell(\alpha), q \in \text{ext}_m(\beta), p \text{ and } q \text{ disjoint}\}.$$

If  $v \in \text{ext}_k(w)$ , then we say that  $w$  is an extended word of degree  $k$  of  $w$  by outer commutators.

For brevity, in the remainder we will simply speak of extended words when we mean extended words by outer commutators. Observe that an extended word  $v$  of an outer commutator  $w = w(x_1, \dots, x_r)$  is again an outer commutator, in the variables  $\{x_1, \dots, x_r\} \cup Y$ . Whenever it is convenient we will assume, after renaming the variables, that  $v = v(x_1, \dots, x_r, y_{r+1}, \dots, y_s)$ .

Next we generalize Lemma 2.6 to extended words of an outer commutator word.

**Lemma 3.2.** *Let  $v = v(x_1, \dots, x_r, y_{r+1}, \dots, y_s)$  be an extended word of degree  $k$  of an outer commutator word  $w = w(x_1, \dots, x_r)$ . Assume that  $\mathbf{S} = (S_1, \dots, S_r)$  is a tuple of normal subsets of a group  $G$ . If  $\mathbf{t} = (t_1, \dots, t_s)$  is a tuple of elements of  $G$ , such that  $t_i \in S_i^{*m_i}$  for every  $i = 1, \dots, r$ , then*

$$v(\mathbf{t}) \in w\{\mathbf{S}\}^{*m_1 \dots m_r 2^k}.$$

*Proof.* We use induction on  $k + r$ . If  $k = 0$  then  $v = w$  and the result is Lemma 2.6. This gives in particular the basis of the induction. Suppose now that the result holds for smaller values of  $k + r$ , and that  $k \geq 1$ . According to Definition 3.1, we may assume that  $v(\mathbf{t}) = [p(\mathbf{t}'), q(\mathbf{t}'')]$ , where  $p$  and  $q$  are disjoint and

- (i) either  $p$  is an outer commutator word in  $Y$  and  $q \in \text{ext}_{k-1}(w)$ ,
- (ii) or  $p \in \text{ext}_\ell(\alpha)$ ,  $q \in \text{ext}_m(\beta)$ , with  $w = [\alpha, \beta]$  and  $\ell + m = k$ .

In case (i), all elements  $t_1, \dots, t_r$  appear in the vector  $\mathbf{t}''$ , and by the induction hypothesis we have  $q(\mathbf{t}'') \in w\{\mathbf{S}\}^{*m_1 \dots m_r 2^{k-1}}$ . Then the result follows by applying Lemma 2.5 to the commutator word  $[x_1, x_2]$  and the normal subset  $w\{\mathbf{S}\}^{*m_1 \dots m_r 2^{k-1}}$ .

Suppose now that we are in case (ii), and assume without loss of generality that  $\alpha = \alpha(x_1, \dots, x_q)$  and  $\beta = \beta(x_{q+1}, \dots, x_r)$ . Set  $\mathbf{S}' = (S_1, \dots, S_q)$  and  $\mathbf{S}'' = (S_{q+1}, \dots, S_r)$ . Since  $\alpha$  and  $\beta$  involve less variables than  $w$ , the result is true for  $p$  and  $q$ , and so

$$p(\mathbf{t}') \in \alpha\{\mathbf{S}'\}^{*m_1 \dots m_q 2^\ell} \quad \text{and} \quad q(\mathbf{t}'') \in \beta\{\mathbf{S}''\}^{*m_{q+1} \dots m_r 2^m}.$$

Now the result follows by applying Lemma 2.6 to the commutator word  $[x_1, x_2]$  and the pair of normal subsets  $(\alpha\{\mathbf{S}'\}, \beta\{\mathbf{S}''\})$ . □

We also need to define a type of extensions of tuples of normal subgroups and of verbal subgroups on normal subgroups. The idea behind the definition is to be able to deal with tuples like the ones appearing in (3.3) and with the corresponding verbal subgroups on normal subgroups in that paragraph.

