Astérisque

JOHN S. WILSON

Finite index subgroups and verbal subgroups in profinite groups

Astérisque, tome 339 (2011), Séminaire Bourbaki,
exp. n° 1026, p. 387-408

<http://www.numdam.org/item?id=AST_2011__339__387_0>

Séminaire BOURBAKI Juin 2010
62e année, 2009-2010, n° 1026, p. 387 à 408

FINITE INDEX SUBGROUPS AND VERBAL SUBGROUPS IN PROFINITE GROUPS

by **John S. WILSON**

In the 1970s J-P. Serre proved that if H is a subgroup of finite index in a finitely generated pro-p-group G then H is necessarily an open subgroup of G, and he commented that he did not know whether the same conclusion holds for arbitrary finitely generated profinite groups G. This was finally shown to be the case by Nikolov and Segal in 2003. The result depends on important properties of values of group words in finite groups, which are likely to have other applications. We shall discuss the background to these results, aspects of the proof, related results, and some ways in which the results have already been used.

1. INTRODUCTION TO PROFINITE GROUPS

A profinite group is by definition an inverse limit of finite groups. The profinite groups can be characterized among topological groups as the compact Hausdorff totally disconnected groups.

Among the profinite groups, pro-p-groups—inverse limits of finite p-groups, where p is a fixed prime—play a distinguished role, like the role of finite p-groups among finite groups. They are precisely the compact Hausdorff totally disconnected groups such that $x^{p^n} \to 1$ as $n \to \infty$ for all elements x.

Profinite groups arise naturally in many parts of mathematics: in analysis and the general theory of topological groups as the quotients of compact groups modulo the connected component of the identity; in algebraic number theory as Galois groups, etc.; and in combinatorics and model theory as automorphism groups of structures (in connection for example with the small index property). They arise in finite group theory as objects encoding information about infinite families of groups and allowing concise statements of asymptotic results, and in infinite group theory as examples, and as completions, etc.

Since profinite groups are compact and Hausdorff, their subgroups of principal interest, those that are again profinite, are just the closed subgroups. Therefore, for example, we shall say that a subset S *topologically* generates a profinite group G if the closure of the abstract subgroup generated by S is G. A profinite group is called (*topologically*) *finitely generated* if it is topologically generated by some finite set.

Consideration of coset decompositions shows that a subgroup of finite index in a profinite group is open if and only if it is closed, and, since profinite groups are compact, that open subgroups have finite index. It is natural to ask under what circumstances the converse holds, i.e. when all subgroups of finite index are open. Since profinite groups can also be described as the topological groups (topologically) isomorphic to closed subgroups of Cartesian products of finite groups with the product topology, useful insight can be gained by considering the special case of Cartesian products.

2. CARTESIAN PRODUCTS AND THEIR IMAGES

We write $C = \prod_{i\in I} G_i$ for the Cartesian product of a family of sets $(G_i)_{i\in I}$. Its elements are the vectors $(g_i)_{i\in I}$, with entries $g_i \in G_i$ for each $i \in I$. For each i write $\pi_i : C \to G_i$ for the projection map.

If each G_i is a group, then so is C with componentwise multiplication. If each G_i is a topological space, then so is C, with subbase of open sets

$$\{\pi_i^{-1}(U) \mid i \in I,\ U \text{ open in } G_i\}.$$

If each G_i is a topological group, then so is C.

Tychonoff's theorem asserts that if each G_i is a compact topological space, then so is C. Hence if each G_i is a compact Hausdorff topological group, then so is each closed subgroup of C.

In particular, we may take each G_i to be a finite group with the discrete topology. All profinite groups are (topologically) isomorphic to closed subgroups of such products; and the countably based (i.e. separable) profinite groups are precisely the (groups isomorphic to) closed subgroups of $\prod_{n\in\mathbb{N}} \mathrm{Sym}(n)$.

In general, not all subgroups of finite index in a profinite group are open. Consider, for example, the product $C = \prod_{i\in\mathbb{N}} G_i$ with each G_i a non-trivial finite group. The definition of the product topology shows that C has $\aleph_0$ open normal subgroups. However it can be proved that if for some non-trivial finite group F we have $G_i \cong F$ for all i, then C has $2^{2^{\aleph_0}}$ subgroups of index $|F|$. In the case when $F \cong \mathbb{Z}/p\mathbb{Z}$ this is clear, for then $\dim_{\mathbb{F}_p} C = 2^{\aleph_0}$, and the dual of a vector space of infinite dimension d has

dimension 2^d. To see where non-open subgroups come from when F is non-abelian, we consider ultraproducts.

Let I be an infinite index set. We recall that an *ultrafilter* on I is a family $\mathscr{U}$ of subsets of I such that

(i) $\varnothing \notin \mathscr{U}$;
(ii) if $X_1, X_2 \in \mathscr{U}$ then $X_1 \cap X_2 \in \mathscr{U}$;
(iii) for each $X \subseteq I$, either $X \in \mathscr{U}$ or $I \setminus X \in \mathscr{U}$.

