ON THE RANK OF A VERBAL SUBGROUP OF A FINITE GROUP

ELOISA DETOMI, MARTA MORIGI⬛ and PAVEL SHUMYATSKY

Communicated by Ben Martin

Abstract

We show that if w is a multilinear commutator word and G a finite group in which every metanilpotent subgroup generated by w-values is of rank at most r, then the rank of the verbal subgroup $w(G)$ is bounded in terms of r and w only. In the case where G is soluble, we obtain a better result: if G is a finite soluble group in which every nilpotent subgroup generated by w-values is of rank at most r, then the rank of $w(G)$ is at most $r + 1$.

1. Introduction

Guralnick [11] and Lucchini [17] independently proved that if all Sylow subgroups of a finite group G can be generated by d elements, then the group G itself can be generated by $d + 1$ elements. This was an improvement over an earlier result due to Longobardi and Maj [15] giving the bound $2d$.

It follows that if all nilpotent subgroups of a finite group G have rank at most r, then the group G has rank at most $r + 1$. Here, the rank of a finite group is the minimum number r such that every subgroup can be generated by r elements.

In the present paper, we are concerned with the question of whether the rank of a verbal subgroup $w(G)$ can be bounded in terms of ranks of nilpotent subgroups generated by w-values. Recall that, given a group G and a group-word w, the verbal subgroup $w(G)$ is the one generated by the set G_w of all w-values in G. In general, elements of $w(G)$ are not w-values but there are many results showing that the set G_w has a strong influence on the structure of G (see, for example, [19]). Thus, the main theme of this article is as follows.

145

QUESTION 1.1. *Let w be a group-word and let G be a finite group in which every nilpotent subgroup generated by w-values has rank at most r. Is the rank of the verbal subgroup $w(G)$ bounded in terms of r and w only?*

In the case where $w = x^n$ is the power word for $n \geq 2$, the answer to the above question is negative. Indeed, let $G = FK$ be a Frobenius group with kernel F and cyclic complement K satisfying the conditions that $F^n = 1$ and $K^n = K$. It is straightforward that $G^n = G$ while any nilpotent subgroup generated by nth powers is cyclic. However, F can be chosen to be of arbitrarily large rank. Hence the rank of G cannot be bounded in terms of n.

In view of this example, we focus on commutator words. Similar issues with respect to the exponent of a verbal subgroup of a finite group were addressed in [20] where, in particular, it was proved that if w is a multilinear commutator word and G a finite group in which the exponent of any nilpotent subgroup generated by w-values divides some fixed number e, then the exponent of $w(G)$ is bounded in terms of e and w only. Later, this was extended in [3] to some words that are not necessarily multilinear. We recall that a finite group G is said to have exponent e if e is the least positive number such that $g^e = 1$ for each $g \in G$. Multilinear commutator words, also known as outer commutator words, are obtained by nesting commutators, but using always different indeterminates. For example, the word $[[x_1, x_2], [x_3, x_4, x_5], x_6]$ is a multilinear commutator word.

The question about the rank of a verbal subgroup is more complex than that of the exponent. The main catch is that the condition that every nilpotent subgroup generated by w-values has rank at most r may fail in homomorphic images of G. The corresponding condition on the exponent was shown in [20] to survive under homomorphisms.

In this paper, we find a way around that obstacle in the case of soluble groups. Our first result is the following theorem.

THEOREM 1.2. *Let w be a multilinear commutator word and let G be a finite soluble or finite simple group in which every nilpotent subgroup generated by w-values has rank at most r. Then the rank of the verbal subgroup $w(G)$ is at most $r + 1$.*

Theorem 1.2 is almost a straightforward consequence of the verification of the Ore conjecture [14] and the following proposition, which is a generalisation of the focal subgroup theorem to multilinear commutator words in soluble finite groups. This might be of independent interest (see Section 3).

PROPOSITION 1.3. *Let w be a multilinear commutator word and let G be a finite soluble group. Let P be a Sylow p-subgroup of G. Then $P \cap w(G)$ can be generated by w-values lying in P.*

It remains unknown whether the above proposition can be extended to arbitrary finite groups. The main result of [1] says that if w is a multilinear commutator word and G is a finite group, then $P \cap w(G)$ can be generated by powers of w-values. We also mention that, for derived words, the proposition was established in [2].

In the general case, that is, where G is not assumed to be either soluble or simple, we are able to answer only a weaker version of the main question: namely, the condition on the rank is imposed on the metanilpotent subgroups generated by w-values rather than the nilpotent ones.

THEOREM 1.4. *Let w be a multilinear commutator word and let G be a finite group in which every metanilpotent subgroup generated by w-values has rank at most r. Then the rank of the verbal subgroup $w(G)$ is bounded in terms of r and w only.*

Recall that a group G is metanilpotent if there is a normal subgroup N such that N and G/N are both nilpotent. Obviously, Theorem 1.4 furnishes evidence that, for multilinear commutator words, the answer to our question should be positive. We do not know for which other words results of a similar nature can be obtained. In particular, it would be interesting to see whether Theorems 1.2 and 1.4 remain valid when w is an Engel word.

2. Multilinear commutator words

We recall that multilinear commutator words are recursively defined as follows. The group-word $w(x) = x$ in one indeterminate is a multilinear commutator; if α and β are multilinear commutators involving disjoint sets of indeterminates, then the word $w = [\alpha, \beta]$ is a multilinear commutator, and all multilinear commutators are obtained in this way. Here and in the rest of the paper, given two elements a and b of a group G, we use the standard notation $[a, b] = a^{-1}b^{-1}ab$ and $a^b = b^{-1}ab$. Examples of multilinear commutators include the familiar lower central words $\gamma_n(x_1, \ldots, x_n) = [x_1, \ldots, x_n]$ and derived words δ_n, on 2^n variables, defined recursively by

$$\delta_0 = x_1, \qquad \delta_n = [\delta_{n-1}(x_1, \ldots, x_{2^{n-1}}), \delta_{n-1}(x_{2^{n-1}+1}, \ldots, x_{2^n})].$$

Of course, $\delta_n(G) = G^{(n)}$ is the nth derived group of a group G.

