Submitted exclusively to the London Mathematical Society
doi:10.1112/0000/000000

Strong conciseness in profinite groups

Eloisa Detomi, Benjamin Klopsch and Pavel Shumyatsky

Abstract

A group word w is said to be strongly concise in a class $\mathcal{C}$ of profinite groups if, for every group G in $\mathcal{C}$ such that w takes less than $2^{\aleph_0}$ values in G, the verbal subgroup $w(G)$ is finite. Detomi, Morigi and Shumyatsky established that multilinear commutator words – and the particular words x^2 and $[x^2, y]$ – have the property that the corresponding verbal subgroup is finite in a profinite group G whenever the word takes at most countably many values in G. They conjectured that, in fact, this should be true for every word. In particular, their conjecture included as open cases power words and Engel words.

In the present paper, we take a new approach via parametrised words that leads to stronger results. First we prove that multilinear commutator words are strongly concise in the class of all profinite groups. Then we establish that every group word is strongly concise in the class of nilpotent profinite groups. From this we deduce, for instance, that, if w is one of the group words x^2, x^3, x^6, $[x^3, y]$ or $[x, y, y]$, then w is strongly concise in the class of all profinite groups. Indeed, the same conclusion can be reached for all words of the infinite families $[x^m, z_1, \ldots, z_r]$ and $[x, y, y, z_1, \ldots, z_r]$, where $m \in \{2, 3\}$ and $r \geq 1$.

1. *Introduction*

Let $w = w(x_1, \ldots, x_r)$ be a group word, i.e. an element of the free group on $x_1, \ldots, x_r$. We take an interest in the set of all w-values in a group G and the verbal subgroup generated by it; they are

$$G_w = \{w(g_1, \ldots, g_r) \mid g_1, \ldots, g_r \in G\} \qquad \text{and} \qquad w(G) = \langle G_w \rangle.$$

In the context of topological groups G, we write $w(G)$ to denote the closed subgroup generated by all w-values in G.

The word w is said to be *concise* in a class $\mathcal{C}$ of groups if, for each G in $\mathcal{C}$ such that G_w is finite, also $w(G)$ is finite. For topological groups, especially profinite groups, a variation of the classical notion arises quite naturally: we say that w is *strongly concise* in a class $\mathcal{C}$ of topological groups if, for each G in $\mathcal{C}$, already the bound $|G_w| < 2^{\aleph_0}$ implies that $w(G)$ is finite.

A conjecture proposed by Philip Hall (e.g. see [**18**]) predicted that every word w would be concise in the class of all groups, but almost three decades later the assertion was famously refuted by Ivanov [**10**]. On the other hand, Merzlyakov [**12**] showed already in the 1960s that every word is concise in the class of linear groups. This naturally leads to the question whether every word is concise in the class of residually finite groups, or equivalently in the class of profinite groups. Lately, this question was highlighted by Jaikin-Zapirain [**11**], who used Merzlyakov's theorem in his investigations of verbal width in finitely generated pro-p groups; compare also [**15**].

In [**2**], Detomi, Morigi and Shumyatsky suggested a strengthened profinite version of Hall's conciseness conjecture, namely that for every word w and every profinite group G, the bound $|G_w| \leq \aleph_0$ implies that $w(G)$ is finite. They verified this for multilinear commutator words, also known as outer-commutator words (see Section 3), as well as for the particular words

---

2000 *Mathematics Subject Classification* 20E18, 20E26, 20F10, 20F18, 20F12 .

x^2 and $[x^2, y]$. Their considerations relied on the Baire category theorem, but a more direct argument (see Section 2) allows us to deal with a natural stronger form of the conjecture.

For short, we say that a word w is strongly concise if it is strongly concise in the class of all profinite groups.

STRONG CONCISENESS CONJECTURE. Every group word w is strongly concise.

In the present paper, we initiate a systematic investigation of this conjecture and produce positive evidence for it. Among the words treated in [**2**], the special power word x^2 is the only one for which a simple replacement of the Baire category theorem by Proposition 2.1 below yields that it is strongly concise. More work is needed to confirm the Strong Conciseness Conjecture for multilinear commutator words.

THEOREM 1.1. *Every multilinear commutator word is strongly concise.*

Guided by an interest in power words x^m of exponent $m \geq 3$ and n-Engel words $[x,_n y] = [x, y, \ldots, y]$, where y appears $n \geq 2$ times, we began an investigation of some specific words, such as x^3 and $[x, y, y]$. Later we discovered that the relevant computations could be subsumed under a common approach. The main outcome of this consolidation is the following result.

THEOREM 1.2. *Every group word w is strongly concise in the class of nilpotent profinite groups.*

A straightforward and well-known argument shows that every group word is strongly concise in the class of abelian profinite groups; compare Proposition 2.3. But strong conciseness does not behave well under group extensions; Theorem 1.2 and, more importantly, the considerations that enter into its proof are new, even for nilpotent groups of class 2.

The following corollaries can be derived from Theorem 1.1 and Theorem 1.2 without further difficulty.

COROLLARY 1.3. *Let F be a free group of countably infinite rank and let w be a group word such that $F/w(F)$ is nilpotent. Then w is strongly concise.*

COROLLARY 1.4. *The following group words w are strongly concise:*

$$x^2, \quad x^3, \quad x^6, \quad [x, y, y] \qquad \text{and}$$

$$[x^2, z_1, \ldots, z_r], \quad [x^3, z_1, \ldots, z_r], \quad [x, y, y, z_1, \ldots, z_r] \qquad \text{for } r \geq 1,$$

where $x, y, z_1, z_2, \ldots$ are independent variables.

Our proof of Theorem 1.2 is based on parametrised words; see Section 5. Nilpotency is a key ingredient for setting up induction parameters that help us to reduce the complexity of the word w as well as the complexity of the group G under consideration.

As a byproduct, our approach highlights the relevance of the following two weaker versions of the Strong Conciseness Conjecture.

Conjecture 1.5. *Suppose that the group word w has less than $2^{\aleph_0}$ values in a profinite group G. Then $w(G)$ is generated by finitely many w-values.*

Conjecture 1.6. *Suppose that the group word w has less than $2^{\aleph_0}$ values in a profinite group G. Then there is an open subgroup H of G such that $w(H) = 1$.*

To illustrate the relevance of Conjecture 1.5, we summarise some conditional results that we obtained. For this we recall that if a group word w 'implies virtual nilpotency', then for a large class of groups G, including all finitely generated residually finite groups, $w(G) = 1$ implies that G is nilpotent-by-finite, due to results of Burns and Medvedev [**1**]. Furthermore, following [**7**] we say that a group word w is 'weakly rational' if for every finite group G and for every positive integer e with $\gcd(e, |G|) = 1$, the set G_w is closed under taking eth powers of its elements. We refer to Section 4 for a more detailed discussion of these notions.

Theorem 1.7. *Let w be a group word that* (i) *implies virtual nilpotency or* (ii) *is weakly rational. Let G be a profinite group such that $|G_w| < 2^{\aleph_0}$. If $w(G)$ is generated by finitely many w-values, then $w(G)$ is finite.*

Notation and Organisation. Our notation is mostly standard. All repeated commutators are left-normed, e.g. $\gamma_3(x, y, z) = [x, y, z] = [[x, y], z]$.

In Section 2 we collect some known results and several basic observations; the elementary Proposition 2.1 is one of the early key insights. In Section 3 we prove that multilinear commutator words are strongly concise. The main results in Section 4 are Propositions 4.7, 4.8 and 4.9; in particular, the latter two yield Theorem 1.7. In Section 5 we set up the reduction arguments based on parametrised words. In Section 6 we prove Theorem 1.2 as well as Corollaries 1.3 and 1.4.

2. *Preliminaries*

In this section we collect some known results as well as several straightforward consequences and basic observations.

For simplicity and to steer clear of the Continuum Hypothesis (or Martin's Axiom), we record the following proposition that helps us to avoid references to the Baire category theorem, which appear frequently in [**2**] and related articles.

Proposition 2.1. *Let $\varphi \colon X \to Y$ be a continuous map between non-empty profinite spaces that is nowhere locally constant, i.e. there exists no non-empty open subset $U \subseteq_o X$ such that $\varphi|_U$ is constant. Then $|X\varphi| \geq 2^{\aleph_0}$.*

Proof. For every non-empty closed open subset $U \subseteq X$ choose a continuous map $\vartheta_U : Y \to Z_U$ onto a finite discrete space Z_U such that $\varphi \vartheta_U : X \to Z_U$ is not constant on U. Choose non-empty distinct fibers U_1, U_2 of the restriction of $\varphi \vartheta_U$ to U; then $U_1, U_2 \subseteq U$ are non-empty closed open subsets of X with $U_1\varphi \cap U_2\varphi = \varnothing$.