From this it follows that if $X_1 \in \mathscr{U}$ and $X_1 \subseteq X_2$ then $X_2 \in \mathscr{U}$. For $i \in I$ the family $\mathscr{U}_i = \{X \subseteq I \mid i \in X\}$ is an ultrafilter. An ultrafilter that is not of the form $\mathscr{U}_i$ must evidently contain all subsets of I with finite complements, and the existence of such *non-principal ultrafilters* follows from Zorn's lemma. We fix a non-principal ultrafilter $\mathscr{U}$.

Let $(G_i)_{i\in I}$ be a family of groups. Write $C = \prod G_i$ and let $K_{\mathscr{U}}$ be the normal subgroup $\{(g_i) \in C \mid \{i \mid g_i = 1\} \in \mathscr{U}\}$ of C. The *ultraproduct* $\prod G_i/\mathscr{U}$ is defined to be the quotient group $C/K_{\mathscr{U}}$.

Now take all groups G_i to be equal to a non-trivial finite group F. If $(g_i) \in C$ then $\{i \in I \mid g_i = f\}$ belongs to $\mathscr{U}$ for just one $f \in F$, and it follows easily that $C/K_{\mathscr{U}} = \prod G_i/\mathscr{U}$ is isomorphic to F. Clearly $K_{\mathscr{U}}$ is not open.

This method for obtaining non-open subgroups of finite index in Cartesian powers is of course non-constructive. This is necessarily the case, since it is consistent with (ZF and) the principle of dependent choice that all subgroups of finite index in all countably based profinite groups are open, as observed by Lascar [**8**].

The above argument gives the implication (i) $\Rightarrow$ (ii) in the following result:

THEOREM 2.1 (Saxl and Wilson [**20**]; Martínez and Zel'manov [**11**])
Let $C = \prod_{i\in\mathbb{N}} S_i$ with each S_i a non-abelian finite simple group. Then the following are equivalent:

(i) *all subgroups of finite index are open in C;*
(ii) *there are only finitely many groups S_i of each isomorphism type.*

It is reasonable to ask what restrictions there are on abstract images of profinite groups. They can be countably infinite: for example this holds trivially for $C = \prod G_i$ with all groups G_i cyclic of order p, and it holds for $\mathbb{Z}_p$, as shown by considering the composite of the inclusion from $\mathbb{Z}_p$ to its field of fractions $\mathbb{Q}_p$ and a surjective $\mathbb{Q}$-linear map from $\mathbb{Q}_p$ to $\mathbb{Q}$.

If $(S_i)_{i\in\mathbb{N}}$ is a family of Chevalley groups ${}^{\varepsilon}X_r(F_i)$ of the same (twisted or untwisted) type ${}^{\varepsilon}X_r$, then for any ultrafilter $\mathscr{U}$ on $\mathbb{N}$ we have $\prod_i S_i/\mathscr{U} \cong {}^{\varepsilon}X_r(\prod F_i/\mathscr{U})$ by a result of Point [**18**]; in particular this group is simple. These Chevalley groups are the

only completely explicit examples of infinite simple images of profinite groups known to the author.

Now let $\mathscr{U}$ be an ultrafilter on $\mathbb{N}$. The *ultralimit* $\lim_{\mathscr{U}} r_i$ of a bounded sequence $(r_i)_{i\in\mathbb{N}}$ of real numbers is the unique $\alpha \in \mathbb{R}$ such that for any $\varepsilon > 0$ the set $\{i \in \mathbb{N} \mid |r_i - \alpha| < \varepsilon\}$ is in $\mathscr{U}$. Again let $C = \prod_{i\in\mathbb{N}} S_i$ with each S_i a non-abelian finite simple group. Define $h = h_{\mathscr{U}} : C \to [0,1]$ by

$$h(a) = \lim_{\mathscr{U}} \frac{\log |a_i^{S_i}|}{\log |S_i|} \quad \text{for } a = (a_i) \in C;$$

here, $a_i^{S_i}$ denotes the conjugacy class of a_i in S_i. Clearly we have $h(b^{-1}ab) = h(a)$ and $h(ab) \leqslant h(a) + h(b)$ for all $a, b \in C$, and so $L_{\mathscr{U}} = \{g \in C \mid h(g) = 0\}$ is a normal subgroup of C.

Nikolov [**12**] has proved the following result.

THEOREM 2.2. — *With the above notation, the simple quotients of C are exactly the groups $C/L_{\mathscr{U}}$ for some ultrafilter $\mathscr{U}$ on $\mathbb{N}$; the infinite simple quotients all have cardinality $2^{\aleph_0}$.*

In [**4**], Holt constructed a sequence $(K_i)_{i\in\mathbb{N}}$ of finite groups with no non-trivial abelian images and with $(\log |K_i|)/(\log |\{x^2 \mid x \in K_i\}|)$ unbounded. These two properties imply that $\prod K_i$ is a profinite group with no non-trivial abelian continuous homomorphic image but with an abstract subgroup of index 2.