Let X be a subset of a group and let $w = w(x_1, \ldots, x_n)$ be a multilinear commutator word. We set

$$X_w = \{w(y_1, \ldots, y_n) \mid y_i \in X\}.$$

We say that the subset X is commutator-closed if $[x, y] \in X$ whenever $x, y \in X$.

If $G = \langle X \rangle$ is a group generated by a commutator-closed set X, then the commutator subgroup G' is generated by commutators $[x, y]$, where $x, y \in X$ (see [2, Lemma 2.2]). Here we generalise this result to multilinear commutator words.

PROPOSITION 2.1. *Let $G = \langle X \rangle$ be a group generated by a commutator-closed set X and let w be a multilinear commutator word. Then $w(G) = \langle X_w \rangle$.*

To prove Proposition 2.1, we need the concepts of *height* and *defect* of a multilinear commutator word, introduced in [8]. For the reader's convenience, we now describe some results from [8].

DEFINITION 2.2. The height and the labelled tree of a multilinear commutator word are defined recursively as follows.
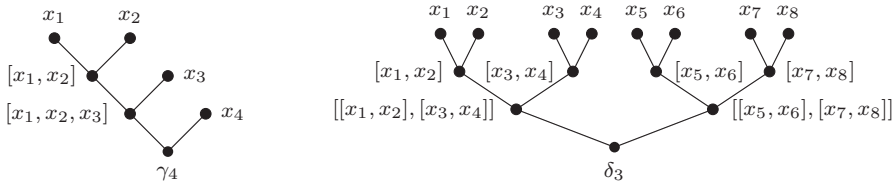
FIGURE 1. The trees of the words γ_4 and δ_3.

(1) An indeterminate has height zero, and its tree is an isolated vertex, labelled with the name of the indeterminate.

(2) If $w = [\alpha, \beta]$, where α and β are disjoint multilinear commutator words, then the height ht(w) of the word w is taken to be the maximum of the heights of α and β plus one, and the tree of w is obtained by adding a new vertex with label w and connecting it to the vertices labelled α and β of the corresponding trees of these words.

The tree of a multilinear commutator word w provides a visual way of reading how w is constructed by nesting commutators, which is easier than writing the actual expression for w using commutator brackets. We draw these trees by going downwards whenever we form a new commutator, so that the vertex with label w is placed at the root of the tree. Every vertex v is labelled with a multilinear commutator word, which we denote by w_v. Note that the indeterminates correspond exactly to the vertices of degree one. Also, the height of w coincides with the height of the tree, that is, the largest distance from the root to another vertex of the tree (which is necessarily labelled by an indeterminate). For example, Figure 1 shows the trees for the words γ_4 and δ_3.

More generally, the full tree of height h corresponds to the derived word δ_h.

All labels of the tree of a multilinear commutator word are determined, up to permuting the indeterminates, by the tree itself (as a graph without labels): given the tree, we only need to associate an indeterminate to every vertex of degree one, and then proceed downwards by labelling each vertex with the commutator of the labels of its immediate ascendants.

DEFINITION 2.3. Let v be a vertex of the tree of a multilinear commutator word w of height h. We say that v is in the *ith level* of the tree if it lies at distance $h - i$ from the root of the tree.

Thus the upmost level is level zero and the root is at level h, but note that a vertex v at level i is not necessarily labelled with a word w_v of height i: it might happen that w_v is an indeterminate.

It is also useful to associate a *companion* vertex to each vertex of the tree that is different from the root, defined as follows.

DEFINITION 2.4. Let p be a vertex of the tree of a multilinear commutator word w that is different from the root and let u be the immediate descendant of p. Then the

companion of p is the only other vertex q of the tree which has u as an immediate descendant.

It is clear that companion vertices lie on the same level of the tree.

We prove Proposition 2.1 for a general multilinear commutator word w by induction on the 'distance' of w to the closest derived word. We make this notion of distance precise in the following definition.

DEFINITION 2.5. Let w be a multilinear commutator word of height h. Then the *defect* of w, which is denoted by def w, is defined as

$$\text{def } w = 2^{h+1} - 1 - V,$$

where V is the number of vertices of the tree of w.

So, if the height of w is h, then the defect is the number of vertices that need to be added to the tree of w in order to get the tree of δ_h. Thus the defect is zero if and only if w is a derived word, and we have def $\gamma_4 = 8$ and def $[\gamma_3, \gamma_3] = 4$.

DEFINITION 2.6. Let T be the tree associated to a multilinear commutator word w. A subset S of vertices of T is called a *section* of T if S is maximal (with respect to inclusion) subject to the condition that S does not contain two vertices where one is a descendant of the other. Equivalently, in terms of labels, this means that every indeterminate involved in w appears in exactly one word w_v with $v \in S$.

A very natural way of obtaining a section is by cutting a tree below level i, that is, we consider the section S containing all vertices at level $i + 1$ and all the vertices of the tree lying below level $i + 1$ labelled by an indeterminate. This is the type of section that we use in the proof of Proposition 2.1. Figure 2 shows an example of a section obtained by cutting below level 0.
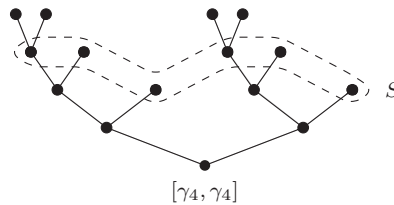
FIGURE 2. Example of a section.