**Definition 3.3.** *Let  $G$  be a group and let  $\mathbf{N} = (N_1, \dots, N_r)$  and  $\mathbf{M} = (M_1, \dots, M_s)$  be two tuples of normal subgroups of  $G$ . We say that  $\mathbf{M}$  is an outer commutator extension of  $\mathbf{N}$ , if the following conditions hold:*

- (i)  $s \geq r$ ;
- (ii) For every  $i = 1, \dots, s$ , we have  $M_i = w_i(\mathbf{N}_i)$ , where  $w_i$  is an outer commutator word and all components of  $\mathbf{N}_i$  belong to  $\mathbf{N}$ ;
- (iii) For every  $i = 1, \dots, r$ , the subgroup  $N_i$  is a component of  $\mathbf{N}_i$ , and consequently  $M_i \leq N_i$ .

**Definition 3.4.** Let  $w = w(x_1, \dots, x_r)$  be a word and let  $\mathbf{N}$  be an  $r$ -tuple of normal subgroups of a group  $G$ . An extension of degree  $k$  of  $w(\mathbf{N})$  by outer commutators is a subgroup of the form  $v(\mathbf{M})$ , where  $v$  is an extended word of degree  $k$  of  $w$  and  $\mathbf{M}$  is an outer commutator extension of  $\mathbf{N}$ .

For example, the subgroup in (3.2) can be seen as an extension of  $\delta_2(N_1, N_2, N_3, N_4)$ , by taking  $v = [[x_1, x_2], [[y_1, y_2], [x_3, x_4]]]$  and  $\mathbf{M} = (N_1, N_2, N_3, [N_3, N_4], N_1, N_2)$ . Note that  $v(\mathbf{M})$  is linear in the fourth component modulo the subgroup that appears below it in Figure 4.

We now prove the existence of a linear series for derived words.

**Theorem 3.5.** Let  $k \in \mathbb{N}$  and let  $\mathbf{N} = (N_1, \dots, N_{2^k})$  be a tuple of normal subgroups of a group  $G$ . Then there exists a series

$$[\delta_k(\mathbf{N}), \delta_k(\mathbf{N})] = V_0^k \leq V_1^k \leq \dots \leq V_t^k = \delta_k(\mathbf{N})$$

of normal subgroups of  $G$  of length  $t = 2^k + 2^{k-1} - 1$  such that, for every  $i = 1, \dots, t$ , the following hold:

- (i) The section  $V_i^k/V_{i-1}^k$  is the image of an extension  $v_i^k(\mathbf{M}_i^k)$  of  $\delta_k(\mathbf{N})$  of degree at most  $k - 1$ ;
- (ii) In the section  $V_i^k/V_{i-1}^k$ , the word  $v_i^k$  is linear in one component of the tuple  $\mathbf{M}_i^k$ .

Furthermore, the words  $v_i^k$  and the words appearing in the outer commutator extensions  $\mathbf{M}_i^k$  depend only on  $k$ , and not on the group  $G$  or on the tuple  $\mathbf{N}$ .

*Proof.* We prove the theorem by induction on  $k$ . If  $k = 1$  then, as already mentioned, it suffices to take the series

$$V_0^1 = [[N_1, N_2], [N_1, N_2]] \leq V_1^1 = [N_1, [N_1, N_2]] \leq V_2^1 = [N_1, N_2],$$

so that  $v_1^1 = v_2^1 = \delta_1$ , and  $M_1^1 = (N_1, [N_1, N_2])$  and  $M_2^1 = (N_1, N_2)$ .

Assume now that  $k \geq 2$  and that the result is true for  $k - 1$ . Set  $\mathbf{N}_1 = (N_1, \dots, N_{2^{k-1}})$  and  $\mathbf{N}_2 = (N_{2^{k-1}+1}, \dots, N_{2^k})$ . By the induction hypothesis, if we put  $s = 2^{k-1} + 2^{k-2} - 1$ , then there exist series

$$(3.4) \quad V_{0,1}^{k-1} = [\delta_{k-1}(\mathbf{N}_1), \delta_{k-1}(\mathbf{N}_1)] \leq \dots \leq V_{i,1}^{k-1} \leq \dots \leq V_{s,1}^{k-1} = \delta_{k-1}(\mathbf{N}_1)$$