The implication (ii) $\Rightarrow$ (i) in Theorem 2.1 is a consequence of assertion (a) of the following result, and the proof of Theorem 2.2 depends on assertion (c). These results in turn depend on the classification of the finite simple groups (CFSG).

THEOREM 2.3. — (a) (Saxl and Wilson [**20**]; Martínez and Zel'manov [**11**]) *For each integer $q > 0$, there is a number $n(q)$ such that in every non-abelian finite simple group S either each element of S is a product of $n(q)$ q-th powers or all q-th powers equal the identity element.*

(b) (Wilson [**24**]) *There is a constant c such that every element of every non-abelian finite simple group S is a product of c commutators.*

(c) (Liebeck and Shalev [**10**]) *There is a constant d such that if S is a non-abelian finite simple group and $1 \neq h \in S$ then every element of S is a product of at most n conjugates of h, where $n = \lfloor d \log |S| / \log |h^S| \rfloor$.*

It is now known that the constant c above can be taken to be 1: thus every element of a non-abelian finite simple group is a commutator. This was established in [**9**], completing the work of many mathematicians over many years.

3. SUBGROUPS OF FINITE INDEX

We return to arbitrary profinite groups and ask which of them have the property that all subgroups of finite index are open. Because of the following elementary lemma, such groups are sometimes said to be *strongly complete.*

LEMMA 3.1. — *For a profinite group G the following are equivalent:*

(i) *all subgroups of finite index in G are open;*

(ii) *the map from G to its profinite completion is bijective;*

(iii) *all abstract homomorphisms from G to profinite groups are continuous.*

Here is a characterization of the strongly complete profinite groups.

THEOREM 3.2 (Smith and Wilson [**22**]). — *Let G be profinite. Then G is strongly complete if and only if G has only countably many subgroups of finite index, or equivalently if and only if G has only finitely many subgroups of index n for each integer n.*

A group satisfying the conditions of Theorem 3.2 is certainly countably based (with base consisting of all cosets of all subgroups of finite index). Among the countably based groups are the (topologically) finitely generated groups. In about 1975, J-P. Serre raised the tantalizing prospect that all finitely generated profinite groups might be strongly complete. This was confirmed by Nikolov and Segal in 2003:

THEOREM 3.3 (Nikolov and Segal [**14**], [**15**]). — *If G is a finitely generated profinite group then all subgroups of finite index are open in G.*

We shall discuss some aspects of the proof of this important result later.

Serre had shown in the 1970s that finitely generated pro-p-groups are strongly complete. One proof of this earlier result proceeds by showing that if G is a finitely generated pro-p-group then the subgroup $[G,G]G^p$ generated (algebraically) by all p-th powers and all commutators is open in G. This subgroup is the verbal subgroup corresponding to the word $[x,y]z^p$ in the free group on x,y,z. To deal with arbitrary profinite groups we need to consider verbal subgroups corresponding to more general words.

4. SOME NOTATION

For elements x, y and subsets S, T of a group G and an integer $m > 1$ we write

$x^y = y^{-1}xy$, $[x,y] = x^{-1}y^{-1}xy$,
$[S,y] = \{[s,y] \mid s \in S\}$, $\mathfrak{c}(S,T) = \{[s,t] \mid s \in S, t \in T\}$,
$ST = \{st \mid s \in S, t \in T\}$,
$S^{\{m\}} = \{s^m \mid s \in S\}$,
$S^{(m)}$ for the Cartesian product of m copies of S,
S^{*m} for the set of all products $s_1^{\pm 1} \dots s_m^{\pm 1}$ in G.

For subgroups H, K, we write

$[H,K]$ for the *subgroup* generated by $\{[h,k] \mid h \in H, k \in K\}$,
$H' = [H,H]$, the *derived group* of H,
$H'' = [H',H']$, the *second derived group* of H,

and we define $[H,_n K]$ recursively by
$[H,_1 K] = [H,K]$, $[H,_{n+1} K] = [[H,_n K], K]$ for $n \geqslant 1$.

5. BAIRE'S CATEGORY THEOREM AND VERBAL SUBGROUPS

Baire's category theorem from general topology plays an important part in the reduction of problems about profinite groups to problems about finite groups. In a profinite group G the cosets Nt with $N \triangleleft G$ open, $t \in G$, form a basis for the topology, and Baire's category theorem takes the following form: *if G is a profinite group and $(C_n)_{n\in\mathbb{N}}$ is a countable family of closed sets whose union contains a non-empty open set, then some C_n contains an open coset Nt.* (A proof is sketched in [**25**, (2.7.10)].) The next result is an easy application:

PROPOSITION 5.1. — *Let S be a non-empty closed subset of a profinite group. Then the subgroup A generated algebraically by S and the subgroup generated topologically by S coincide if and only if for some integer n every element of A is a product of at most n elements of $S \cup S^{-1}$.*