LEMMA 2.7 [8, Lemma 3.2]. *Let w be a multilinear commutator word and let T be the tree of w. If η is another multilinear commutator word, then, for every $v \in T$, we define $\pi^{(v)}$ to be the word whose tree is obtained by replacing the tree of w_v at vertex v by the tree of $[w_v, \eta]$. Then, for every section S of T and for every group G,*

$$[w(G), \eta(G)] \leq \prod_{v \in S} \pi^{(v)}(G).$$

DEFINITION 2.8. Let φ and w be two multilinear commutator words.

1. We say that w is a *constituent* of φ if w is the label of a vertex in the tree of φ.
2. We say that φ is an *extension* of w if the tree of φ is an upward extension of the tree of w (simply as a tree, without labels).

Thus, in order to get an extension of w, we only need to draw new binary trees at some of the vertices which are labelled by indeterminates in the tree of w.

$$[\gamma_4, \delta_2]$$

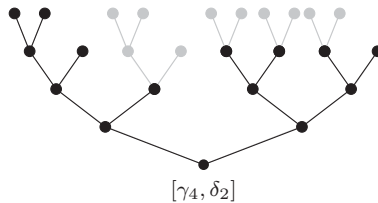FIGURE 3. An extension of $[\gamma_4, \delta_2]$.

In Figure 3, the black tree represents the word $w = [\gamma_4, \delta_2]$ and the extension of w, which is obtained by adding the grey trees, is $\varphi = [[\gamma_3, \gamma_3], [\delta_2, \gamma_3]]$. Clearly, the derived word δ_h is an extension of all words of height less than or equal to h.

Observe that, if G is a group and φ is an extension of w, then $\varphi(G) \leq w(G)$. Moreover, if α is a constituent of w, then $w(G) \leq \alpha(G)$.

The proof of Proposition 2.1 depends on the following result which is implicit in the proof of Theorem B of [8].

THEOREM 2.9. *Let $w = [\alpha, \beta]$ be a multilinear commutator word of height h. If $w \neq \delta_h$, then at least one of the subgroups $[w(G), \alpha(G)]$ and $[w(G), \beta(G)]$ is contained in a product of verbal subgroups corresponding to words that are proper extensions of w of height h.*

For the reader's convenience we include a proof.

PROOF. Let Φ be the (finite) set of all multilinear commutator words of height h that are a proper extension of w and set $H = \prod_{\varphi \in \Phi} \varphi(G)$.
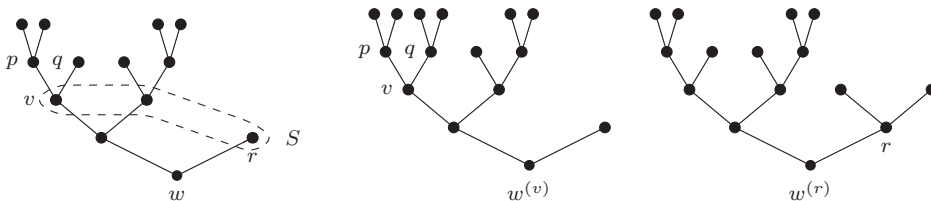

FIGURE 4. The two different cases for the construction of $w^{(v)}$ with $v \in S$. Observe that $i = 1$ in this example.

Let i be the largest integer for which there is a vertex in the tree of w at level i with label δ_i. Note that $1 \leq i < h$, since w is not a derived word. Let S be the section of the tree of w obtained by cutting the tree below level i, so that S contains all vertices at level $i + 1$ and all the vertices of the tree lying below level $i + 1$ that are labelled with an indeterminate. For every vertex v in S, we construct a word $w^{(v)}$, as follows (see Figure 4 for an example). If the label w_v of v is not an indeterminate, then we can write $w_v = [w_p, w_q]$, where p and q are the companion vertices at level i having v as immediate descendant. By the maximality of i, one of these vertices is labelled with a word which is different from δ_i. For simplicity, let us assume that this happens for q, the vertex on the right (the argument is exactly the same otherwise). We define $w^{(v)}$ to be the word whose tree is obtained by replacing w_q with δ_i in the tree of w. Thus the label of $w^{(v)}$ at the vertex v is the commutator $[w_p, \delta_i]$. On the other hand, if w_v is an indeterminate, then $w^{(v)}$ is defined simply by putting the tree corresponding to δ_i on top of the vertex v in the tree of w.

In any case, it is clear that $\mathrm{ht}(w^{(v)}) = h$ and that $w^{(v)}$ is a proper extension of w, so that $w^{(v)}$ belongs to Φ. Consequently, we have $w^{(v)}(G) \leq H$ for every vertex v in the section S.

On the other hand, if we apply Lemma 2.7 to the section S with δ_i playing the role of η, then

$$[w(G), \delta_i(G)] \leq \prod_{v \in S} \pi^{(v)}(G). \tag{2-1}$$

Here, $\pi^{(v)}$ is the word whose tree is obtained by inserting the tree of $[w_v, \delta_i]$ at vertex v into the tree of w. Now it is easy to compare the two words $w^{(v)}$ and $\pi^{(v)}$: they look the same at all vertices of the original tree of w, except for the vertex v, where $\pi^{(v)}$ has the label $[w_v, \delta_i]$ and $w^{(v)}$ has either $[w_p, \delta_i]$ or δ_i. In either of the two cases,

$$(\pi^{(v)})_v(G) \leq (w^{(v)})_v(G),$$

and then, since $\pi^{(v)}$ and $w^{(v)}$ have the same labels outside the tree above v, also

$$\pi^{(v)}(G) \leq w^{(v)}(G).$$

Since this happens for all vertices in S, it follows from (2-1) that

$$[w(G), \delta_i(G)] \leq H.$$

Now, by the definition of i, the derived word δ_i is a constituent of either α or β and, consequently, either $[w(G), \alpha(G)] \leq H$ or $[w(G), \beta(G)] \leq H$. This concludes the proof of the theorem. □

We also need the following observation.