and

$$(3.5) \quad V_{0,2}^{k-1} = [\delta_{k-1}(\mathbf{N}_2), \delta_{k-1}(\mathbf{N}_2)] \leq \dots \leq V_{i,2}^{k-1} \leq \dots \leq V_{s,2}^{k-1} = \delta_{k-1}(\mathbf{N}_2)$$

such that, for every  $i = 1, \dots, s$ , the factors  $V_{i,1}^{k-1}/V_{i-1,1}^{k-1}$  and  $V_{i,2}^{k-1}/V_{i-1,2}^{k-1}$  are the images of  $v_i^{k-1}(\mathbf{M}_{i,1}^{k-1})$  and  $v_i^{k-1}(\mathbf{M}_{i,2}^{k-1})$ , respectively, where:

- (a)  $v_i^{k-1}$  is an extended word of  $\delta_{k-1}$  of degree at most  $k - 2$ ;
- (b)  $\mathbf{M}_{i,1}^{k-1}$  is an outer commutator extension of  $\mathbf{N}_1$ ;
- (c)  $\mathbf{M}_{i,2}^{k-1}$  is an outer commutator extension of  $\mathbf{N}_2$ ;
- (d) In the sections  $V_{i,1}^{k-1}/V_{i-1,1}^{k-1}$  and  $V_{i,2}^{k-1}/V_{i-1,2}^{k-1}$ , the word  $v_i^{k-1}$  is linear in one component of the tuples  $\mathbf{M}_{i,1}^{k-1}$  and  $\mathbf{M}_{i,2}^{k-1}$ , respectively.

Let us now see how to obtain the series for  $k$  and for the tuple  $\mathbf{N}$  from the two series (3.4) and (3.5). We start by taking the commutator of all terms of the series (3.4) with  $\delta_{k-1}(\mathbf{N}_2)$ . This way we obtain the series

$$(3.6) \quad [V_{0,1}^{k-1}, \delta_{k-1}(\mathbf{N}_2)] \leq \dots \leq [V_{i,1}^{k-1}, \delta_{k-1}(\mathbf{N}_2)] \leq \dots \leq [\delta_{k-1}(\mathbf{N}_1), \delta_{k-1}(\mathbf{N}_2)] = \delta_k(\mathbf{N}).$$

By P. Hall’s Three Subgroup Lemma, we have

$$\begin{aligned} [V_{0,1}^{k-1}, \delta_{k-1}(\mathbf{N}_2)] &= [\delta_{k-1}(\mathbf{N}_1), \delta_{k-1}(\mathbf{N}_1), \delta_{k-1}(\mathbf{N}_2)] \\ &\leq [\delta_{k-1}(\mathbf{N}_1), \delta_{k-1}(\mathbf{N}_2), \delta_{k-1}(\mathbf{N}_1)] = [\delta_{k-1}(\mathbf{N}_1), \delta_k(\mathbf{N})]. \end{aligned}$$

Now we multiply all terms of the series (3.6) by  $[\delta_{k-1}(\mathbf{N}_1), \delta_k(\mathbf{N})]$ , and this is the rightmost part of the series we seek (where  $t = 2^k + 2^{k-1} - 1$ , as above):

$$(3.7) \quad V_{t-s}^k = [\delta_{k-1}(\mathbf{N}_1), \delta_k(\mathbf{N})] \leq \dots \leq V_{t-s+i}^k = [V_{i,1}^{k-1}, \delta_{k-1}(\mathbf{N}_2)] [\delta_{k-1}(\mathbf{N}_1), \delta_k(\mathbf{N})] \leq \dots \leq V_t^k = \delta_k(\mathbf{N}).$$

Note that  $t - s = 2^{k-1} + 2^{k-2}$ . The factors in this series are the images of the subgroups

$$[v_i^{k-1}(\mathbf{M}_{i,1}^{k-1}), \delta_{k-1}(\mathbf{N}_2)],$$

which can be represented in the form  $v_i^k(\mathbf{M}_i^k)$  by taking

$$v_i^k = [v_i^{k-1}, \delta_{k-1}(x_{2^{k-1}+1}, \dots, x_{2^k})]$$

and defining  $\mathbf{M}_i^k$  to be the concatenation of  $\mathbf{M}_{i,1}^{k-1}$  and  $\mathbf{N}_2$ , where the elements of  $\mathbf{N}_2$  occupy the positions  $2^{k-1} + 1, \dots, 2^k$  (which are the positions corresponding to the variables  $x_{2^{k-1}+1}, \dots, x_{2^k}$ ). Note that  $\mathbf{M}_i^k$  is an outer commutator extension of  $\mathbf{N}$ .