Proof. — For each integer m let $C_m = (S \cup S^{-1})^{*m}$ (in the notation of Section 4). Thus C_m is the image of $T^{(m)}$ under the continuous map $(t_1, \dots, t_m) \mapsto t_1 \dots t_m$ where $T = S \cup S^{-1}$, and so, since T is compact, the set C_m is compact and hence closed. Therefore if $A = C_n$ for some n then A equals the subgroup generated topologically by S. On the other hand, in general we have $A = \bigcup_m C_m$, and so if A is closed then some C_m contains a coset Nu with $u \in A$ and N open in A; thus $N \subset (Nu)(Nu)^{-1} \subset C_{2m}$ and it follows that $A = C_n$ where $n = 2m + |A : N|$. □

In particular, the abstract derived group $G' = \langle [g,h] \mid g,h \in G \rangle$ is closed if and only if it consists of products of n commutators for some integer n. This is because the set $S = \{[g,h] \mid g,h \in G\}$ is closed, being the image of the compact space $G \times G$ under the continuous map $(g,h) \mapsto [g,h]$. A similar conclusion holds if the word $[x,y]$ is replaced by any other word w from a free (abstract) group.

DEFINITIONS. — *Let w be a group word (i.e. an element of an abstract free group on a set X) and G a group.*

1) *A* w-value *in G is an element obtained by substituting elements of G in place of the elements of X in w or w^{-1}.*

2) *The* verbal subgroup $w(G)$ *is $\langle w(g_1,\dots,g_k) \mid g_1,\dots,g_k \in G \rangle$, the subgroup generated* algebraically *(whether or not G is a topological group) by all w-values in G.*

3) *A subset S of G has* width at most m *in G if the (abstract) subgroup generated by S is equal to $(S \cup S^{-1})^{*m}$.*

4) *w has width at most m in G if every element of $w(G)$ is a product of at most m w-values.*

The convention of including inverses in 2) above follows **[14]**, **[15]**. We note that the commutators $[h_1,h_2] = h_1^{-1}h_2^{-1}h_1h_2$ are the values of the word $[x,y]$, since $[x,y]^{-1} = [y,x]$, and the values of the word x^q for an integer $q > 0$ are just the q-th powers in G.

The following result is essentially due to Hartley **[3]**.

PROPOSITION 5.2. — *Let G be a profinite group and w an abstract group word.*

(a) *w has width at most k in G if and only if w has width at most k in G/N for all open normal subgroups N.*

(b) *The following are equivalent:*

(i) *$w(G)$ is closed in G;*

(ii) *w has finite width in G.*

Proof. — We use the fact that if Y is a closed set in G then $L = \bigcap(YN)$, where the intersection is over all open normal subgroups N (see **[25]**, (0.3.3)]). If w has width at most k in each G/N then $w(G) \subseteq NT^{*k}$ for each N, where T is the set of w-values, and since T^{*k} is closed we have $w(G) = T^{*k}$. The other implication in (a) is clear, and (b) is provided by Proposition 5.1. □

The following proposition is another simple application of Baire's category theorem.

PROPOSITION 5.3. — *A verbal subgroup of a profinite group cannot have countably infinite index.*

Proof. — It suffices to note that if H is a subgroup of countable index in a profinite group G and H is a union of countably many closed sets C_m, then G is a union of countably many sets $C_m g$; hence some C_m contains an open coset Nu and H is open. □

6. WIDTH OF WORDS; SERRE'S THEOREM

The width of words in abstract groups was studied by Philip Hall and his student Peter Stroud in the 1960s. Stroud proved that if G is finitely generated and is an extension of an abelian group by a nilpotent group, then all words have finite width in G. Hall also gave the following example in lectures in Cambridge in 1967.

PROPOSITION 6.1. — *There is a 2-generator (abstract) group G with the following properties*:

(i) *G has a normal subgroup G_1 that is nilpotent of class 2 and with $G/G_1 \cong \mathbb{Z}$;*

(ii) *the word $[[x,y],[z,t]]$ does not have finite width in G.*

Proof. — Let F be a free (abstract) group on s,t and let K be the kernel of the homomorphism from F to $\langle t \rangle$ mapping s to 1 and t to itself. Let L be the third term of the lower central series of K; thus $L = [[K,K],K]$ and $L \lhd F$. Evidently any value of the word $[[x,y],[z,t]]$ in F/L is a value of the word $[x,y]$ in K/L, so if the width in F/L of the former word is finite then so is the width of the latter in K/L. But K is a free group with basis $\{s^{t^n} \mid n \in \mathbb{Z}\}$, and so the quotient K'/L is isomorphic to the exterior square $E = (K/K') \wedge (K/K')$ under the map $E \to K'/L$ induced by the commutator map. However clearly there is no bound on the lengths of elements of E as sums of elements of form $\pm u \wedge v$. □

To focus attention more sharply on widths of words in groups, Hall asked about the special case of words w taking only finitely many values in a group G: then w has finite width if and only if the subgroup $w(G)$ is finite. Ivanov [**5**] gave an example, consisting of an abstract group G and a word w such that there is only one non-trivial element of the form $w(\mathbf{g})$ with $\mathbf{g}$ a sequence of elements of G; this element has infinite order and so $w(G)$ is infinite. It is not known whether there is a similar example for profinite groups.