LEMMA 2.10. *Let H, K be two subgroups of a group G such that $H = \langle A \rangle$, $K = \langle B \rangle$ and let $C = \{[a, b] \mid a \in A, b \in B\}$. Let $T = \langle C \rangle$. If $[a, b]^{\tilde{a}}, [a, b]^{\tilde{b}} \in T$ for all $a, \tilde{a} \in A$ and $b, \tilde{b} \in B$, then $T = [H, K]$.*

PROOF. It follows from the hypotheses that T is a normal subgroup of $\langle H, K \rangle$, and thus, in the quotient group $\langle \bar{H}, \bar{K} \rangle = \langle HT/T, KT/T \rangle$, we have that every generator of $\bar{H}$ commutes with every generator of $\bar{K}$. Therefore $[\bar{H}, \bar{K}] = 1$, that is, $[H, K] \leq T$. The reverse inclusion is obvious.   □

PROOF OF PROPOSITION 2.1. We argue by double induction: we first use induction on the height of the word w, and then, for a fixed value of the height, induction on the defect of w. If w has height zero, then $w = x_1$ and the result holds. Now assume that $h = \mathrm{ht}(w) \geq 1$ and that the result has been proved for any multilinear commutator word whose height is less than h. If $\mathrm{def}\, w = 0$, then w is a derived word, and the result holds by [2, Lemma 2.2]. So we assume that $\mathrm{def}\, w > 0$. Let us write $w = [\alpha, \beta]$, where α and β are multilinear commutator words of height smaller than h. Then, by induction on the height, $\alpha(G) = \langle X_\alpha \rangle$ and $\beta(G) = \langle X_\beta \rangle$.

Let Φ be the (finite) set of all multilinear commutator words of height h that are a proper extension of w. Then, by induction on the defect, $\varphi(G) = \langle X_\varphi \rangle$ for each $\varphi \in \Phi$.

Let $T = \langle X_w \rangle$. Since X is commutator-closed, we have that $X_\varphi \subseteq X_w$ for each $\varphi \in \Phi$. Therefore

$$\prod_{\varphi \in \Phi} \varphi(G) \leq T.$$

In view of Theorem 2.9, at least one of the subgroups $[w(G), \alpha(G)]$ and $[w(G), \beta(G)]$ is contained in T.

Now, by Lemma 2.10, in order to prove that $T = w(G)$ it is enough to prove that $[a, b]^{\tilde{a}}, [a, b]^{\tilde{b}} \in T$ for all $a, \tilde{a} \in X_\alpha$ and $b, \tilde{b} \in X_\beta$.

Let us assume that $[w(G), \alpha(G)] \leq T$; the other case is similar.

Then $[a, b]^{\tilde{a}} = [a, b][a, b, \tilde{a}] \in T$, as $[a, b] \in X_w$ and $[a, b, \tilde{a}] \in [w(G), \alpha(G)] \leq T$. This proves that T is normalised by $\alpha(G)$.

Moreover, as $[a, b] \in \alpha(G) = \langle X_\alpha \rangle$, we can write $[a, b] = c_1 c_2 \cdots c_r$ with $c_i \in X_\alpha \cup X_\alpha^{-1}$. Then $[a, b]^{\tilde{b}} = [a, b][a, b, \tilde{b}]$ and, by the standard commutator identities, we can write $[a, b, \tilde{b}] = [c_1 c_2 \cdots c_r, \tilde{b}]$ as the product of r $\alpha(G)$-conjugates of elements of the form $[c_i, \tilde{b}]$. If $c_i \in X_\alpha$, then $[c_i, \tilde{b}] \in X_w \leq T$. If $c_i^{-1} \in X_\alpha$, then again $[c_i, \tilde{b}] = ([c_i^{-1}, \tilde{b}]^{-1})^{c_i} \in T^{\alpha(G)} = T$. It follows that $[a, b, \tilde{b}] \in T$ and thus $[a, b]^{\tilde{b}} = [a, b][a, b, \tilde{b}] \in T$, as desired.   □

3. Proof of Proposition 1.3

The focal subgroup theorem (see, for example, [9, Theorem 7.3.4]) says that if P is a Sylow subgroup of a finite group G, then $P \cap G'$ is generated by elements of the form $[x, y] \in P$, where $x \in P$ and $y \in G$. In particular, it follows that the Sylow subgroups of G' are generated by commutators. Thus, the following question arises.

Let w be a commutator word, let G be a finite group and let P be a Sylow p-subgroup of $w(G)$. Is it true that P can be generated by w-values lying in P?

The above question was introduced in [1] where it was proved that if w is a multilinear commutator word, then P is generated by powers of w-values. In this

section, we prove that if G is soluble, then indeed P can be generated by w-values. For the derived words, this was established in [2, Lemma 2.5]. To deal with arbitrary multilinear commutator words, we require the following combinatorial lemma.

Let i be a positive integer. We denote by I the set of all n-tuples $(i_1, \ldots, i_n)$, where all entries i_k are nonnegative integers. We view I as a partially ordered set with the partial order given by the rule that

$$(i_1, \ldots, i_n) \leq (j_1, \ldots, j_n)$$

if and only if $i_1 \leq j_1, \ldots, i_n \leq j_n$.