Fix a non-empty closed open subset $A \subseteq X$, e.g. $A = X$. For every sequence $\mathbf{i} = (i_1, i_2, i_3, \ldots)$ in $\{1, 2\}$, the consideration above yields a descending chain of non-empty closed open subsets

$A_{i_1} \supseteq (A_{i_1})_{i_2} \supseteq ((A_{i_1})_{i_2})_{i_3} \supseteq \ldots$, and we set

$$A_{\mathbf{i}} = \bigcap\nolimits_{n\in\mathbb{N}} (\cdots((A_{i_1})_{i_2})_{i_3}\cdots)_{i_n} \subseteq_{\mathrm{c}} X.$$

Since X is compact, each $A_{\mathbf{i}}$ is non-empty, and we choose $a_{\mathbf{i}} \in A_{\mathbf{i}}$. By construction we have $a_{\mathbf{i}}\varphi \neq a_{\mathbf{j}}\varphi$ for $\mathbf{i} \neq \mathbf{j}$. Hence

$$B = \left\{a_{\mathbf{i}} \mid \mathbf{i} \in \{1,2\}^{\mathbb{N}}\right\} \subseteq X$$

is mapped injectively into Y under φ, and $|X\varphi| \geq |B\varphi| = 2^{\aleph_0}$. □

LEMMA 2.2. *Let G be a profinite group and let $x \in G$. If the conjugacy class $\{x^g \mid g \in G\}$ contains less than $2^{\aleph_0}$ elements, then it is finite.*

Proof. The set $\{x^g \mid g \in G\}$ is in bijection with the coset space $G/\mathrm{C}_G(x)$, a homogeneous profinite space. Alternatively, one can adapt the proof of [**2**, Lemma 3.1], using Proposition 2.1 in place of the Baire category theorem. □

PROPOSITION 2.3. *Every group word is strongly concise in the class of abelian profinite groups.*

Proof. Let G be an abelian profinite group. It is enough to consider power words $w(x) = x^n$, where $n \in \mathbb{N}$. For these we observe that $w(G) = \{g^n \mid g \in G\} = G_w$, as $G \to G$, $g \mapsto g^n$ is a homomorphism. Hence $w(G) = G_w$ is finite or has cardinality at least $2^{\aleph_0}$. □

LEMMA 2.4. *Let $w \in F$ be an element of a free group F such that $w \notin [F,F]$. Let G be a profinite group such that $|G_w| < 2^{\aleph_0}$. Then G is periodic.*

Proof. Write $w(x_1,\ldots,x_r) = x_1^{e_1}\cdots x_r^{e_r}v$, where $e_1,\ldots,e_r \in \mathbb{Z}$ are not all zero and $v \in [F,F]$. Then the word $y^m = w(y^{f_1},\ldots,y^{f_r})$, where $m = \sum_{i=1}^{r} e_i f_i = \gcd(e_1,\ldots,e_r) \in \mathbb{N}$, takes less than $2^{\aleph_0}$ values in G. By Proposition 2.3, every procyclic subgroup of G is finite, and thus G is periodic. □

3. *Multilinear commutator words*

In this section we prove that every multilinear commutator word is strongly concise. Recall that a multilinear commutator word, also known as an outer-commutator word, is obtained by nesting commutators and using each variable only once. Thus the word $[[x_1,x_2],[x_3,x_4,x_5],x_6]$ is a multilinear commutator word while the 3-Engel word $[x,y,y,y]$ is not. An important family of multilinear commutator words consists of the repeated commutator words γ_k on k variables, given by $\gamma_1 = x_1$ and $\gamma_k = [\gamma_{k-1}, x_k] = [x_1,\ldots,x_k]$ for $k \geq 2$. The verbal subgroup $\gamma_k(G)$ of a group G is the kth term of the lower central series of G. The derived words δ_k, on 2^k variables, form another distinguished family of multilinear commutators; they are defined by $\delta_0 = x_1$ and $\delta_k = [\delta_{k-1}(x_1,\ldots,x_{2^{k-1}}), \delta_{k-1}(x_{2^{k-1}+1},\ldots,x_{2^k})]$. The verbal subgroup $\delta_k(G) = G^{(k)}$ is the kth derived subgroup of G.

Relying on the Baire category theorem, Detomi, Morigi and Shumyatsky [**2**] proved that, if w is a multilinear commutator word, then for every profinite group G the bound $|G_w| \leq \aleph_0$ implies that $w(G)$ is finite. Proposition 2.1 enables us to strengthen this result: we show – without recourse to the Continuum Hypothesis (or Martin's Axiom) – that every multilinear commutator word is strongly concise. For this we employ combinatorial techniques that were developed in [**3**, **4**] specifically for handling multilinear commutator words.

Throughout this section, we fix $r \in \mathbb{N}$ and a multilinear commutator word

$$w = w(x_1, \dots, x_r).$$

Furthermore, G is a profinite group. For $A_1, \dots, A_r \subseteq G$, we denote by

$$w(A_1, \dots, A_r)$$

the subgroup generated by all w-values $w(a_1, \dots, a_r)$, where $a_i \in A_i$ for $1 \le i \le r$. For $I \subseteq \{1, \dots, r\}$ we write $\overline{I} = \{1, \dots, r\} \smallsetminus I$. For families of variables $\mathbf{y} = (y_i)_{i \in I}$, $\mathbf{z} = (z_i)_{i \in \overline{I}}$ we define

$$w_I(\mathbf{y}; \mathbf{z}) = w(u_1, \dots, u_r), \quad \text{where } u_s = \begin{cases} y_s & \text{if } s \in I, \\ z_s & \text{if } s \notin I. \end{cases}$$

The notation extends to families $\mathbf{A} = (A_i)_{i \in I}$, $\mathbf{B} = (B_i)_{i \in \overline{I}}$ of subsets of G in the natural way: $w_I(\mathbf{A}; \mathbf{B})$ denotes the subgroup generated by the relevant w-values. For short, we write $w_I(y_i; z_i)$ in place of $w_I(\mathbf{y}; \mathbf{z})$ and $w_I(A_i; B_i)$ in place of $w_I(\mathbf{A}; \mathbf{B})$.

The following are corollaries of [**4**, Lemma 2.5] and [**3**, Lemma 4.1].

Corollary 3.1. *Let $H \trianglelefteq_{\mathrm{c}} G$. Suppose that $g_1, \dots, g_r \in G$ and $g \in G$ are such that $w(g_1 h_1, \dots, g_r h_r) = g$ for all $h_1, \dots, h_r \in H$. Then $w_I(g_i H; H) = 1$ for every proper subset $I \subsetneqq \{1, \dots, r\}$.*

Corollary 3.2. *Let $H \trianglelefteq_{\mathrm{c}} G$, and suppose that $I \subseteq \{1, \dots, r\}$ is such that $w_J(G; H) = 1$ for all $J \subsetneqq I$. Then $w_I(g_i h_i; h_i) = w_I(g_i; h_i)$ for all $g_i \in G$, $i \in I$, and all $h_1, \dots, h_r \in H$.*

Next we employ the hypothesis $|G_w| < 2^{\aleph_0}$.

Lemma 3.3. *Let $H \trianglelefteq_{\mathrm{o}} G$ and $I \subsetneqq \{1, \dots, r\}$ be such that*

$$w_J(G; H) = 1 \text{ for all } J \subsetneqq I. \qquad (*)$$

Suppose that $|G_w| < 2^{\aleph_0}$. Let $(g_i)_{i \in I}$ be an arbitrary family in G. Then there exists $U \trianglelefteq_{\mathrm{o}} G$, with $U \subseteq H$, such that

$$w_I(g_i; U) = 1.$$

Proof. The image of the continuous map

$$H \times \cdots \times H \to G, \quad (h_1, \dots, h_r) \mapsto w_I(g_i h_i; h_i)$$

contains less than $2^{\aleph_0}$ elements. By Proposition 2.1, there exist $b_1, \dots, b_r \in H$ and $U \trianglelefteq_{\mathrm{o}} G$, with $U \subseteq H$, such that

$$w_I(g_i b_i u_i; b_i u_i) = w_I(g_i b_i; b_i) \quad \text{for all } u_1, \dots, u_r \in U.$$

As $I \subsetneqq \{1, \dots, r\}$, we conclude from Corollary 3.1 that

$$w_I(g_i b_i U; U) = 1. \qquad (3.1)$$

On the other hand, based on $(*)$ and the fact that $b_i U \subseteq H$, we deduce from Corollary 3.2 that

$$w_I(g_i b_i U; U) = w_I(g_i; U). \qquad (3.2)$$

From (3.1) and (3.2) we conclude that $w_I(g_i; U) = 1$. □

Lemma 3.4. *Suppose that $|G_w| < 2^{\aleph_0}$. Suppose further that $H \trianglelefteq_{\mathrm{o}} G$ satisfies $w(H) = 1$. Then G_w is finite.*

Proof. Below we construct $V \trianglelefteq_{\mathrm{o}} G$ such that

$$w_J(G; V) = 1 \qquad \text{for every proper subset } J \subsetneqq \{1, \ldots, r\}. \tag{3.3}$$

Let S be a transversal, i.e., a set of coset representatives, for V in G. From (3.3) and Corollary 3.2 we deduce that

$$w(g_1v_1, \ldots, g_rv_r) = w(g_1, \ldots, g_r) \quad \text{for all } g_1, \ldots, g_r \in S \text{ and all } v_1, \ldots, v_r \in V.$$

Since $G = \bigcup\{gV \mid g \in S\}$, this shows that $G_w = \{w(g_1, \ldots, g_r) \mid g_1, \ldots, g_r \in S\}$ is finite.

It remains to produce $V \trianglelefteq_{\mathrm{o}} G$ such that (3.3) holds. Indeed, we prove for $I \subsetneqq \{1, \ldots, r\}$, by induction on $|I|$, that there exists $U_I \trianglelefteq_{\mathrm{o}} G$ such that $w_I(G; U_I) = 1$. The group V then results from intersecting the finitely many groups U_I, where $I \subsetneqq \{1, \ldots, r\}$.