An *outer commutator word* in a free group on a set X is a word obtained by finitely many operations of taking commutators of two words involving disjoint sets of elements of X, starting from elements of $X \cup X^{-1}$. Strengthening earlier results,

Fernandéz-Alcober and Morigi **[2]** have recently shown that if w is an outer commutator word on $x_1, \dots, x_d$ and G is a group such that $|\{w(\mathbf{g}) \mid \mathbf{g} \in G^{(d)}\}| = n$ then $|w(G)| \leqslant (n-1)^{n-1}$.

The following elementary result played a role in the early investigations of verbal width, and generalizations of it remain important today. For example, Theorem 8.1 and the Key Theorem discussed below make assertions of a similar character but are valid in a vastly wider setting. The notation is as defined in Section 4.

LEMMA 6.2. — *Let G be a group and suppose that $G = G'\langle g_1, \dots, g_d\rangle$. Let H be a normal subgroup of G. Then*

(a) $[H,G] = [H,g_1]\dots[H,g_d][H,{}_nG]$ *for all integers* $n \geqslant 1$.

(b) *If in addition G is a pro-p-group and H is closed, then $[H,G] = [H,g_1]\dots[H,g_d]$; in particular, $[H,G]$ is closed.*

Proof. — (a) The crucial point is that the commutator map $(x,y) \mapsto [x,y]$ behaves like a bilinear map when elements of its image lie in the centre.

The result is clear for $n = 1$. When $n = 2$ we can pass to quotients modulo $[H,{}_2G]$; thus we can assume that $[H,G] \leqslant Z(G)$. For each $h \in H$ the map $x \mapsto [h,x]$ from G to $Z(G)$ is a homomorphism; so we have $[H,G'] = 1$ and $[H,G] = [H,\langle g_1,\dots,g_d\rangle]$. For each i the map $h \mapsto [h,g_i]$ is a homomorphism from H to $Z(G)$, and so its image $[H,g_i]$ is a subgroup. Let $L = [H,g_1]\dots[H,g_d]$. The image of H in G/L lies in the centre, and hence $[H,G] = L$.

Now let $n \geqslant 3$ and assume the obvious induction hypothesis. We may assume that $[H,{}_nG] = 1$; thus $[K,G] \leqslant Z(G)$ where $K = [H,{}_{n-2}G]$. By induction,

$$[H,G] = [H,g_1]\dots[H,g_d]\,[K,G] \quad \text{and} \quad [K,G] = [K,g_1]\dots[K,g_d].$$

Since $[kh,g_i] = [k,g_i][h,g_i]$ for all $k \in K$, $h \in H$ it follows that

$$[H,G] = ([K,g_1][H,g_1])\dots([K,g_d][H,g_d]) = [H,g_1]\dots[H,g_d].$$

(b) Since the set $Y = [H,g_1]\dots[H,g_d]$ is closed and finite p-groups are nilpotent, (a) and the fact used to prove Lemma 5.2 give $[H,G] \subseteq \bigcap(YN) = Y$, as required. □

Serre's result on finitely generated pro-p-groups follows easily from this:

THEOREM 6.3 (Serre). — *Every subgroup of finite index in a finitely generated pro-p-group is open.*

Proof. — First suppose that G is a finitely generated abelian profinite group and H a subgroup of finite index n. The map $g \mapsto g^n$ is a continuous homomorphism from G and so its kernel K is closed; since G/K is finite and $K \leqslant H$ it follows that H is open.

To complete the proof, it suffices to show that normal subgroups of finite index in finitely generated pro-p-groups are open (since the intersection of all conjugates of a subgroup of finite index again has finite index). Suppose that $n > 1$ and that all normal subgroups of index less than n in finitely generated pro-p-groups are open. Let G be a finitely generated pro-p-group and H a normal subgroup of index n. If K is a normal subgroup with $H < K < G$ then by induction K is open; hence K is a pro-p-group and is finitely generated, so that H must be open in K and hence open in G. Thus we may assume that G/H is simple. Let $g \in G \setminus H$ and let C be the closed subgroup generated by g. Then $C \cap H$ is open in C from the first paragraph, and so $|CH/H| = |C/(C \cap H)|$ is a power of p. Therefore G/H is a simple p-group and so is cyclic; thus $[G, G] \leqslant H$. By Lemma 6.2 the normal subgroup $[G, G]$ is closed; therefore $G/[G, G]$ is a finitely generated abelian pro-p-group, and its subgroup $H/[G, G]$ is open, from the first paragraph. Because the quotient map is continuous it follows that H is open in G. □

Lemma 6.2 shows that the word $[x, y]$ has finite width in all finitely generated pro-p-groups. Roman'kov **[19]** proved that each word w has finite width in every finitely generated virtually nilpotent profinite group. Jaikin-Zapirain **[6]** made a major advance by proving the following result.