Given a group G, a multilinear commutator word $w = w(x_1, \ldots, x_n)$ and $\mathbf{i} = (i_1, \ldots, i_n) \in I$, we write

$$w(\mathbf{i}) = w(G^{(i_1)}, \ldots, G^{(i_n)})$$

for the subgroup generated by the w-values $w(g_1, \ldots, g_n)$ with $g_j \in G^{(i_j)}$. Further, set

$$w(\mathbf{i}^+) = \prod w(\mathbf{j}),$$

where the product is taken over all $\mathbf{j} \in I$ such that $\mathbf{j} > \mathbf{i}$.

Observe that $w(\mathbf{i}) = w_{\mathbf{i}}(G)$, where $w_{\mathbf{i}}$ is the extension of w obtained by replacing, in $w(x_1, \ldots, x_n)$, each x_j with the word δ_{i_j}, for $j = 1, \ldots, n$.

Note that if $\delta_h(G) = 1$, then there is an n-tuple $\mathbf{i}$ such that $w(\mathbf{i}) \neq 1$ but $w(\mathbf{i}^+) = 1$.

LEMMA 3.1 [4, Corollary 6]. *Let G be a group, let $w = w(x_1, \ldots, x_n)$ be a multilinear commutator word and let $\mathbf{i} \in I$. If $w(\mathbf{i}^+) = 1$, then $w(\mathbf{i})$ is abelian.*

The next lemma is taken from [1].

LEMMA 3.2 [1, Lemma 1.1]. *Let G be a finite group and let N be a normal subgroup of G. If P is a Sylow p-subgroup of G and Y is a normal subset of G consisting of p-elements, then $YN \cap PN = (Y \cap P)N$.*

The following lemma also plays an important role.

LEMMA 3.3 [2, Lemma 2.1]. *Any finite soluble group is generated by a commutator-closed set, all of whose elements have prime power order.*

Now we are ready to prove that if w is a multilinear commutator word and G is a finite soluble group, then, for any Sylow p-subgroup P of G, the corresponding Sylow p-subgroup $P \cap w(G)$ of $w(G)$ can be generated by the w-values lying in P.

PROOF OF PROPOSITION 1.3. Recall that G is a finite soluble group and that w is a multilinear commutator word. We know from Lemma 3.3 that there is a commutator-closed subset $X \subseteq G$ such that X generates G and every element of X has prime power order.

Recall that X_w stands for the set $\{w(x_1, \ldots, x_n) \mid x_i \in X\}$. For a prime $p \in \pi(G)$, set

$$X_{w,p} = \{x \in X_w \mid x \text{ is a } p\text{-element}\}$$

and

$$Y_{w,p} = \{X_{w,p}^G\},$$

that is, $Y_{w,p}$ is the union of the conjugacy classes of elements of $X_{w,p}$.

We claim that if P is a Sylow p-subgroup of G, then $P \cap w(G) = \langle P \cap Y_{w,p} \rangle$.

Without loss of generality, assume that G is a minimal counterexample to the above claim. If N is a nontrivial normal subgroup of G, then the set $\bar{X} = \{xN \mid x \in X\}$ has the required properties. Since X is a commutator-closed set, we deduce that every element of X_w has prime power order. Hence,

$$(\bar{X})_{w,p} = \overline{X_{w,p}} = \{xN \in \bar{X}_w \mid xN \text{ is a } p\text{-element}\},$$

so that $\bar{Y}_{w,p} = \overline{Y_{w,p}}$. By minimality of G,

$$\bar{P} \cap w(\bar{G}) = \langle \bar{P} \cap \bar{Y}_{w,p} \rangle = \langle \bar{P} \cap \overline{Y_{w,p}} \rangle.$$

By virtue of Lemma 3.2,

$$P \cap w(G) = \langle P \cap N, P \cap Y_{w,p} \rangle. \tag{3-1}$$

So, if N is a p'-group, then $P \cap N = 1$ and $P \cap w(G) = \langle P \cap Y_{w,p} \rangle$, which is a contradiction. Hence G has no nontrivial normal p'-subgroups.

We now use the notation introduced before Lemma 3.1. Let $\mathbf{i} \in I$ such that $w(\mathbf{i}) \neq 1$ but $w(\mathbf{i}^+) = 1$. By Lemma 3.1, $w(\mathbf{i})$ is abelian. Since G has no nontrivial normal p'-subgroups, $w(\mathbf{i})$ is a p-group. It follows from Proposition 2.1 that

$$w(\mathbf{i}) = w_{\mathbf{i}}(G) = \langle X_{w_{\mathbf{i}}} \rangle.$$

Since $w(\mathbf{i})$ is a p-group, $X_{w_{\mathbf{i}}} = X_{w_{\mathbf{i},p}}$ and $P \cap w(\mathbf{i}) = w(\mathbf{i})$. Moreover, every $w_{\mathbf{i}}$-value is a w-value, and hence $X_{w_{\mathbf{i},p}} \leq Y_{w,p}$. We deduce from (3-1), applied with $N = w(\mathbf{i})$, that

$$P \cap w(G) = \langle P \cap Y_{w,p}, P \cap w(\mathbf{i}) \rangle = \langle P \cap Y_{w,p}, w(\mathbf{i}) \rangle = \langle P \cap Y_{w,p} \rangle,$$

which is contrary to our assumptions. $\qquad\square$

4. The proofs of Theorem 1.2 and Theorem 1.4

As mentioned in the introduction, now the proof of Theorem 1.2 follows easily.

PROOF OF THEOREM 1.2. Recall that G is a finite group in which every nilpotent subgroup generated by w-values has rank at most r. Let H be a subgroup of $w(G)$ and let P be a Sylow p-subgroup of H. By the aforementioned result of Guralnick [11] and Lucchini [17], it is sufficient to prove that the rank of P is at most r.