Let $I \subsetneqq \{1, \ldots, r\}$. If $I = \varnothing$ then $U_\varnothing = H$ satisfies $w_\varnothing(G; U_\varnothing) = w(H) = 1$. Now suppose that $|I| \geq 1$. For each $J \subsetneqq I$ induction yields $U_J \trianglelefteq_{\mathrm{o}} G$ such that $w_J(G; U_J) = 1$. Then $U = \bigcap\{U_J \mid J \subsetneqq I\} \trianglelefteq_{\mathrm{o}} G$ satisfies

$$w_J(G; U) = 1 \qquad \text{for every proper subset } J \subsetneqq I. \tag{3.4}$$

Let R be a transversal for U in G. For each family $\mathbf{g} = (g_i)_{i \in I}$ in R, Lemma 3.3 yields $U_{\mathbf{g}} \trianglelefteq_{\mathrm{o}} G$, with $U_{\mathbf{g}} \subseteq U$, such that $w_I(g_i; U_{\mathbf{g}}) = 1$. Intersecting the finitely many groups $U_{\mathbf{g}}$, parametrised by $\mathbf{g}$, we obtain $U_I \trianglelefteq_{\mathrm{o}} G$, with $U_I \subseteq U$, such that

$$w_I(g_i; U_I) = 1 \qquad \text{for all families } \mathbf{g} = (g_i)_{i \in I} \text{ in } R.$$

From (3.4) and Corollary 3.2 we deduce that

$$w_I(g_iU; U_I) = w_I(g_i; U_I) = 1 \qquad \text{for all families } \mathbf{g} = (g_i)_{i \in I} \text{ in } R.$$

Since $G = \bigcup\{gU \mid g \in R\}$, this shows that $w_I(G; U_I) = \langle \bigcup_{\mathbf{g}} w_I(g_iU; U_I)\rangle = 1$. □

With these preparations we are ready to prove Theorem 1.1.

Proof of Theorem 1.1. Recall that $w = w(x_1, \ldots, x_r)$ is a multilinear commutator word and that G is a profinite group such that $|G_w| < 2^{\aleph_0}$. Clearly, we may assume that $r \geq 1$. By Proposition 2.1 and Corollary 3.1, there exists $H \trianglelefteq_{\mathrm{o}} G$ such that $w(H) = 1$. Thus Lemma 3.4 shows that $|G_w| < \infty$ and the claim follows from [**19**, Theorem 1] (or [**2**, Theorem 1.1]). □

4. *The case where $w(G)$ is generated by finitely many w-values*

In theory, the task of establishing the strong conciseness of a group word w for a class $\mathcal{C}$ of profinite groups can be divided into two steps: Given w and a profinite group G in $\mathcal{C}$ such that $|G_w| < 2^{\aleph_0}$, it suffices to show that

- ∘ $w(G)$ is generated by finitely many w-values and
- ∘ using this extra information, the group $w(G)$ is finite.

If $w(G)$ is a pro-p group, for some prime p, the situation simplifies further: the verbal subgroup $w(G)$ is generated by finitely many w-values if and only if it is finitely generated. Indeed, it suffices to look at the Frattini quotient of $w(G)$, an elementary abelian pro-p group. In addition, we have the following useful lemma.

LEMMA 4.1. *Let w be a group word and let G be a profinite group such that $|G_w| < 2^{\aleph_0}$. Suppose that $N = w(G)$ is a pro-p group, for some prime p, and that $N/[N,G]N^p$ is finite. Then $w(G)$ is generated by finitely many w-values.*

Proof. Let X be a finite set of w-values such that $N = \langle X\rangle[N,G]N^p$. Since N is a pro-p group, the set $\{x^g \mid x \in X, g \in G\}$ generates N modulo $N \cap K$ for every open normal subgroup $K \trianglelefteq_{\mathrm{o}} G$. Hence $N = \langle x^g \mid x \in X, g \in G\rangle$, and from Lemma 2.2 we conclude that $\{x^g \mid x \in X, g \in G\}$ is finite. □

We recall that a group word w has finite width in an abstract group G if there exists $m \in \mathbb{N}$ such that every element $g \in w(G)$ can be written as a product $g = g_1 \cdots g_m$, where each g_i is a w-value or the inverse of a w-value in G. This notion extends naturally to profinite groups. If w has finite width in a profinite group G, then $w(G)$ coincides with the abstract subgroup generated by G_w; see [**15**, Proposition 4.1.2].

COROLLARY 4.2. *Let w be a group word and let G be a profinite group such that $|G_w| < 2^{\aleph_0}$. Suppose that $w(G)$ is a pro-p group, for some prime p, and that w has finite width in every finitely generated subgroup H of G. Then $N = w(G)$ is finite if and only if $N/[N,G]N^p$ is finite.*

Proof. Suppose that $N/[N,G]N^p$ is finite. By Lemma 4.1, N is generated as a subgroup by finitely many w-values. Thus we may further suppose that G is finitely generated. By our assumptions, w has finite width in G. Hence $|w(G)| < 2^{\aleph_0}$ and $w(G)$ is finite. □

We now extend our considerations to general profinite groups G, but impose a priori the condition that $w(G)$ is generated by finitely many w-values.

LEMMA 4.3. *Let w be a group word and let G be a profinite group such that $|G_w| < 2^{\aleph_0}$. Suppose that $w(G)$ is generated by finitely many w-values. Then the commutator subgroup of $w(G)$ is finite.*

Proof. Suppose that $w(G) = \langle g_1, \ldots, g_r\rangle$ for $g_1, \ldots, g_r \in G_w$. Lemma 2.2 implies that $\mathrm{C}_G(g_1), \ldots, \mathrm{C}_G(g_r)$ are open in G. Therefore $\mathrm{C}_G(w(G)) = \bigcap_{i=1}^n \mathrm{C}_G(g_i) \le_{\mathrm{o}} G$ and $G/\mathrm{C}_G(w(G))$ is finite. By Schur's Theorem (see [**13**, p. 102]), the commutator subgroup of $w(G)$ is finite. □

For $n, k \in \mathbb{N}_0$, the nth power of the derived word δ_k is written as δ_k^n. We say that a quantity is $(a, b, c, \ldots)$-bounded if it can be bounded from above by a number depending only on the specified parameters $a, b, c, \ldots$.

LEMMA 4.4. *[**17**, Lemma 3.2] Let $k, n, t \in \mathbb{N}$. Let G be a group satisfying $\delta_k^n(G) = 1$. Let H be a nilpotent subgroup of G generated by a set of δ_k-values and suppose, in addition, that H is t-generated. Then the order of H is (k, n, t)-bounded.*

LEMMA 4.5. *[**5**, Lemma 2.1] Let $d, k \in \mathbb{N}$. There exists a number $t = t(d, k)$, depending on d and k only, such that, if G is a finite d-generated group, then every δ_{k-1}-value in elements of G' is a product of at most t elements that are δ_k-values in elements of G.*

LEMMA 4.6. *[**5**, Lemma 2.2] Let G be a soluble group of derived length l, and suppose that X is a symmetric, normal and commutator-closed set of generators for G. Let g be an arbitrary element of G, written as $g = x_1 \cdots x_t$, where $x_i \in X$ for all $i \in \{1, \dots t\}$. Then, for every $n \in \mathbb{N}$, we have*

$$g^{n^l} = y_1^n \cdots y_s^n,$$

where $y_1, \dots, y_s \in X$ and s is (n, t, l)-bounded.

PROPOSITION 4.7. *Let $w = \delta_k^n$, where $k, n \in \mathbb{N}_0$. Let G be a profinite group such that $|G_w| < 2^{\aleph_0}$. Suppose that the kth derived subgroup $G^{(k)}$ is pronilpotent and that $w(G)$ is finitely generated. Then $w(G)$ is finite.*

Proof. We argue by induction on n. For $n \leq 1$, the result is immediate from Theorem 1.1. Now suppose that $n \geq 2$. Let π be the set of prime divisors of n. If $G^{(k)}$ is a pro-p' group for some $p \in \pi$, then for $v = \delta_k^{n/p}$ the map $x \mapsto x^p$ provides a bijection from G_v onto G_w and, by induction $w(G) = v(G)$ is finite. Hence, we may suppose that $G^{(k)}$ has non-trivial Sylow pro-p subgroup for each $p \in \pi$. Moreover, if P is the Sylow pro-p subgroup of $G^{(k)}$ for some $p \in \pi$, then the image of $w(G)$ in G/P is finite. Suppose that $|\pi| \geq 2$, and let P_1 and P_2 be the Sylow subgroups of $G^{(k)}$ for distinct primes $p_1, p_2 \in \pi$. Then images of $w(G)$ in G/P_1 and G/P_2 are finite, and from $P_1 \cap P_2 = 1$ we deduce that $w(G)$ is finite.

Thus, it is sufficient to deal with the case where n is a p-power for some prime p. Passing to the quotient $G/O_{p'}(G^{(k)})$, we may suppose that $G^{(k)}$ is a pro-p group. Since $w(G)$ is a finitely generated pro-p group, it is actually generated by finitely many w-values. By Lemma 4.3, the commutator subgroup of $w(G)$ is finite. Passing to the quotient $G/w(G)'$, we may suppose that $w(G)$ is abelian. If $k = 0$, we deduce from Lemma 2.4 that G is periodic, hence $w(G)$ is finite.

Suppose that $k \geq 1$. Since $w(G)$ is generated by finitely many w-values, we may choose finitely many elements $g_1, \dots, g_d \in G$ such that $w(\langle g_1, \dots, g_d \rangle) = w(G)$. It is sufficient to work with $\langle g_1, \dots, g_d \rangle$ in place of G and so without loss of generality we suppose that G is finitely generated, by d elements, say. By Lemma 4.5 there exists a number $t = t(d, k)$, depending on d and k only, such that every δ_{k-1}-value in elements of G' is a product of at most t elements which are δ_k-values in elements of G.

Consider a subgroup $H = \langle x_1, \dots, x_t \rangle$, where $x_1, \dots, x_t$ are δ_k-values in G. By Lemma 4.4, applied to finite quotients of $G/w(G)$, every finite quotient of $H/(H \cap w(G))$ and hence the entire group $H/(H \cap w(G))$ is finite of (k, n, t)-bounded order. In particular, H is soluble of derived length at most l, where $l = l(k, n, t)$ depends on k, n, t only.