THEOREM 6.4. — *Let G be a compact p-adic analytic group. Then all words have finite width in G.*

The lovely proof of this result uses in an essential way the structure of the groups as manifolds and is strongly Lie-theoretic. Segal **[21]** has generalized this result further to certain Cartesian products of p-adic analytic groups.

7. SUBGROUPS OF FINITE INDEX AND d-FINITE WORDS

For an integer $d \geqslant 1$ we say that a word w is *d-locally finite* if every d-generator abstract group H with $w(H) = 1$ is finite or, equivalently, if $F_d/w(F_d)$ is finite, where F_d is the free (abstract) group of rank d.

Nikolov and Segal prove the following:

THEOREM 7.1. — *Let w be a d-locally finite word and let G be a d-generator profinite group. Then there is a number $f = f(d, w)$ such that w has width at most f in G. Consequently (by Lemma 5.2) the verbal subgroup $w(G)$ is closed in G.*

Now suppose that H is a subgroup of finite index in a d-generator profinite group G. Then the intersection K of the conjugates of H in G is a normal subgroup of finite index. Suppose that there is a d-locally finite word w such that $w(G/K) = 1$. Then

$w(G) \leqslant K \leqslant H$. By Theorem 7.1, $w(G)$ is closed in G and so equals the intersection of all open normal subgroups N containing $w(G)$. For each such N, G/N is a d-generator group with $w(G/N) = 1$, and so G/N is an image of the finite group $F_d/w(F_d)$. Therefore all such subgroups N have index at most $|F_d/w(F_d)|$, and their intersection $w(G)$ must have index at most $|F_d/w(F_d)|$. We conclude that both $w(G)$ and H are open.

Thus Theorem 3.3 is an easy consequence of Theorem 7.1 and the following fact:

PROPOSITION 7.2. — *Let $d > 1$ and let G be a finite group. Then there is a d-locally finite word w such that $w(G) = 1$.*

This proposition is a consequence of the Oates–Powell theorem, a deep theorem on varieties of groups, proved in 1964 in [**17**]. Nikolov and Segal give the following elementary proof of Proposition 7.2.

Proof. — Let F be a free group with basis $x_1, \ldots, x_d$, let Θ be the set of all homomorphisms from F to G, and let $D = \bigcap_{\theta \in \Theta} \ker \theta$. Since Θ is finite, D has finite index in F, and so D is finitely generated; let $D = \langle w_1, \ldots, w_m \rangle$. If $g_1, \ldots, g_d \in G$ then consideration of the homomorphism given by $x_k \mapsto g_k$ shows that $w_i(g_1, \ldots g_d) = 1$; thus $w_i(G) = 1$. Similarly, since the composite of a homomorphism from F to itself and a member of Θ necessarily belongs to Θ, we have $w_i(\mathbf{u}) \in D$ for each i and each d-tuple $\mathbf{u}$ of elements of F. Now define a word w in the free group of rank md as follows:

$$w(\mathbf{y_1}, \ldots, \mathbf{y_m}) = w_1(\mathbf{y_1}) \ldots w_m(\mathbf{y_m});$$

here $\mathbf{y_1}, \ldots, \mathbf{y_m}$ are disjoint d-tuples of variables. From above we have $w(G) = 1$ and $w(F) = D$, and the result is proved. □

8. OTHER CLOSED VERBAL SUBGROUPS

In [**14**], [**15**], the methods used to prove Theorem 7.1 were also used to prove the following result.

THEOREM 8.1. — *Let G be a finitely generated profinite group and H a closed normal subgroup. Then the (abstract) subgroup $[H, G]$ generated by $\{[h, g] \mid h \in H, g \in G\}$ is closed. Consequently all terms $\gamma_n(G)$ of the lower central series of G are closed.*

The subgroups $\gamma_n(G)$ are defined by $\gamma_1(G) = G$ and $\gamma_{n+1}(G) = [\gamma_n(G), G]$ for $n \geqslant 1$; they are well known to be the verbal subgroups corresponding to the words γ_n defined by $\gamma_1 = x_1$ and $\gamma_{n+1} = [\gamma_n, x_{n+1}]$ for $n \geqslant 1$.

In the later paper [**16**], Nikolov and Segal proved the following result.