In a symmetric way, by first taking the commutator of  $\delta_{k-1}(\mathbf{N}_1)$  with all terms of the series (3.5) and then multiplying by  $[\delta_k(\mathbf{N}), \delta_{k-1}(\mathbf{N}_2)]$ , we get the series

$$(3.8) \quad W_{t-s}^k = [\delta_k(\mathbf{N}), \delta_{k-1}(\mathbf{N}_2)] \leq \dots \leq W_{t-s+i}^k = [\delta_{k-1}(\mathbf{N}_1), V_{i,2}^{k-1}] [\delta_k(\mathbf{N}), \delta_{k-1}(\mathbf{N}_2)] \leq \dots \leq W_t^k = \delta_k(\mathbf{N}).$$

In this series, the factors are given by the images of the subgroups  $w_i^k(\mathbf{L}_i^k)$ , where

$$w_i^k = [\delta_{k-1}(y_{2^{k-1}+1}, \dots, y_{2^k}), v_{i,2}^{k-1}],$$

$v_{i,2}^{k-1}$  being the same word as  $v_i^{k-1}$ , with  $x_1, \dots, x_{2^{k-1}}$  replaced with  $x_{2^{k-1}+1}, \dots, x_{2^k}$ , and  $\mathbf{L}_i^k$  being the concatenation of  $\mathbf{M}_{i,2}^{k-1}$  and  $\mathbf{N}_1$ , where we put the components of the second tuple after the components of the first. Note that  $w_i^k$  is an extended word of  $\delta_{k-1}(x_{2^{k-1}+1}, \dots, x_{2^k})$  of degree at most  $k - 1$ .

Now we take the commutator of  $\delta_{k-1}(\mathbf{N}_1)$  with the terms of the last series, and subtract  $s$  to all indices, getting

$$(3.9) \quad Z_1^k = [\delta_{k-1}(\mathbf{N}_1), [\delta_k(\mathbf{N}), \delta_{k-1}(\mathbf{N}_2)]] \leq \dots \leq Z_{i+1}^k = [\delta_{k-1}(\mathbf{N}_1), W_{t-s+i}^k] \leq \dots \leq Z_{t-s}^k = [\delta_{k-1}(\mathbf{N}_1), \delta_k(\mathbf{N})],$$

since  $t - 2s = 1$ . Finally, we define  $V_0^k = [\delta_k(\mathbf{N}), \delta_k(\mathbf{N})]$  and multiply all terms of (3.9) with this subgroup, setting  $V_i^k = Z_i^k V_0^k$  for  $i = 1, \dots, t - s$ . Since  $V_0^k \leq [\delta_{k-1}(\mathbf{N}_1), \delta_k(\mathbf{N})]$ , we get the series

$$(3.10) \quad V_0^k = [\delta_k(\mathbf{N}), \delta_k(\mathbf{N})] \leq \dots \leq V_{i+1}^k = Z_{i+1}^k [\delta_k(\mathbf{N}), \delta_k(\mathbf{N})] \\ \leq \dots \leq V_{t-s}^k = [\delta_{k-1}(\mathbf{N}_1), \delta_k(\mathbf{N})].$$

In this series, the factors  $V_i^k/V_{i-1}^k$  for  $i = 2, \dots, t - s$  are given by the images of the subgroups  $v_i^k(\mathbf{M}_i^k)$ , where

$$v_i^k = [\delta_{k-1}(x_1, \dots, x_{2^{k-1}}), w_{i+s}^k]$$

is an extended word of  $\delta_k$  of degree at most  $k - 1$ , and  $\mathbf{M}_i^k$  is the concatenation of  $\mathbf{N}_1$  and  $\mathbf{L}_{i+s}^k$ , with the components of the second tuple after the components of the first. On the other hand, the quotient  $V_1^k/V_0^k$  is given by the image of  $v_1^k(\mathbf{M}_1^k)$ , where

$$v_1^k = [\delta_{k-1}(x_1, \dots, x_{2^{k-1}}), [y, \delta_{k-1}(x_{2^{k-1}+1}, \dots, x_{2^k})]]$$

and  $\mathbf{M}_1^k = (\mathbf{N}, \delta_k(\mathbf{N}))$ .