Set $v = (\delta_{k-1})^{n^l}$. By Lemma 4.6, every v-value in elements of G' is a product of an (n, t, l)-bounded number of w-values and inverses of w-values. This gives $|(G')_v| < 2^{\aleph_0}$ and, by induction on k, the verbal subgroup $v(G')$ is finite.

Passing to the quotient $G/v(G')$, we may suppose that δ_{k-1}-values in elements of G' are of finite order. Then also δ_k-values in elements of G are of finite order. As $w(G)$ is abelian and finitely generated, we conclude that $w(G)$ is finite. □

Recall that a group word w is a *law* in a group G if $w(G) = 1$. We say that w *implies virtual nilpotency* if every finitely generated metabelian group for which w is a law has a nilpotent subgroup of finite index. Burns and Medvedev [**1**] showed that if w implies virtual nilpotency, then for a much larger class of groups G, including all finitely generated residually finite groups, $w(G) = 1$ implies that G is nilpotent-by-finite. Moreover, the word w implies virtual nilpotency if and only if, for all primes p, the word w is not a law in the wreath product $C_p \wr C_\infty$ of the cyclic group of order p by the infinite cyclic group; see [**1**]. In particular, every word of the form uv^{-1}, where u and v are positive words (i.e. semigroup words in finitely many free generators), implies virtual nilpotency. Furthermore, by a result of Gruenberg [**6**], all Engel words imply virtual nilpotency. Other examples of words implying virtual nilpotency include generalisations of Engel words, such as words of the form $w = w(x, y) = [x^{e_1}, y^{e_2}, \ldots, y^{e_r}]$, where $r \in \mathbb{N}$ and $e_1, \ldots, e_r \in \mathbb{Z} \smallsetminus \{0\}$. To see that such a word implies virtual nilpotency, we employ the criterion of Burns and Medvedev. The case $r = 1$ is easy; now suppose that $r \geq 2$. Let p be a prime and consider the wreath product

$$C_p \wr C_\infty = \langle a, t \mid a^p = 1, [a^{t^i}, a^{t^j}] = 1 \text{ for } i, j \in \mathbb{Z} \rangle \xrightarrow{\cong} \mathbb{F}_p[T, T^{-1}] \rtimes \langle T \rangle,$$

where $a \mapsto 1 \in \mathbb{F}_p[T, T^{-1}]$ in the base group and $t \mapsto T \in \langle T \rangle \leq \mathbb{F}_p[T, T^{-1}]^*$ in the top group. We may suppose that $e_1 > 0$. Then the indicated isomorphism maps $w(ta, t)$ to

$$(T^{e_1 - 1} + \ldots + T + 1)(T^{e_2} - 1) \cdots (T^{e_r} - 1) \neq 0,$$

in the base group. Thus $w(ta, t) \neq 1$ and w is not a law in $C_p \wr C_\infty$.

Proposition 4.8. *Let w be a word implying virtual nilpotency and let G be a profinite group such that $|G_w| < 2^{\aleph_0}$. If the verbal subgroup $w(G)$ is generated by finitely many w-values, then $w(G)$ is finite.*

Proof. Without loss of generality we may assume that G is finitely generated. Using Lemma 4.3, we may further assume that $w(G)$ is abelian. Clearly, w is a law in $G/w(G)$. Hence [**1**, Theorem A] shows that $G/w(G)$ is nilpotent-by-finite. Thus G is abelian-by-nilpotent-by-finite. Every word has finite width in every finitely generated abelian-by-nilpotent-by-finite group; compare [**15**, Theorem 4.1.5].

Thus $w(G)$ has less than $2^{\aleph_0}$ elements, hence it is finite. □

Following [**7**] we say that a group word w is *weakly rational* if for every finite group G and for every positive integer e with $\gcd(e, |G|) = 1$, the set G_w is closed under taking eth powers of its elements. By [**7**, Lemma 1], the word w is weakly rational if and only if for every finite group G, every $g \in G_w$ and every $e \in \mathbb{N}$ with $\gcd(e, |\langle g \rangle|) = 1$ we have $g^e \in G_w$. According to [**7**, Theorem 3], the word $w = [x_1, \ldots, x_r]^q$ is weakly rational for all $r, q \in \mathbb{N}$.

Proposition 4.9. *Let w be a weakly rational word and let G be a profinite group such that $|G_w| < 2^{\aleph_0}$. If the verbal subgroup $w(G)$ is generated by finitely many w-values, then $w(G)$ is finite.*

Proof. By Lemma 4.3 we may suppose that $w(G)$ is abelian, and it suffices to show that elements of G_w have finite order.

Let $h \in G_w$, and let g be any generator of the procyclic group $H = \langle h \rangle$. For every $N \trianglelefteq_o G$, there exists $e \in \mathbb{N}$ with $\gcd(e, |HN/N|) = 1$ such that $g \equiv_N h^e$ and, because w is weakly rational, we obtain $g \in G_w N$. Hence $g \in \bigcap_{N \trianglelefteq_o G} G_w N = G_w$. Therefore the procyclic group H has less than $2^{\aleph_0}$ single generators.

This implies that H is finite. Indeed, consider the Frattini subgroup $\Phi(H)$ of H. Since $\langle g \rangle = H$ for every $g \in h\Phi(H)$, the group $\Phi(H)$ has less than $2^{\aleph_0}$ elements. Hence $\Phi(H)$ is finite, and without loss of generality we assume that $\Phi(H) = 1$. Then $H \cong \prod_{p \in \pi} C_p$ for a set of primes π. Each factor C_p has $p - 1$ single generators. Since H has less than $2^{\aleph_0}$ single generators, π and hence H is finite. □

5. *Reduction via parametrised words*

Throughout this section, we fix a profinite group G, a positive integer $r \in \mathbb{N}$ and a normal subgroup $\mathbf{G} \trianglelefteq G \times \ldots \times G$ of the direct product of r copies of G. A typical situation would be $\mathbf{G} = G_1 \times \cdots \times G_r$, where $G_1, \ldots, G_r \trianglelefteq G$.

Our intention is to consider (products of) 'parametrised group words' in variables $x_1, \ldots, x_r$, with parameters coming from G where each x_i is intended to take values in G_i and where we formally distinguish repeated occurrences of the same variable. This elementary concept requires a flexible but precise set-up.

Let $\Omega = \Omega_{G,r}$ be the free group on free generators

$$\xi_h, \quad \eta_{1,i}, \eta_{2,i}, \ldots, \eta_{r,i} \qquad \text{for } h \in G \text{ and } i \in \mathbb{N}.$$

Informally, we think of each free generator ξ_h as a 'parameter variable' that is to take the value h and each free generator $\eta_{q,i}$ as a 'free variable' that can be specialised to x_q, irrespective of the additional index i.

We refer to elements $\omega \in \Omega$ as *r-valent parametrised words* for G or, since r is fixed throughout, simply as *parametrised words* for G. For $\mathbf{g} = (g_1, \ldots, g_r) \in \mathbf{G}$, we write

$$\underline{\omega}(\mathbf{g}) = \underline{\omega}_{\mathbf{G}}(\mathbf{g}) = \underline{\omega}(g_1, \ldots, g_r) \in G$$

for the ω-value that results from replacing each ξ_h by h and each $\eta_{q,i}$ by g_q, for all $h \in G$, $q \in \{1, \ldots, r\}$ and $i \in \mathbb{N}$. In this way we obtain a parametrised word map $\underline{\omega}(\cdot) \colon \mathbf{G} \to G$.

The *degree* $\deg(\omega)$ of the parametrised word ω is the number of free generators $\eta_{q,i}$, with $q \in \{1, \ldots, r\}$ and $i \in \mathbb{N}$, appearing in (the reduced form of) ω; here we care whether a generator $\eta_{q,i}$ appears, but not whether it appears repeatedly. The degree $\deg(\omega)$ is a non-negative integer and plays a role in defining appropriate induction parameters. We remark that, if ω has degree 0, then the map $\underline{\omega}(\cdot)$ is constant, i.e. there exists $h \in G$ such that for all $\mathbf{g} \in \mathbf{G}$ we have $\underline{\omega}(\mathbf{g}) = h$.

EXAMPLE 5.1. Our main interest will be in iterated commutator words, such as $w(x_1, x_2, x_3) = [[x_1, x_2, x_2], [x_2, x_3]]$, and the w-values in a profinite group G. We set $r = 3$, $\mathbf{G} = G \times G \times G$ and $\omega = [[\eta_{1,1}, \eta_{2,1}, \eta_{2,2}], [\eta_{2,3}, \eta_{3,1}]]$ to model w in the sense that

$$w(g_1, g_2, g_3) = [[g_1, g_2, g_2], [g_2, g_3]] = \underline{\omega}(g_1, g_2, g_3) \quad \text{for all } g_1, g_2, g_3 \in G.$$

The 3-valent parametrised word ω has degree 5; moreover, ω is a multilinear commutator word of weight 5 (meaning that it involves 5 variables). In this example, we are not yet using the possibility to involve parameters.

We fix a set $\mathfrak{E} = \mathfrak{E}_{G,r} \subseteq \Omega$ of r-valent parametrised words for G, which we think of as 'elementary' words, and we consider finite products of such. To write down these products we use finite index sets $T, S, \ldots \subseteq \mathbb{N}$ that are implicitly ordered so that the products are unambiguous in a typically non-commutative setting.

Formally, an r-valent *$\mathfrak{E}$-product* for G is a finite sequence $(\varepsilon_t)_{t \in T}$, where $\varepsilon_t \in \mathfrak{E}$ for each $t \in T$; more suggestively, we denote it by

$$\dot{\prod}_{t \in T} \varepsilon_t,$$

where the dot indicates that we consider a formal product and not the parametrised word that results from actually carrying out the multiplication in Ω.