THEOREM 8.2. — *If d, $q \in \mathbb{N}$ and H is a finite d-generator group, then the word x^q has (d,q)-bounded width in H. Consequently, if G is a finitely generated profinite group and $q \in \mathbb{N}$, then the subgroup G^q is closed.*

Zelmanov's solution of the restricted Burnside problem **[28]**, **[29]** shows that for all $d \geqslant 1$ and all $q \geqslant 1$ there is a bound on the orders of all finite d-generator groups of exponent dividing q. Using this, Theorem 8.2 and the argument given in the deduction of Theorem 3.3 from Theorem 7.1, we obtain

COROLLARY 8.3. — *If G is a finitely generated profinite group and $q \in \mathbb{N}$, then G^q is open.*

The use of Zel'manov's result here is inevitable: Jaikin-Zapirain **[6]** has shown that the statement of Corollary 8.3 with $q = p^{n+1}$ implies a positive solution to the restricted Burnside problem for exponent p^n.

A *commutator word* is a word that lies in the derived group of the free group on a set X. Clearly w is a non-commutator word if and only if for some $x \in X$ the sum q of the exponents of x in w is non-zero, and, since the values obtained by mapping $X \setminus \{x\}$ to 1 are q-th powers, the non-commutator words are precisely the 1-locally finite words.

COROLLARY 8.4. — *If G is a finitely generated profinite group and w is a non-commutator word, then the verbal subgroup $w(G)$ is open in G.*

This follows directly from Corollary 8.3, because $w(G) \geqslant G^q$ where $q = |\mathbb{Z}/w(\mathbb{Z})|$.

On the other hand, not all words w have the property that $w(G)$ is closed for every finitely generated profinite group G. Roman'kov **[19]** gave an example of a finitely generated pro-p-group G with $w(G)$ not closed and with $w(G) \leqslant Z(G)$, where $w = [[x,y],[z,t]]$.

In **[6]**, Jaikin-Zapirain showed that if $1 \neq w \in F''(F')^p$ where F is a free group and p is any prime, then $w(G)$ is not closed in the free pro-p-group of rank 2. He proved the following result.

THEOREM 8.5. — *For $1 \neq w \in F$ and a prime p, the following are equivalent:*

(i) *$w(G)$ is closed in G for every finitely generated pro-p-group G;*

(ii) *$w \notin F''(F')^p$;*

(iii) *$H/\overline{w(H)}$ is virtually nilpotent for every finitely generated pro-p-group H.*

Here, $\overline{w(H)}$ denotes the closure of $w(H)$, and we call a profinite group virtually nilpotent if it has an open nilpotent subgroup.

For example x^{p^d} satisfies (iii) by the solution to the restricted Burnside problem, and so does the Engel word $[x,_n y]$ for all n, by a result in **[23]**.

The implication (iii) $\Rightarrow$ (ii) in the above theorem is elementary, and (ii) $\Rightarrow$ (iii) is an easy consequence of results of Burns and Medvedev **[1]**. To prove (i) $\Rightarrow$ (ii), Jaikin-Zapirain generalizes Roman'kov's example mentioned above. If $w \in F''(F')^p$ has finite width in a non-abelian free pro-p-group H it quickly follows that $[x,y]z^p$ has finite width in the closure L of the derived group of H; since L is a free pro-p-group of infinite rank it follows that there is a uniform bound on the width of $[x,y]z^p$ in all free pro-p-groups of finite rank. A study of growth rates of dimensions of subspaces in associated graded Lie algebras now leads easily to a contradiction. Finally, to prove (ii) $\Rightarrow$ (i), by Theorem 6.4 or the earlier result of Roman'kov **[19]** it suffices to prove that $w(G)$ contains a closed normal subgroup K with G/K virtually nilpotent, since then $w(G)/K = w(G/K)$, which is closed in G/K. This is established with $K = \gamma_{n+1}(G^{p^t})$, where t, n are sufficiently large integers depending on w, by solving congruences successively in quotients $G/\gamma_m(G^{p^t})$, rather in the fashion of Hensel's lemma.

9. UNIFORM BOUNDS

Write $G^{\{w\}}$ for the set of values of a word w in a group G. Because of Proposition 5.2, the results Theorem 7.1 and Theorem 8.1 are consequences of the following results about finite groups.

THEOREM 9.1. — *Let $d \geqslant 1$ and let w be a d-locally finite word. Then there is a function $f = f(w,d)$ such that $w(G) = G^{\{w\}*f}$ for every finite d-generator group G.*

THEOREM 9.2. — *Let $d \geqslant 1$. Then for every finite d-generator group G and normal subgroup H we have $[H,G] = \{[h,g] \mid h \in H, g \in G\}^{*l(d)}$, where $l(d) = 12d^3 + O(d^2)$.*

These results depend on the Key Theorem stated below.