Now the concatenation of (3.7) and (3.10) is the desired series for  $k$  and  $\mathbf{N}$ . The discussion of the previous paragraphs shows that  $v_i^k(\mathbf{M}_i^k)$  is an extension of  $\delta_k(\mathbf{N})$  for every  $i = 1, \dots, t$ . Thus we only need to check the linearity of every word  $v_i^k$  in one component of the vector  $\mathbf{M}_i^k$ . For  $i = t - s + 1, \dots, t$ , if  $v_i^{k-1}$  is linear in component  $j$  of  $\mathbf{M}_{i,1}^{k-1}$  of the initial series (3.4), then combining this fact with Lemma 2.8, it follows that  $v_i^k$  is linear in the same component of  $\mathbf{M}_i^k$ . For  $i = 1, \dots, t - s$ , we can use similarly the linearity of the series (3.5). Finally for  $i = 1$ , since  $V_0^k = [\delta_k(\mathbf{N}), \delta_k(\mathbf{N})]$ , we have linearity in the component corresponding to  $y$ , which takes values in  $\delta_k(\mathbf{N})$ . □

We can now prove the corresponding version of Theorem 2.11 for the words  $\delta_k$ .

**Theorem 3.6.** *Let  $k \in \mathbb{N}$  and let  $\mathbf{N} = (N_1, \dots, N_{2^k})$  be a tuple of normal subgroups of a group  $G$ . Assume that  $N_i = \langle S_i \rangle$  for every  $i = 1, \dots, 2^k$ , where:*

- (i)  $S_i$  is a normal subset of  $G$ ;
- (ii) There exists  $n_i \in \mathbb{N}$  such that all  $n_i$ th powers of elements of  $N_i$  are contained in  $S_i$ .

*If for the tuple  $\mathbf{S} = (S_1, \dots, S_{2^k})$  the set of values  $\delta_k\{\mathbf{S}\}$  is finite of order  $m$ , then the subgroup  $\delta_k(\mathbf{N})$  is also finite and of  $(m, k, n_1, \dots, n_{2^k})$ -bounded order.*

*Proof.* Let us consider the series

$$[\delta_k(\mathbf{N}), \delta_k(\mathbf{N})] = V_0^k \leq V_1^k \leq \dots \leq V_t^k = \delta_k(\mathbf{N}),$$

of Theorem 3.5, where  $t = 2^k + 2^{k-1} - 1$ . We prove that every  $V_i^k$  is finite of bounded order by induction on  $i$ . The result for  $i = 0$  follows from Lemma 2.10. Assume now that  $i \geq 1$  and that the result is true for  $V_{i-1}^k$ . By Theorem 3.5, the section  $V_i^k/V_{i-1}^k$  coincides with the image of a subgroup  $v_i^k(\mathbf{M}_i^k)$ , that is an extension of  $\delta_k(\mathbf{N})$  of degree at most  $k - 1$ .

For simplicity, let us write  $v$  for  $v_i^k$  and  $\mathbf{M}$  for  $\mathbf{M}_i^k$ . Let  $\mathbf{M} = (M_1, \dots, M_s)$ , which is an outer commutator extension of  $\mathbf{N}$ . Hence  $s \geq 2^k$  and for every  $j = 1, \dots, s$  we have  $M_j = w_j(\mathbf{N}_j)$ , where  $w_j$



is an outer commutator word, all components in  $\mathbf{N}_j$  belong to  $\mathbf{N}$ , and one of these components must be  $N_j$  for  $j = 1, \dots, 2^k$ .