The Ore conjecture that every element of a nonabelian finite simple group is a commutator was famously verified in [14]. Since w is a multilinear commutator word, we easily deduce that if G is a nonabelian simple group, then every element of G is a w-value. So P is a nilpotent subgroup of G generated by w-values. By hypotheses, the rank of P is at most r.

Now assume that G is soluble and let $\tilde{P}$ be a Sylow p-subgroup of $w(G)$ containing P. By Proposition 1.3, $\tilde{P}$ can be generated by w-values. Hence, the rank of $\tilde{P}$ is at most r. So, P also has rank at most r. $\qquad\square$

We now start the preparations for the proof of Theorem 1.4.

The Frattini subgroup of a group G is denoted by $\mathrm{Frat}(G)$. Let us denote by $\mathrm{Fit}(G)$ the Fitting subgroup of G and by $F_i(G)$ the ith term of the upper Fitting series of G, defined recursively by $F_1(G) = \mathrm{Fit}(G)$ and $F_i(G)/F_{i-1}(G) = \mathrm{Fit}(G/F_{i-1}(G))$. If G is a finite soluble group, then the least number h with the property that $F_h(G) = G$ is called the Fitting height of G.

The next lemma is quite well known.

LEMMA 4.1. *Let G be a finite soluble group of rank at most r. Then the Fitting height of G is at most $2r + 1$.*

PROOF. Since $\mathrm{Fit}(G)/\mathrm{Frat}(G) = \mathrm{Fit}(G/\mathrm{Frat}(G))$ (see, for example, [18, 5.2.15]), without loss of generality, we can assume that $\mathrm{Frat}(G) = 1$. In this case, $F = \mathrm{Fit}(G)$ is a direct product of abelian minimal normal subgroups of G, say,

$$F = N_1 \times \cdots \times N_t,$$

where each N_i is an elementary abelian p_i-group of rank at most r.

Set $H_i = G/C_G(N_i)$, for $i = 1, \ldots t$. Every H_i is isomorphic to a soluble linear group acting on N_i, where N_i is a vector space of dimension at most r. As a soluble subgroup of $GL(n, \mathbb{F})$, where $\mathbb{F}$ is any field, has derived length at most $2n$ (see, for instance, [7, Theorem 6.2A]), the derived length of each H_i is bounded by $2r$. Therefore $G/\cap_{i=1}^{t} C_G(N_i)$ has derived length at most $2r$. Since $\cap_{i=1}^{t} C_G(N_i) = C_G(F) \leq F$ (see, for example, [18, 5.4.4]) it follows that G/F has derived length at most $2r$. We conclude that G has Fitting height at most $2r + 1$. $\qquad\square$

As a corollary of Theorem 1.2, we deduce the following.

COROLLARY 4.2. *Let w be a multilinear commutator word and let K be a finite soluble group in which every nilpotent subgroup generated by w-values has rank at most r. Then the Fitting height of K is bounded in terms of r and w.*

PROOF. Let n be the height of w. As every δ_n-value of K is a w-value, $K/w(K)$ is soluble of derived length at most n. So it is sufficient to bound the Fitting height of $w(K)$. By Theorem 1.2 applied to K, the verbal subgroup $w(K)$ has rank at most $r + 1$. Therefore, by Lemma 4.1, the Fitting height of $w(K)$ is bounded in terms of r. $\qquad\square$

Every finite group G has a normal series, each of whose quotients is either soluble or is a direct product of nonabelian simple groups. The nonsoluble length of G, denoted by $\lambda(G)$, was defined in [12] as the minimal number of nonsoluble factors in a series of this kind: if

$$1 = G_0 \leq G_1 \leq \cdots \leq G_{2k+1} = G$$

is a shortest normal series in which, for i even, the quotient G_{i+1}/G_i is soluble (possibly trivial) and, for i odd, the quotient G_{i+1}/G_i is a (nonempty) direct product of nonabelian simple groups, then the nonsoluble length $\lambda(G)$ is equal to k.

PROPOSITION 4.3 [6, Proposition 2.2]. *Let N, M be normal subgroups of G such that $\lambda(G/N) \leq \lambda(G/M) \leq 1$. Then $\lambda(G/N \cap M) \leq 1$.*

Given a finite group G, we define $T(G)$ as the intersection of all normal subgroups N of G such that $\lambda(G/N) \leq 1$. It is easy to deduce from Proposition 4.3 that $\lambda(G/T(G)) \leq 1$ and $\lambda(G/T(G)) = 1$ if and only if G is nonsoluble.

It is proved in [12] that the nonsoluble length $\lambda(G)$ does not exceed the maximum Fitting height of soluble subgroups of a finite group G. A straightforward consequence of this result and Corollary 4.2 is the following lemma.

LEMMA 4.4. *Let w be a multilinear commutator word and let G be a finite group in which every nilpotent subgroup generated by w-values has rank at most r. Then the nonsoluble length of G is bounded in terms of r and w.*

The following well-known lemma is useful.

LEMMA 4.5. *Let N be a normal subgroup of a finite group G. Then there exists a subgroup H of G such that $G = HN$ and $H \cap N \leq \mathrm{Frat}(H)$.*

PROOF. The lemma clearly holds if $N \leq \mathrm{Frat}(G)$, with $H = G$. On the other hand, if N is not a subgroup of $\mathrm{Frat}(G)$, then there exists a proper subgroup of G supplementing N. Let H be a subgroup of G which is minimal with respect to the property that $G = HN$. If $N \cap H$ is not contained in $\mathrm{Frat}(H)$, then there exists a proper subgroup M of H such that $H = M(N \cap H)$. Thus $G = NH = MN$, against the minimality of H.   □

If w is a multilinear commutator word and N is a normal subgroup of a group G containing no nontrivial w-values, then N centralises $w(G)$ (see, for example, [21, Theorem 2.3] or the comment after Lemma 4.1 in [5]).