By a *length function* on $\mathfrak{E}$ we mean any map $\ell \colon \mathfrak{E} \to \mathsf{W}$ from $\mathfrak{E}$ into a well-ordered set $\mathsf{W} = (\mathsf{W}, \leq)$ such that elements $\varepsilon \in \mathfrak{E}$ whose length $\ell(\varepsilon)$ is minimal with respect to $\leq$ also have minimal degree $\deg(\varepsilon) = 0$. As usual, we agree that the maximum of the empty subset of W is the least element of W. A length function ℓ induces a total pre-order $\preceq_\ell$ on the set of all

r-valent $\mathfrak{E}$-products, as follows:

$$\dot{\prod}_{s\in S}\widetilde{\varepsilon}_s \preceq_\ell \dot{\prod}_{t\in T}\varepsilon_t \qquad \text{if} \quad \max\{\ell(\widetilde{\varepsilon}_s) \mid s\in S\} \le \max\{\ell(\varepsilon_t)\mid t\in T\};$$

we write $\dot{\prod}_{s\in S}\widetilde{\varepsilon}_s \prec_\ell \dot{\prod}_{t\in T}\varepsilon_t$ if $\max\{\ell(\widetilde{\varepsilon}_s) \mid s\in S\} < \max\{\ell(\varepsilon_t)\mid t\in T\}$. Clearly, there are no infinite descending chains of $\mathfrak{E}$-products, with respect to $\prec_\ell$. This fact allows us to give the following recursive definition.

Definition 5.2 Friendly products. Let $\ell\colon \mathfrak{E}\to \mathsf{W}$ be a length function. We define recursively the set $\mathfrak{F} = \mathfrak{F}_{\mathbf{G},r,\ell}$ of *ℓ-friendly* r-valent $\mathfrak{E}$-products for $\mathbf{G}$ as follows. An r-valent $\mathfrak{E}$-product $\dot{\prod}_{t\in T}\varepsilon_t$ for G belongs to $\mathfrak{F}$ if for every $\mathbf{b}\in\mathbf{G}$ there exists an r-valent $\mathfrak{E}$-product $\dot{\prod}_{s\in S(\mathbf{b})}\widetilde{\varepsilon}_{\mathbf{b},s}$ such that

(F1) $\dot{\prod}_{s\in S(\mathbf{b})}\widetilde{\varepsilon}_{\mathbf{b},s}$ belongs to $\mathfrak{F}$ and $\dot{\prod}_{s\in S(\mathbf{b})}\widetilde{\varepsilon}_{\mathbf{b},s} \prec_\ell \dot{\prod}_{t\in T}\varepsilon_t$ and

(F2) the parametrised words $\omega = \prod_{t\in T}\varepsilon_t$ and $\nu_{\mathbf{b}} = \prod_{s\in S(\mathbf{b})}\widetilde{\varepsilon}_{\mathbf{b},s}$ satisfy

$$\underline{\omega}(\mathbf{bg}) = \underline{\omega}(\mathbf{b})\cdot\underline{\omega}(\mathbf{g})\cdot\underline{\nu}_{\mathbf{b}}(\mathbf{g}) \qquad \text{for all } \mathbf{g}\in\mathbf{G}. \tag{5.1}$$

Remark 5.3. (1) In the definition, the product $\dot{\prod}_{s\in S(\mathbf{b})}\widetilde{\varepsilon}_{\mathbf{b},s}$ is allowed to be empty, in which case (5.1) simplifies to

$$\underline{\omega}(\mathbf{bg}) = \underline{\omega}(\mathbf{b})\cdot\underline{\omega}(\mathbf{g}). \tag{5.2}$$

Such a strong relation holds, for instance, if the parametrised word $\omega = \prod_{t\in T}\varepsilon_t$ has degree 1 and defines a homomorphism $\mathbf{G}\to G$ that factors through the qth coordinate, if the single free variable occurring in ω is $\eta_{q,i}$ for some $i\in\mathbb{N}$. In this special situation, (5.2) holds uniformly for all $\mathbf{b}\in\mathbf{G}$.

(2) If the ℓ-friendly $\mathfrak{E}$-product $\dot{\prod}_{t\in T}\varepsilon_t$ is minimal with respect to $\preceq_\ell$, then $\dot{\prod}_{s\in S(\mathbf{b})}\widetilde{\varepsilon}_{\mathbf{b},s}$ is necessarily empty for every choice of $\mathbf{b}\in\mathbf{G}$. Furthermore, each ε_t has degree $\deg(\varepsilon_t)=0$, so there is $h\in G$ such that for all $\mathbf{g}\in\mathbf{G}$ we have $\underline{\omega}(\mathbf{g}) = \prod_{t\in T}\underline{\varepsilon}_t(\mathbf{g}) = h$. Thus (5.1) yields $h = h\cdot h\cdot 1 = h^2$, and hence $h=1$.

In this sense there is only one parametrised word map coming from an ℓ-friendly $\mathfrak{E}$-product for $\mathbf{G}$ that is minimal with respect to $\preceq_\ell$, namely the constant map with value 1. In particular, for every ℓ-friendly $\mathfrak{E}$-product $\dot{\prod}_{t\in T}\varepsilon_t$ that is second smallest with respect to $\preceq_\ell$, the parametrised word $\omega = \prod_{t\in T}\varepsilon_t$ satisfies (5.2).

Remark 5.4. In this paper we use the terminology introduced above in the context of nilpotent groups. We indicate how the general set-up specialises.

Let G be a nilpotent profinite group of class at most c, i.e. $\gamma_{c+1}(G)=1$. Denote by $\mathfrak{E}$ the set of all left-normed repeated commutators in the free generators ξ_h and $\eta_{q,i}$ of Ω, subject to the restriction that each $\eta_{q,i}$ appears at most once. In other words, $\mathfrak{E}$ consists of all γ_m-values, for $m\ge 2$, that result from replacing the m variables in $\gamma_m = [x_1,x_2,\ldots,x_m]$ by arbitrary free generators ξ_h and $\eta_{q,i}$ of Ω, subject to the restriction that each $\eta_{q,i}$ appears at most once.

For instance, given some element $a\in G$,

$$\varepsilon_1 = [\xi_a,\eta_{1,1},\xi_a,\eta_{2,1},\eta_{2,2},\eta_{2,3}] \in \mathfrak{E},$$

whereas $\varepsilon_2 = [\xi_a,\eta_{1,1},\xi_a,\eta_{2,1},\eta_{2,2},\eta_{2,1}]$ does not lie in $\mathfrak{E}$, even though $\underline{\varepsilon}_1(\cdot) = \underline{\varepsilon}_2(\cdot)$. We set $\mathsf{W} = \mathbb{N}_0\times\mathbb{N}_0$, equipped with the lexicographic order $\le$; so, for instance, $(2,10)\le(3,0)$ and $(5,4)\le(5,7)$.

Every $\varepsilon\in\mathfrak{E}$, by definition, belongs to $\gamma_2(\Omega)$. Let $k(\varepsilon)$ denote the maximal $j\in\{1,\ldots,c+1\}$ such that $\varepsilon\in\gamma_j(\Omega)$, and define a length function on $\mathfrak{E}$ by associating to ε the length

$$\ell(\varepsilon) = \big(c+1-k(\varepsilon), \deg(\varepsilon)\big) \in \mathsf{W}.$$

For instance, if $c = 8$ then $\ell(\varepsilon_1) = (8+1-6, 4) = (3, 4)$.

Lemma 5.5. *Suppose that the profinite group G is nilpotent of class at most c, and let $\mathfrak{E}$ be defined as in Remark 5.4. Let $\omega \in \gamma_k(\Omega)$, where $k \geq 2$, be such that $\underline{\omega}(\mathbf{1}) = 1$. Then there exists an $\mathfrak{E}$-product $\dot{\prod}_{t \in T} \varepsilon_t$, with $\varepsilon_t \in \gamma_k(\Omega)$ and $\deg(\varepsilon_t) \geq 1$ for all $t \in T$, such that*

$$\underline{\omega}(\mathbf{g}) = \prod\nolimits_{t \in T} \underline{\varepsilon}_t(\mathbf{g}) \qquad \text{for all } \mathbf{g} \in \mathbf{G}.$$

Proof. It suffices to prove, by induction on $c+1-k$, that

$$\omega \equiv \prod_{\substack{t \in T \text{ s.t.} \\ \deg(\varepsilon_t)=0}} \varepsilon_t \cdot \prod_{\substack{t \in T \text{ s.t.} \\ \deg(\varepsilon_t)\neq 0}} \varepsilon_t \mod \gamma_{c+1}(\Omega). \tag{5.3}$$

for a suitable index set T and suitable left-normed repeated commutators $\varepsilon_t = [\varkappa_{t,1}, \ldots, \varkappa_{t,n(t)}]$, where $n(t) \geq k$ and the terms $\varkappa_{t,j}$ stand for suitable free generators ξ_h and $\eta_{q,i}$ of Ω. Indeed, using the infinite supply of generators $\eta_{q,i}$, we can rename the free generators entering into the commutators ε_t to ensure that $\varepsilon_t \in \mathfrak{E}$, without changing the resulting word maps $\underline{\varepsilon}_t(\cdot)$. Furthermore, the identity

$$1 = \underline{\omega}(\mathbf{1}) = \prod_{\substack{t \in T \text{ s.t.} \\ \deg(\varepsilon_t)=0}} \underline{\varepsilon}_t(\mathbf{1}) \cdot \prod_{\substack{t \in T \text{ s.t.} \\ \deg(\varepsilon_t)>0}} \underline{\varepsilon}_t(\mathbf{1}) = \prod_{\substack{t \in T \text{ s.t.} \\ \deg(\varepsilon_t)=0}} \underline{\varepsilon}_t(\mathbf{1})$$

and the fact that $\underline{\varepsilon}_t(\cdot)$ is constant whenever $\deg(\varepsilon_t) = 0$, shows that

$$\underline{\omega}(\mathbf{g}) = \prod_{\substack{t \in T \text{ s.t.} \\ \deg(\varepsilon_t)>0}} \underline{\varepsilon}_t(\mathbf{g}) \qquad \text{for all } \mathbf{g} \in \mathbf{G}.$$