Let  $T_j = w_j\{\mathbf{S}_j\}$ , where  $\mathbf{S}_j$  is obtained from  $\mathbf{N}_j$  by replacing each subgroup  $N_\ell$  with its given generating set  $S_\ell$ . Hence  $T_j \subseteq M_j$ . Recall from Theorem 3.5 that the word  $v$  (and hence also the number  $s$  of variables of  $v$ ) and the words  $w_1, \dots, w_s$  only depend on  $k$ , and not on  $G$  or on  $\mathbf{N}$ . From this fact, and since  $\mathbf{S}_j$  consists of normal subsets of  $G$ , it follows from Lemma 2.5 that there is a function  $h : \mathbb{N} \rightarrow \mathbb{N}$  such that  $T_j \subseteq S_j^{*h(k)}$  for every  $j = 1, \dots, 2^k$ . If we set  $\mathbf{T} = (T_1, \dots, T_s)$  then, by Lemma 3.2, we get  $v\{\mathbf{T}\} \subseteq w\{\mathbf{S}\}^{*n}$ , where

$$n = h(k)^{2^k} 2^{k-1}.$$

Consequently

$$(3.11) \quad |v\{\mathbf{T}\}| \leq (2m + 1)^n,$$

and  $v\{\mathbf{T}\}$  is finite of  $(m, k)$ -bounded cardinality. On the other hand, it follows from Lemma 2.1 that  $v(\mathbf{M})$  can be generated by the set of values  $v\{\mathbf{T}\}$ . From Theorem 3.5, we know that the word  $v$  is linear in some position  $j \in \{1, \dots, s\}$  of the tuple  $\mathbf{M}$  modulo  $V_{i-1}^k$ . Since  $M_j = w_j(\mathbf{N}_j)$  is as above, we have  $M_j \leq N_\ell$  for some  $\ell \in \{1, \dots, 2^k\}$ , and actually  $\ell = j$  if  $j \in \{1, \dots, 2^k\}$ . Now, from linearity, for every tuple  $\mathbf{t} = (t_1, \dots, t_s) \in \mathbf{T}$  and every  $\lambda \in \mathbb{Z}$ , we have

$$(3.12) \quad v(\mathbf{t})^{\lambda n_\ell} = v(t_1, \dots, t_j, \dots, t_s)^{\lambda n_\ell} \equiv v(t_1, \dots, t_j^{\lambda n_\ell}, \dots, t_s) \pmod{V_{i-1}^k}.$$

We have  $t_j^\lambda \in M_j \leq N_\ell$  and then by (ii) in the statement of the theorem,  $t_j^{\lambda n_\ell} \in S_\ell$ . So we get

$$v(t_1, \dots, t_j^{\lambda n_\ell}, \dots, t_s) \in v\{\mathbf{T}_j\},$$

where  $\mathbf{T}_j$  is the tuple obtained from  $\mathbf{T}$  after replacing  $T_j$  with  $S_\ell$ . Similarly to (3.11) and taking into account that  $\ell = j$  if  $j \in \{1, \dots, 2^k\}$ , it follows that the set  $v\{\mathbf{T}_j\}$  is finite of  $(m, k)$ -bounded cardinality. Hence there exist  $(m, k)$ -bounded integers  $\lambda$  and  $\mu$ , with  $\lambda \neq \mu$ , such that

$$v(\mathbf{t})^{\lambda n_\ell} \equiv v(\mathbf{t})^{\mu n_\ell} \pmod{V_{i-1}^k}.$$

This implies that  $v(\mathbf{t})$  has finite order modulo  $V_{i-1}^k$ , bounded in terms of  $m, k$  and  $n_\ell$ .

Summarizing, the abelian quotient  $V_i^k/V_{i-1}^{k-1}$  is the image of the verbal subgroup  $v(\mathbf{M})$ , which is generated by the set  $v\{\mathbf{T}\}$  of  $(m, k)$ -bounded cardinality, and each element of  $v\{\mathbf{T}\}$  has  $(m, k, n_\ell)$ -bounded order. We conclude that the order of  $V_i^k/V_{i-1}^k$  is  $(m, k, n_\ell)$ -bounded, which completes the proof. □

Exactly as in the case of lower central words, we obtain Theorems A and B as special cases of this last result.