The next two lemmas deal with particular cases of Theorem 1.4. Clearly, if G is perfect, then $G = w(G)$.

LEMMA 4.6. *Let w be a multilinear commutator word and let G be a finite group in which every metanilpotent subgroup generated by w-values has rank at most r. If $G/\mathrm{Frat}(G)$ is a direct product of nonabelian simple groups, then the rank of G is bounded in terms of r and w.*

PROOF. Let P be a Sylow p-subgroup of G and set $\Phi = \mathrm{Frat}(G)$. As $P\Phi$ is metanilpotent, by assumption, the set $Y = G_w \cap P\Phi$ generates a subgroup of rank at most r. It follows, from the Ore conjecture [14], that every element of G/Φ is a w-value. Thus P is contained in the set $Y\Phi$ and so $P\Phi/\Phi \leq \langle Y \rangle \Phi/\Phi$ has rank at most r.

First, assume that G/Φ is a simple group. The subgroup N generated by $G_w \cap \Phi$ is nilpotent, and hence, by assumption, its rank is at most r. Since the image of Φ in G/N contains no nontrivial w-values of G/N, and w is a multilinear commutator, Φ/N

centralises $w(G/N) = G/N$. So Φ/N is a quotient of the Schur multiplier of the simple group G/Φ. A corollary of the classification of finite simple groups is that the rank of the Schur multiplier of any such group is at most two (see, for example, [10, Table 4.1]). As $P\Phi/\Phi$, Φ/N and N have bounded rank, we deduce that P has bounded rank. Since this holds for every prime p, the aforementioned result of Guralnick [11] and Lucchini [17] implies that the rank of G is bounded.

Now assume that $G/\Phi = S_1 \times \cdots \times S_t$ is a direct product of $t > 1$ nonabelian simple groups S_i. Let Q be a Sylow 2-subgroup of G. By the above argument, $Q\Phi/\Phi$ has rank at most r. Since each S_i has a nontrivial Sylow 2-subgroup, we deduce that t is at most r. Therefore the lemma follows from the case where G/Φ is simple. $\qquad\square$

LEMMA 4.7. *Let w be a multilinear commutator word and let G be a perfect finite group such that $\lambda(G) = 1$. Assume that every metanilpotent subgroup of G generated by w-values has rank at most r. Then the rank of G is bounded in terms of r and w.*

PROOF. As G is perfect and $\lambda(G) = 1$, the quotient group of G over its soluble radical $R(G)$ is a direct product of nonabelian simple groups. Moreover, by Corollary 4.2, the Fitting height of $R(G)$ is bounded in terms of r and w. So G has a normal series of bounded length

$$1 = G_0 < G_1 < \cdots < G_{s-1} < G_s = G, \qquad (4\text{-}1)$$

where G/G_{s-1} is a direct product of nonabelian simple groups and each section G_i/G_{i-1} is nilpotent for $i = 1, \ldots, s - 1$. We argue by induction on the minimal length s of such a series. The case $s = 1$ is handled in Lemma 4.6, so we assume that $s > 1$.

Let H be a subgroup of G which is minimal with respect to the properties that $G = HG_1$ and $H \cap G_1 \leq \mathrm{Frat}(H)$ (see Lemma 4.5). Note that $H/H \cap G_1$ is perfect since it is isomorphic to G/G_1. We have $H = H'(H \cap G_1)$, and hence $H'G_1 = HG_1 = G$. We therefore conclude that H is perfect, by minimality of H. Moreover, $\lambda(H) = \lambda(G) = 1$.

Consider the series of H

$$1 \leq G_1 \cap H \leq \cdots \leq G_{s-1} \cap H \leq G_s \cap H = H.$$

If $s > 2$, then $(G_2 \cap H)/(G_1 \cap H)$ is nilpotent. Taking into account that $G_1 \cap H \leq \mathrm{Frat}(H)$, we deduce that $G_2 \cap H$ is nilpotent. By induction on the minimal length of a series as in (4-1), H has bounded rank.

On the other hand, if $s = 2$, then $H/\mathrm{Frat}(H)$ is a homomorphic image of G/G_1, which is a direct product of nonabelian simple groups, and we can apply Lemma 4.6 to conclude that also, in this case, H has bounded rank.

Now consider the subgroup $N = \langle G_w \cap G_1 \rangle$ generated by the w-values of G lying in G_1. Since the image of G_1 in G/N contains no nontrivial w-values of G/N and w is a multilinear commutator, G_1/N centralises $w(G/N) = G/N$. Set $K = G/N$ and note that $K/Z(K)$ is a homomorphic image of H. Therefore $K/Z(K)$ has bounded rank. A theorem of Lubotzky and Mann [16] (see also [13]) now implies that the derived group K' of K has bounded rank. Since G is perfect, we conclude that G/N has bounded rank.

Finally, note that since N is a nilpotent subgroup generated by w-values, it has rank at most r by the hypothesis. Therefore G has bounded rank, as claimed.   □

Write $G^{(\infty)}$ for the last term of the derived series of G. Set $T_1(G) = G^{(\infty)}$ and, by induction, $T_{i+1}(G) = T(T_i(G))$. In view of Proposition 4.3, it is clear that if $T_{i-1}(G) \neq 1$, then $T_i(G)$ is the unique smallest normal subgroup N of G such that $\lambda(G/N) = i - 1$. Moreover, $\lambda(T_i(G)/T_{i+1}(G)) = 1$ and $T_i(G)$ is perfect for every $i \geq 1$ such that $T_i(G) \neq 1$.