For $k = c+1$ the congruence (5.3) holds upon setting $T = \varnothing$. Now suppose that $k < c+1$. As $\omega \in \gamma_k(\Omega)$ is a product of γ_k-values in Ω, basic commutator manipulations (compare [**15**, Proposition 1.2.1]) yield that ω can be written as a product

$$\omega = \prod\nolimits_{t \in T(1)} \varepsilon_t \cdot \nu$$

of repeated commutators ε_t of the form $[\varkappa_{t,1}, \ldots, \varkappa_{t,k}]$ or $[\varkappa_{t,1}, \ldots, \varkappa_{t,k}]^{-1}$, where the terms $\varkappa_{t,j}$ stand for suitable free generators of Ω, and an element $\nu \in \gamma_{k+1}(\Omega)$. Modulo $\gamma_{k+1}(\Omega)$, the basic relation

$$[\varkappa_1, \varkappa_2, \varkappa_3, \ldots, \varkappa_k]^{-1} \equiv [[\varkappa_1, \varkappa_2]^{-1}, \varkappa_3, \ldots, \varkappa_k] \equiv [\varkappa_2, \varkappa_1, \varkappa_3, \ldots, \varkappa_k]$$

holds; thus we can even avoid using terms with exponent -1.

By induction, ν can be written as a product

$$\nu \equiv \prod\nolimits_{t \in T(2)} \varepsilon_t \mod \gamma_{c+1}(\Omega).$$

of suitable repeated commutators. Denote by T the ordinal sum of $T(1)$ and $T(2)$, i.e. the disjoint union equipped with the total order in which every $t_1 \in T(1)$ precedes every $t_2 \in T(2)$ and where $T(1)$ and $T(2)$ are ordered as before. This yields

$$\omega \equiv \prod\nolimits_{t \in T} \varepsilon_t \mod \gamma_{c+1}(\Omega).$$

At the expense of creating extra factors of degree at least 1, we can rearrange the factors in the product so that, after enlarging the index set and renaming the relevant factors, we arrive at (5.3). □

LEMMA 5.6. *Suppose that the profinite group G is nilpotent of class at most c, and let $\mathfrak{E}$ be defined as in Remark 5.4. Let $k \geq 2$ and let $\dot{\prod}_{t\in T} \varepsilon_t$ be an $\mathfrak{E}$-product, where $\varepsilon_t \in \gamma_k(\Omega)$ and $\deg(\varepsilon_t) \geq 1$ for all $t \in T$. Let $\mathbf{b} = (b_1, \ldots, b_r) \in \mathbf{G}$. Then there exists a parametrised word $\nu = \nu_{\mathbf{b}} \in \gamma_{k+1}(\Omega)$ such that*

$$\prod_{t\in T} \underline{\varepsilon}_t(\mathbf{bg}) = \prod_{t\in T} \underline{\varepsilon}_t(\mathbf{b}) \cdot \prod_{t\in T} \underline{\varepsilon}_t(\mathbf{g}) \cdot \prod_{t\in T} \underline{\tilde{\varepsilon}}_t(\mathbf{b}/\!\!/\mathbf{g}) \cdot \underline{\nu}(\mathbf{g}) \qquad \text{for all } \mathbf{g} \in \mathbf{G},$$

where $\tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot)$ denotes the $\mathfrak{E}$-product involving (in some implicit order) the $2^{\deg(\varepsilon_t)} - 2$ factors that result from ε_t by replacing a selection of at least one, but not all distinct free variables $\eta_{q,i}$ occurring in ε_t by ξ_{b_q}. Moreover,

$$\dot{\prod}_{t\in T} \tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot) \prec_\ell \dot{\prod}_{t\in T} \varepsilon_t \qquad \text{and} \qquad \underline{\nu}(\mathbf{1}) = 1.$$

Proof. As $\deg(\varepsilon_t) \geq 1$, basic commutator manipulations (compare [**15**, Proposition 1.2.1]) yield that, for each $t \in T$, there exists $\nu_t \in \gamma_{k+1}(\Omega)$ such that

$$\underline{\varepsilon}_t(\mathbf{bg}) = \underline{\varepsilon}_t(\mathbf{b})\, \underline{\varepsilon}_t(\mathbf{g})\, \underline{\tilde{\varepsilon}}_t(\mathbf{b}/\!\!/\mathbf{g})\, \underline{\nu}_t(\mathbf{g}) \qquad \text{for all } \mathbf{g} \in \mathbf{G}.$$

Moreover, by construction each factor of the $\mathfrak{E}$-product $\tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot)$ has degree at least 1, hence $\underline{\tilde{\varepsilon}}_t(\mathbf{b}/\!\!/\mathbf{1}) = 1$.

All the words ε_t, $\tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot)$ and ν_t, for $t \in T$, commute with one another modulo $\gamma_{k+1}(\Omega)$. Hence there exists $\nu \in \gamma_{k+1}(\Omega)$ such that, for all $\mathbf{g} \in \mathbf{G}$,

$$\begin{aligned} \prod_{t\in T} \underline{\varepsilon}_t(\mathbf{bg}) &= \prod_{t\in T} \big(\underline{\varepsilon}_t(\mathbf{b})\, \underline{\varepsilon}_t(\mathbf{g})\, \underline{\tilde{\varepsilon}}_t(\mathbf{b}/\!\!/\mathbf{g})\, \underline{\nu}_t(\mathbf{g})\big) \\ &= \prod_{t\in T} \underline{\varepsilon}_t(\mathbf{b}) \cdot \prod_{t\in T} \underline{\varepsilon}_t(\mathbf{g}) \cdot \prod_{t\in T} \underline{\tilde{\varepsilon}}_t(\mathbf{b}/\!\!/\mathbf{g}) \cdot \underline{\nu}(\mathbf{g}). \end{aligned} \tag{5.4}$$

Since every factor of $\tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot)$ has degree strictly smaller than ε_t, we deduce that

$$\dot{\prod}_{t\in T} \tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot) \prec_\ell \dot{\prod}_{t\in T} \varepsilon_t.$$

Finally, substituting $\mathbf{1}$ for $\mathbf{g}$ in (5.4), we see that $\underline{\nu}(\mathbf{1}) = 1$. □

LEMMA 5.7. *Suppose that the profinite group G is nilpotent of class at most c, and let $\mathfrak{E}$ and ℓ be defined as in Remark 5.4. Let $\omega \in \gamma_k(\Omega)$, where $k \geq 2$, be such that $\underline{\omega}(\mathbf{1}) = 1$. Then there exists an ℓ-friendly $\mathfrak{E}$-product $\dot{\prod}_{t\in T} \varepsilon_t$ such that*

$$\underline{\omega}(\mathbf{g}) = \prod_{t\in T} \underline{\varepsilon}_t(\mathbf{g}) \qquad \text{for all } \mathbf{g} \in \mathbf{G}.$$

Proof. We argue by induction on $c + 1 - k$. If $c + 1 - k \leq 0$, then $\underline{\omega}(\cdot)$ is the constant map with value 1, and the assertion holds trivially; compare Remark 5.3.

Now suppose that $c + 1 - k \geq 1$. We may suppose, in addition, that $\omega \notin \gamma_{k+1}(\Omega)$. By Lemma 5.5, there exists an $\mathfrak{E}$-product $\dot{\prod}_{t\in T} \varepsilon_t$, with $\varepsilon_t \in \gamma_k(\Omega)$ and $\deg(\varepsilon_t) \geq 1$ for all $t \in T$, such that

$$\underline{\omega}(\mathbf{g}) = \prod_{t\in T} \underline{\varepsilon}_t(\mathbf{g}) \qquad \text{for all } \mathbf{g} \in \mathbf{G}.$$

We claim that $\dot{\prod}_{t\in T} \varepsilon_t$ is ℓ-friendly; our task is to check the conditions laid out in Definition 5.2. We argue by induction with respect to the pre-order $\preceq_\ell$. Let $\mathbf{b} \in \mathbf{G}$. By Lemma 5.6, there exists a parametrised word $\nu \in \gamma_{k+1}(\Omega)$ such that

$$\prod_{t\in T} \underline{\varepsilon}_t(\mathbf{bg}) = \prod_{t\in T} \underline{\varepsilon}_t(\mathbf{b}) \cdot \prod_{t\in T} \underline{\varepsilon}_t(\mathbf{g}) \cdot \prod_{t\in T} \underline{\tilde{\varepsilon}}_t(\mathbf{b}/\!\!/\mathbf{g}) \cdot \underline{\nu}(\mathbf{g}) \qquad \text{for all } \mathbf{g} \in \mathbf{G}, \tag{5.5}$$

where the $\mathfrak{E}$-product $\dot{\prod}_{t\in T}\tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot)$ and ν satisfy

$$\dot{\prod}_{t\in T}\tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot) \prec_\ell \dot{\prod}_{t\in T}\varepsilon_t \qquad \text{and} \qquad \underline{\nu}(\mathbf{1}) = 1. \tag{5.6}$$

By Lemma 5.5 there exists an $\mathfrak{E}$-product $\dot{\prod}_{s\in S}\tilde{\varepsilon}_s$, with $\tilde{\varepsilon}_s \in \gamma_{k+1}(\Omega)$ and $\deg(\tilde{\varepsilon}_s) \geq 1$ for $s \in S$, such that

$$\underline{\nu}(\mathbf{g}) = \prod_{s\in S}\underline{\tilde{\varepsilon}}_s(\mathbf{g}) \qquad \text{for all } \mathbf{g} \in \mathbf{G}.$$