**Corollary 3.7.** *Let  $k \in \mathbb{N}$  and let  $u_1, \dots, u_{2^k}$  be non-commutator words. Then the word  $\delta_k(u_1, \dots, u_{2^k})$  is boundedly concise. In particular,  $\delta_k(x_1^{n_1}, \dots, x_{2^k}^{n_{2^k}})$  is boundedly concise for all  $n_i \in \mathbb{Z} \setminus \{0\}$ .*

**Corollary 3.8.** *Let  $k \in \mathbb{N}$  and let  $\mathbf{N} = (N_1, \dots, N_{2^k})$  be a tuple of normal subgroups of a group  $G$ . If  $\delta_k\{\mathbf{N}\}$  is finite of order  $m$ , then the subgroup  $\delta_k(\mathbf{N})$  is also finite, of  $(m, k)$ -bounded order.*

### Acknowledgements

Supported by the Spanish Government, grant PID2020-117281GB-I00, partly with FEDER funds, and by the Basque Government, grant IT483-22. The second author is also supported by a grant FPI-2018 of the Spanish Government.

The authors thank the anonymous referee for helpful suggestions that allowed them to produce an improved version of the paper.

### REFERENCES

- [1] A. Abdollahi and F. G. Russo, On a problem of P. Hall for Engel words, *Arch. Math.* **97** (2011), no. 5, 407–412.
- [2] J. Azevedo and P. Shumyatsky, On finiteness of verbal subgroups, *J. Pure Appl. Algebra*, **226** (2022), no. 6, Paper No. 106976, 11 pages.
- [3] R. Baer, Endlichkeitskriterien für Kommutatorgruppen, *Math. Ann.* **124** (1952) 161–177.
- [4] C. Delizia, P. Shumyatsky, A. Tortora and M. Tota, On conciseness of some commutator words, *Arch. Math. (Basel)* **112** (2019), no. 1, 27–32.
- [5] E. Detomi, M. Morigi and P. Shumyatsky, Words of Engel type are concise in residually finite groups, *Bull. Math. Sci.* **09** (2019), no. 2, 1950012, 19 pages.
- [6] G.A. Fernández-Alcober and M. Morigi, Outer commutator words are uniformly concise, *J. London Math. Soc. (2)* **82** (2010), no. 3, 581–595.
- [7] G. A. Fernández-Alcober, M. Morigi and G. Traustason, A note on conciseness of Engel words, *Comm. Algebra* **40** (2012), no. 7, 2570–2576.
- [8] G. A. Fernández-Alcober and P. Shumyatsky, On bounded conciseness of words in residually finite groups, *J. Algebra* **500** (2018) 19–29.
- [9] I. de las Heras and M. Morigi, Lower central words in finite  $p$ -groups, *Publ. Mat.* **65** (2021), no. 1, 243–269.
- [10] S. V. Ivanov, P. Hall’s conjecture on the finiteness of verbal subgroups, *Izv. Vyssh. Uchebn. Zaved.* **325** (1989), no. 6, 60–70.
- [11] A. Yu. Olshanskii, Geometry of defining relations in groups, Kluwer Academic Publishers Group, Dordrecht, 1991.
- [12] D. J. S. Robinson, Finiteness Conditions and Generalized Soluble Groups, Part 1, Springer-Verlag, New York-Berlin, 1972.
- [13] D. Segal, Words: notes on verbal width in groups, Cambridge University Press, Cambridge, 2009.
- [14] R. F. Turner-Smith, Finiteness conditions for verbal subgroups, *J. London Math. Soc.* **41** (1966) 166–176.
- [15] J. Wilson, On outer-commutator words, *Canad. J. Math.* **26** (1974) 608–620.

#### Gustavo A. Fernández-Alcober

Department of Mathematics, University of the Basque Country UPV/EHU, Bilbao, Spain.

Email: [gustavo.fernandez@ehu.eus](mailto:gustavo.fernandez@ehu.eus)

#### Matteo Pintonello

Department of Mathematics, University of the Basque Country UPV/EHU, Bilbao, Spain.

Email: [matteo.pintonello@ehu.eus](mailto:matteo.pintonello@ehu.eus)