PROOF OF THEOREM 1.4. Recall that w is a multilinear commutator word and G is a finite group in which every metanilpotent subgroup generated by w-values has rank at most r. We want to prove that the rank of the verbal subgroup $w(G)$ is bounded in terms of r and w only. By Corollary 4.4, the nonsoluble length $\lambda = \lambda(G)$ of G is bounded in terms of r and w. We argue by induction on λ. If $\lambda(G) = 0$, then G is soluble, and the result follows from Theorem 1.2. Assume that $\lambda \geq 1$, and let $N = T_\lambda$. Note that N is perfect. Moreover, $\lambda(G/N) = \lambda - 1$ and $\lambda(N) = 1$. By Lemma 4.5, there exists a subgroup H of G such that $G = HN$ and $H \cap N \leq \text{Frat}(H)$. Since $\lambda(H/H \cap N) \leq \lambda(G/N) = \lambda - 1$ and $H \cap N \leq \text{Frat}(H)$ is soluble, the nonsoluble length of H is at most $\lambda - 1$. As H inherits the assumptions, by induction, $w(H)$ has bounded rank. Moreover, as N is perfect and $\lambda(N) = 1$, we deduce from Lemma 4.7 that N has bounded rank. Now $w(G)/w(G) \cap N$ is isomorphic to $w(G)N/N = w(H)N/N$, so it has bounded rank. As N also has bounded rank, we conclude that $w(G)$ has bounded rank.   □

References

[1]   C. Acciarri, G. A. Fernàndez-Alcober and P. Shumyatsky, 'A focal subgroup theorem for outer commutator words', *J. Group Theory* **15** (2012), 397–405.

[2]   J. da Silva Alves and P. Shumyatsky, 'On nilpotency of higher commutator subgroups of a finite soluble group', *Arch. Math.*, **116** (2021), 1–6.

[3]   E. Detomi, M. Morigi and P. Shumyatsky, 'Bounding the exponent of a verbal subgroup', *Ann. Mat. Pura Appl. (4)* **193** (2014), 1431–1441.

[4]   E. Detomi, M. Morigi and P. Shumyatsky, 'On countable coverings of word values in profinite groups', *J. Pure Appl. Algebra* **219** (2015), 1020–1030.

[5]   E. Detomi, M. Morigi and P. Shumyatsky, 'Profinite groups with restricted centralizers of commutators', *Proc. Roy. Soc. Edinburgh Sect. A* **150** (2020), 2301–2321.

[6]   E. Detomi and P. Shumyatsky, 'On the length of a finite group and of its 2-generator subgroups', *Bull. Braz. Math. Soc. (N.S.)* **47** (2016), 845–852.

[7]   J. D. Dixon, *The Structure of Linear Groups* (Van Nostrand Reinhold Company, London, 1971).

[8]   G. A. Fernández-Alcober and M. Morigi, 'Outer commutator words are uniformly concise', *J. Lond. Math. Soc. (2)* **82** (2010), 581–595.

[9]   D. Gorenstein, *Finite Groups* (Chelsea Publishing Company, New York, 1980).

[10] D. Gorenstein, *Finite Simple Groups: An Introduction to Their Classification* (Plenum Press, New York, 1982).

[11] R. Guralnick, 'On the number of generators of a finite group', *Arch. Math.* **53** (1989), 521–523.

[12] E. I. Khukhro and P. Shumyatsky, 'Nonsoluble and non-p-soluble length of finite groups', *Israel J. Math.* **207** (2015), 507–525.

[13]   L. A. Kurdachenko and P. Shumyatsky, 'The ranks of central factor and commutator groups', *Math. Proc. Cambridge Philos. Soc.* **154** (2013), 63–69.

[14]   M. W. Liebeck, E. A. O'Brien, A. Shalev and P. H. Tiep, 'The Ore conjecture', *J. Eur. Math. Soc.* **12**(4) (2010), 939–1008.

[15]   P. Longobardi and M. Maj, 'On the number of generators of a finite group', *Arch. Math. (Basel)* **50** (1988), 110–112.

[16]   A. Lubotzky and A. Mann, 'Powerful *p*-groups. I. Finite groups', *J. Algebra* **105** (1987), 484–505.

[17]   A. Lucchini, 'A bound on the number of generators of a finite group', *Arch. Math.* **53** (1989), 313–317.

[18]   D. J. S. Robinson, *A Course in the Theory of Groups*, 2nd edn, Graduate Texts in Mathematics, 80 (Springer-Verlag, New York, 1996).

[19]   D. Segal, *Words: Notes on Verbal Width in Groups*, London Mathematical Society Lecture Note Series, 361 (Cambridge University Press, Cambridge, 2009).

[20]   P. Shumyatsky, 'On the exponent of a verbal subgroup in a finite group', *J. Aust. Math. Soc.* **93** (2012), 325–332.

[21]   R. F. Turner-Smith, 'Marginal subgroup properties for outer commutator words', *Proc. Lond. Math. Soc. (3)* **14** (1964), 321–341.

ELOISA DETOMI, Dipartimento di Ingegneria dell'Informazione,
Università di Padova, Via G. Gradenigo 6/B,
35121 Padova, Italy
e-mail: eloisa.detomi@unipd.it

MARTA MORIGI, Dipartimento di Matematica,
Università di Bologna, Piazza di Porta San Donato 5,
40126 Bologna, Italy
e-mail: marta.morigi@unibo.it

PAVEL SHUMYATSKY, Department of Mathematics,
University of Brasilia, Brasilia-DF, 70910-900, Brazil
e-mail: pavel@unb.br