Consider the $\mathfrak{E}$-product $\dot{\prod}_{t\in T}\tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot)\dot{\prod}_{s\in S}\tilde{\varepsilon}_s$, formally based on the ordinal sum of T and S, and set $\tilde{\omega} = \prod_{t\in T}\tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot)\prod_{s\in S}\tilde{\varepsilon}_s$. From (5.6) and the fact that $\tilde{\varepsilon}_s \in \gamma_{k+1}(\Omega)$ we deduce that

$$\dot{\prod}_{t\in T}\tilde{\varepsilon}_t(\mathbf{b}/\!\!/\cdot)\dot{\prod}_{s\in S}\tilde{\varepsilon}_s \prec_\ell \dot{\prod}_{t\in T}\varepsilon_t.$$

Moreover, substituting $\mathbf{1}$ for $\mathbf{g}$ in (5.5), we obtain

$$\underline{\tilde{\omega}}(\mathbf{1}) = \prod_{t\in T}\underline{\tilde{\varepsilon}}_t(\mathbf{b}/\!\!/\mathbf{1})\cdot\underline{\nu}(\mathbf{1}) = \Big(\prod_{t\in T}\underline{\varepsilon}_t(\mathbf{1})\Big)^{-1} = \underline{\omega}(\mathbf{1})^{-1} = 1.$$

Hence, by induction with respect to $\preceq_\ell$, we conclude that $\dot{\prod}_{t\in T}\tilde{\varepsilon}_t(\mathbf{b}/\cdot)\dot{\prod}_{s\in S}\tilde{\varepsilon}_s$ is ℓ-friendly and thus all the conditions in Definition 5.2 are satisfied. □

LEMMA 5.8. *Let $\ell\colon \mathfrak{E} \to \mathsf{W}$ be a length function and let $\omega = \prod_{t\in T}\varepsilon_t$, where $\dot{\prod}_{t\in T}\varepsilon_t$ is an ℓ-friendly r-valent $\mathfrak{E}$-product for $\mathbf{G}$.*

Suppose that $\mathbf{V} \leq_{\mathrm{c}} \mathbf{G}$ is such that $\mathbf{V}_\omega = \{\underline{\omega}(\mathbf{v}) \mid \mathbf{v} \in \mathbf{V}\}$ contains less than $2^{\aleph_0}$ elements. Then there exists $\mathbf{U} \leq_{\mathrm{o}} \mathbf{V}$ such that

$$\underline{\omega}(\mathbf{u}) = 1 \qquad \textit{for all } \mathbf{u} \in \mathbf{U}.$$

Proof. We argue by induction, using the pre-order $\preceq_\ell$. If $\dot{\prod}_{t\in T}\varepsilon_t$ is minimal with respect to $\preceq_\ell$, the assertion holds for $\mathbf{U} = \mathbf{V}$, by Remark 5.3.

Now suppose that $\dot{\prod}_{t\in T}\varepsilon_t$ is not minimal. As $|\mathbf{V}_\omega| < 2^{\aleph_0}$, Proposition 2.1 implies that there are $\mathbf{b} \in \mathbf{V}$ and $\mathbf{U}_1 \leq_{\mathrm{o}} \mathbf{V}$ such that $\underline{\omega}(\cdot)$ is constant on the coset $\mathbf{b}\mathbf{U}_1$, i.e.

$$\underline{\omega}(\mathbf{b}) = \underline{\omega}(\mathbf{bu}) \qquad \text{for all } \mathbf{u} \in \mathbf{U}_1.$$

By Definition 5.2, we obtain

$$\underline{\omega}(\mathbf{bg}) = \underline{\omega}(\mathbf{b}) \cdot \underline{\omega}(\mathbf{g}) \cdot \underline{\nu}(\mathbf{g}) \qquad \text{for all } \mathbf{g} \in \mathbf{G},$$

where $\nu = \prod_{s\in S}\tilde{\varepsilon}_s$ for an ℓ-friendly r-valent $\mathfrak{E}$-product $\dot{\prod}_{s\in S}\tilde{\varepsilon}_s$ for $\mathbf{G}$ such that $\dot{\prod}_{s\in S}\tilde{\varepsilon}_s \prec_\ell \dot{\prod}_{t\in T}\varepsilon_t$. This yields

$$\underline{\nu}(\mathbf{u}) = \underline{\omega}(\mathbf{u})^{-1} \qquad \text{for all } \mathbf{u} \in \mathbf{U}_1;$$

in particular, $(\mathbf{U}_1)_\nu = \{\underline{\nu}(\mathbf{u}) \mid \mathbf{u} \in \mathbf{U}_1\}$ has less than $2^{\aleph_0}$ elements. By induction, we find the desired $\mathbf{U} \leq_{\mathrm{o}} \mathbf{U}_1 \leq_{\mathrm{o}} \mathbf{V}$ such that

$$\underline{\omega}(\mathbf{u}) = \underline{\nu}(\mathbf{u})^{-1} = 1 \qquad \text{for all } \mathbf{u} \in \mathbf{U}.$$

□

PROPOSITION 5.9. *Let $\ell\colon \mathfrak{E} \to \mathsf{W}$ be a length function and let $\omega = \prod_{t\in T}\varepsilon_t$, where $\dot{\prod}_{t\in T}\varepsilon_t$ is an ℓ-friendly r-valent $\mathfrak{E}$-product for $\mathbf{G}$. Suppose that $\mathbf{V} \leq_{\mathrm{c}} \mathbf{G}$ is such that $\mathbf{V}_\omega = \{\underline{\omega}(\mathbf{v}) \mid \mathbf{v} \in \mathbf{V}\}$ has less than $2^{\aleph_0}$ elements. Then $\mathbf{V}_\omega$ is already finite.*

Proof. We argue by induction, using the pre-order $\preceq_\ell$. If $\dot{\prod}_{t\in T}\varepsilon_t$ is minimal with respect to $\preceq_\ell$, the assertion holds, by Remark 5.3: indeed, $\mathbf{V}_\omega = \{1\}$.

Now suppose that $\dot{\prod}_{t\in T}\varepsilon_t$ is not minimal. By Lemma 5.8, there exists $\mathbf{U} \leq_{\mathrm{o}} \mathbf{V}$ such that

$$\underline{\omega}(\mathbf{u}) = 1 \qquad \text{for all } \mathbf{u} \in \mathbf{U}.$$

Let B be any set of coset representatives for $\mathbf{U}$ in $\mathbf{V}$ so that $|B| = |\mathbf{V} : \mathbf{U}| < \infty$.

By Definition 5.2, we see that, for each of the finitely many coset representatives $\mathbf{b} \in B$,

$$\underline{\omega}(\mathbf{bu}) = \underline{\omega}(\mathbf{b}) \cdot \underbrace{\underline{\omega}(\mathbf{u})}_{=1} \cdot \underline{\nu}_{\mathbf{b}}(\mathbf{u}) = \underbrace{\underline{\omega}(\mathbf{b})}_{\text{constant}} \underline{\nu}_{\mathbf{b}}(\mathbf{u}) \qquad \text{for } \mathbf{u} \in \mathbf{U},$$

where $\nu_{\mathbf{b}} = \prod_{s\in S(\mathbf{b})} \widetilde{\varepsilon}_{\mathbf{b},s}$ for an ℓ-friendly r-valent $\mathfrak{E}$-product $\dot{\prod}_{s\in S(\mathbf{b})} \widetilde{\varepsilon}_{\mathbf{b},s}$ for $\mathbf{G}$ such that $\dot{\prod}_{s\in S(\mathbf{b})} \widetilde{\varepsilon}_{\mathbf{b},s} \prec_\ell \dot{\prod}_{t\in T}\varepsilon_t$. In particular, for each $\mathbf{b} \in B$ the set

$$\mathbf{U}_{\nu_{\mathbf{b}}} = \{\underline{\nu}_{\mathbf{b}}(\mathbf{u}) \mid \mathbf{u} \in \mathbf{U}\} = \{\underline{\omega}(\mathbf{b})^{-1}\underline{\omega}(\mathbf{bu}) \mid \mathbf{u} \in \mathbf{U}\} \subseteq \underline{\omega}(\mathbf{b})^{-1}\mathbf{V}_\omega$$

has less than $2^{\aleph_0}$ elements. By induction, each $\mathbf{U}_{\nu_{\mathbf{b}}}$ is finite, hence also the finite union

$$\mathbf{V}_\omega = \bigcup_{\mathbf{b}\in B} \underline{\omega}(\mathbf{b})\mathbf{U}_{\nu_{\mathbf{b}}}.$$

□

6. *Nilpotent groups and specific words*

In this section we prove Theorem 1.2 and its corollaries.

LEMMA 6.1. *Let $\mathcal{C}$ be a class of profinite groups such that every commutator word is strongly concise in $\mathcal{C}$. Then every word is strongly concise in $\mathcal{C}$.*

Proof. Let $w = w(x_1, \dots, x_r) \in F$, where $F = \langle x_1, \dots, x_r\rangle$ is a free group of rank $r \geq 2$. Write $w = uv$, where $u = x_1^{e_1} \cdots x_r^{e_r}$ with $e_1, \dots, e_r \in \mathbb{Z}$ and where $v = v(x_1, \dots, x_r)$ is a commutator word, i.e. $v \in F'$.

Suppose that $|G_w| < 2^{\aleph_0}$. Let $m = \gcd(e_1, \dots, e_r)$ and choose $f_1, \dots, f_r \in \mathbb{Z}$ such that $m = \sum_{i=1}^r e_i f_i$. We observe that $g^m = w(g^{f_1}, \dots, g^{f_r})$, for every $g \in G$. Thus $\{g^m \mid g \in G\}$ has less than $2^{\aleph_0}$ elements, and consequently G_u has less than $2^{\aleph_0}$ elements, as $m = \gcd(e_1, \dots, e_r)$. Therefore, also G_v has less than $2^{\aleph_0}$ elements. Since v is strongly concise in G, the group $v(G)$ is finite. Working modulo $v(G)$, we may assume that $v(G) = 1$ and $w = u$. To simplify the notation, we may further assume that $w = x_1^m$.

As G_w has less than $2^{\aleph_0}$ elements, so does $G_{\widetilde{v}}$ for the commutator word $\widetilde{v} = \widetilde{v}(x_1, x_2) = (x_1x_2)^{-m}x_1^m x_2^m$. Since $\widetilde{v}$ is strongly concise in G, the group $\widetilde{v}(G)$ is finite. Working modulo $\widetilde{v}(G)$, we may assume that $\widetilde{v}(G) = 1$ and thus

$$(gh)^m = g^m h^m \qquad \text{for all } g, h \in G.$$

Hence $G_w = w(G)$ is the image of the homomorphism $G \to G$, $g \mapsto g^m$. From $|w(G)| < 2^{\aleph_0}$ we conclude that $w(G)$ is finite. □

Proof of Theorem 1.2. Let w be a group word, and let G be a nilpotent profinite group of class c. Suppose that $|G_w| < 2^{\aleph_0}$. We claim that $w(G)$ is finite. By Lemma 6.1, we may suppose that w is a commutator word.

Clearly, $G = \prod_p H_p$, where the product runs over all primes p and H_p denotes the unique Sylow pro-p subgroup of G. We conclude that $G_w = \prod_p (H_p)_w$ and $w(G) = \prod_p w(H_p)$. As $|G_w| < 2^{\aleph_0}$, this implies $w(H_p) = 1$ for all but finitely many primes p.

Consequently, we may suppose that G is a pro-p group. As G is nilpotent, it satisfies the hypothesis of Corollary 4.2 (see [**15**, Theorem 4.1.5]). Hence we may further suppose that $w(G)$ is central and of exponent p. Using Lemma 5.7, we apply Proposition 5.9 to deduce that G_w is finite. Thus $w(G)$ is a finitely generated elementary abelian group and therefore finite. □

Proof of Corollary 1.3. Let G be a profinite group such that $|G_w| < 2^{\aleph_0}$. Suppose that $F/w(F)$ has nilpotency class c. Then $\gamma_{c+1} = [x_1, \ldots, x_{c+1}]$ can be written as the product of finitely many w-values or their inverses. Hence γ_{c+1} has less than $2^{\aleph_0}$ values in G, and Theorem 1.1 implies that $\gamma_{c+1}(G)$ is finite. Passing to the quotient $G/\gamma_{c+1}(G)$, we can assume that G nilpotent and Theorem 1.2 applies. □

LEMMA 6.2. *Let $m \in \mathbb{N}$ with prime factorisation $m = p_1^{e_1} \cdots p_k^{e_k}$. Suppose that, for each $i \in \{1, \ldots, k\}$, the word $x^{p_i^{e_i}}$ is strongly concise in the class of pro-p_i groups. Then the word x^m is strongly concise in the class of all profinite groups.*

Proof. Put $w = x^m$ and let G be a profinite group such that $|G_w| < 2^{\aleph_0}$. By Lemma 2.4, the group G is periodic. A theorem of Herfort [**8**] yields that the group G has non-trivial Sylow pro-p subgroups for only finitely many primes p.

Suppose that q is a prime not dividing m, and let Q be a Sylow pro-q subgroup of G. Then every element of Q is an mth power. Consequently, Q is finite and each of its elements has only finitely many conjugates in G, by Lemma 2.2. Hence Q is contained in a finite normal subset of G consisting of elements of finite order. By Dicman's Lemma [**14**, 14.5.7], the group Q is contained in a finite normal subgroup of G.

Consequently, there exists a finite normal subgroup $N \trianglelefteq_{\mathrm{c}} G$ that contains all Sylow pro-q subgroups of G, for primes q not dividing m. Passing to G/N, we may suppose that G is a pro-$\{p_1, \ldots, p_k\}$ group. Fix $i \in \{1, \ldots, k\}$ and let P_i be a Sylow pro-p_i subgroup of G. Then the set of $p_i^{e_i}$th powers in P_i is the same as the set of mth powers. Thus the set has less than $2^{\aleph_0}$ elements and our assumptions yield that the group $K_i = P_i^{p_i^{e_i}} = \langle g^{p_i^{e_i}} \mid g \in P_i \rangle$ is finite. Each element of K_i has only finitely many conjugates in G; compare Lemma 2.2. Thus K_i is contained in a finite normal subgroup $N_i \trianglelefteq G$, again by Dicman's Lemma.

Factoring out the finite normal subgroup $N_1 N_2 \cdots N_k$, we may suppose that each of the Sylow pro-p_i subgroups P_i has exponent dividing $p_i^{e_i}$. Thus G has exponent dividing $m = p_1^{e_1} \cdots p_k^{e_k}$ and $w(G) = 1$ □

Proof of Corollary 1.4. By Lemma 6.2, the assertion for x^6 follows once we have dealt with the words x^2 and x^3. Let F be a free group of countably infinite rank, and let w be one of the specific words, other than x^6, that appear in the statement of the corollary. It suffices to show that $F/w(F)$ is nilpotent, so Corollary 1.3 can be applied.

(i) $w = x^2$. It is well known that $F/w(F)$ s abelian.

(ii) $w = [x^2, z_1, \ldots, z_r]$. Extending the argument given in (i), we see that $F/w(F)$ is nilpotent of class at most $1 + r$.

(iii) $w = x^3$. Since $F/w(F)$ has exponent 3, it is a 2-Engel group and thus nilpotent of class at most 3, by a classical result of Hopkins [**9**]; compare [**14**, 12.3.6].

(iv) $w = [x^3, z_1, \ldots, z_r]$ for $r \geq 1$. Extending the argument given in (iii), we see that $F/w(F)$ is nilpotent of class at most $3 + r$.

(v) $w = [x, y, y]$. Every 2-Engel group is nilpotent of class at most 3.

(vi) $w = [x, y, y, z_1, \ldots, z_r]$ for $r \geq 1$. Extending the argument given in (v), we see that $F/w(F)$ is nilpotent of class at most $3 + r$.

□

Acknowledgements. In connection with our original proof of Theorem 1.1 we acknowledge discussions with Marta Morigi; subsequent comments of the referee led to a significant simplification of our argument. We also acknowledge the referee's general feedback which led to several improvements of the presentation.

References

1. R. G. Burns, Y. Medvedev, *Group laws implying virtual nilpotence*, J. Aust. Math. Soc. **74** (2003), 295–312.
2. E. Detomi, M. Morigi, P. Shumyatsky, *On conciseness of words in profinite groups*, J. Pure Appl. Algebra **220** (2016), 3010–3015.
3. E. Detomi, M. Morigi and P. Shumyatsky, *On profinite groups with word values covered by nilpotent subgroups*, Israel J. Math. **226** (2018), 993–1008.
4. E. Detomi, M. Morigi and P. Shumyatsky, *On profinite groups with commutators covered by countably many cosets*, J. Algebra **508** (2018), 431–444.
5. G. A. Fernández-Alcober, P. Shumyatsky, *On bounded conciseness of words in residually finite groups*, J. Algebra **500** (2018), 19–29.
6. K. W. Gruenberg, *Two theorems on Engel groups*, Proc. Cambridge Philos. Soc. **49** (1953), 377–380.
7. R. Guralnick, P. Shumyatsky, *On rational and concise word*, J. Algebra **429** (2015), 213–217.
8. W. N. Herfort, *Compact torsion groups and finite exponent*, Arch. Math. (Basel) **33** (1979/80), 404–410.
9. C. Hopkins, *Finite groups in which conjugate operations are commutative*, Amer. J. Math. **51** (1929), 35–41.
10. S. V. Ivanov, *P. Hall's conjecture on the finiteness of verbal subgroups*, Soviet Math. (Iz. VUZ) **33** (1989), 59–70.
11. A. Jaikin-Zapirain, *On the verbal width of finitely generated pro-p groups*, Rev. Mat. Iberoam. **24** (2008), 617–630.
12. Yu I. Merzlyakov, *Verbal and marginal subgroups of linear groups*, Dokl. Akad. Nauk SSSR **177** (1967), 1008–1011.
13. D. J. S. Robinson, Finiteness conditions and generalized soluble groups, Part 1, Springer-Verlag, New York-Berlin, 1972.
14. D. J. S. Robinson, A course in the theory of groups, Springer-Verlag, New York, 1996.
15. D. Segal, Words: notes on verbal width in groups, Cambridge University Press, Cambridge, 2009.
16. P. Shumyatsky, *Verbal subgroups in residually finite groups*, Q.J. Math. **51** (2000) 523–528.
17. P. Shumyatsky, *On varieties arising from the solution of the restricted Burnside problem*, J. Pure Appl. Algebra **171** (2002), 67–74.
18. R. F. Turner-Smith, *Marginal subgroup properties for outer commutator words*, Proc. London Math. Soc. **14** (1964), 321–341.
19. J. C. Wilson, *On outer-commutator words*, Canadian J. Math. **26** (1974), 608–620.

Eloisa Detomi
Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Padova, Via Gradenigo 6/b, 35131 Padova
Italy

eloisa.detomi@unipd.it

Benjamin Klopsch
Heinrich-Heine-Universität Düsseldorf, Mathematisches Institut, Universitätsstr. 1, 40225, Düsseldorf
Germany

klopsch@math.uni-duesseldorf.de

Pavel Shumyatsky
Department of Mathematics, University of Brasilia, Brasilia-DF 10910-900
Brazil

pavel@unb.br