DEFINITION. — *A normal subgroup H of G is* acceptable *if $[H,G] = H$ and if G has no normal subgroups Z, N with $Z \leqslant N \leqslant H$ such that N/Z is either a non-abelian simple group or the direct square of a non-abelian simple group.*

THEOREM 9.3 (Key Theorem). — *Let $G = \langle g_1, \ldots, g_d \rangle$ be a finite group and H an acceptable normal subgroup, and let $q > 1$. Then*

$$H = ([H,g_1] \ldots [H,g_d])^{*h_1(d,q)} \cdot (H^{\{q\}})^{*z(q)},$$

and

$$H = ([H, g_1] \dots [H, g_d])^{*h_2(d)} \cdot (\mathfrak{c}(H, H))^{*D},$$

where $h_1(d, q)$, $z(q)$, $h_2(d)$ depend only on the indicated variables and D is a constant.

To see how the first assertion of the Key Theorem can be applied we require two lemmas; the first completely elementary and the second a mild extension of Theorem 2.3 (a).

LEMMA 9.4. — *Let K be a finite d-generator group and let $C > 0$. Then K has characteristic subgroups*

$$K_5 \leqslant K_4 \leqslant K_3 \leqslant K_2 \leqslant K_1$$

such that

(i) *K_5 is acceptable in K;*

(ii) *K_3/K_5 is perfect and $K_4/K_5 = Z(K_3/K_5)$;*

(iii) *K_3/K_4 is a direct product of non-abelian simple groups of order greater than C;*

(iv) *K_1/K_3 is soluble, $K_2 = [K_1, {}_n K]$ for some n and K_2/K_3 is acceptable in K/K_3;*

(v) *$|K : K_1|$ is bounded in terms of C and d.*

Proof. — Let $K_1 = \bigcap_{\theta \in \Theta} \ker \theta$, where Θ is the set of all homomorphisms θ from K to groups $\mathrm{Aut}\,(S \times S)$ with S non-abelian simple and $|S| \leqslant C$.

Let $K_2 = \bigcap_{n \in \mathbb{N}} [K_1, {}_n K]$, let K_3 be the smallest normal subgroup with K_1/K_3 soluble, let $K_4 = \bigcap (K \lhd K_3 \mid H_3/K \in \mathscr{T})$, where $\mathscr{T}$ is the family of non-abelian simple groups of order greater than C, and let $K_5 = [K_4, K_3]$.

Since there are at most d^{C^2} homomorphisms from K to each of the finitely many groups $S \times S$ with $|S| \leqslant C$, assertion (v) is clear, and all other statements follow easily, except that $K_5 = [K_5, K_3]$ in (i). To see this, we use the three-lemma, which asserts that if N_1, N_2, N_3 are normal subgroups of a group then

$$[[N_1, N_2], N_3] \leqslant [[N_2, N_3], N_1]\,[[N_3, N_1], N_2].$$

Since $K_3 = [K_3, K_3]$ we have

$$K_5 = [K_5, K_3] = [[K_3, K_3], K_5] \leqslant [K_5, K_3]$$

and the result follows. □

LEMMA 9.5. — *For each integer $q > 1$, there is an integer $k = k(q)$ with the following property. Let H be a finite group such that $H = H'$ and such that $H/Z(H)$ is a direct product of simple groups of exponent not a divisor of q. Then $H = H^{\{q\}*k}$.*

The two lemmas and the Key Theorem have the following easy but striking consequence.

COROLLARY 9.6. — *Suppose that K is a finite group and $X = X^{-1}$ is a subset containing $K^{\{q\}}$ and invariant under conjugation in K. Suppose also that d' elements of X generate K. Then the width of X in K is bounded in terms of d' and q.*

Proof. — Construct $K_1, \dots, K_5$ as in Lemma 9.4 taking C to be the greatest order of a finite simple group S with $S^q = 1$. Since K_5 is acceptable in K and since each $[k, x]$ with $k \in K$ and $x \in X$ is a product of two elements of X, we conclude from the Key Theorem that $K_5 \leqslant X^{*n_5}$ for some (d', q)-bounded n_5. Then Lemma 9.5 implies that $K_3 \leqslant K_5 X^{*n_3}$ for some q-bounded n_3. Next, applying the Key Theorem for the acceptable normal subgroup K_2/K_3 of K/K_3, we conclude that $K_2 \leqslant K_3 X^{*n_2}$ for some (d', q)-bounded n_2. Now Lemma 6.2 (a) shows that $K_1 \leqslant K_2 X^{*2d'}$. Since $K = \langle X \rangle$ we have $K \subseteq X^{*m}$ where $m = n_5 + n_3 + n_2 + 2d' + |K : K_1|$ and the result follows. □

It is now simple to deduce Theorem 9.1: one takes $q = |C/w(C)|$ where $C \cong \mathbb{Z}$, and applies Corollary 9.6 with $K = w(G)$, with $X = G^{\{w\}}$ and with d' the number of w-values required to generate $w(F_d)$, where F_d is a free group of rank d.

10. HENSEL'S LEMMA FOR GROUPS

Hensel's lemma in its simplest form states that if $f(x)$ is a polynomial with coefficients in $\mathbb{Z}_p$ and if $u \in \mathbb{Z}_p$ satisfies $f(u) \equiv 0 \pmod{p^r}$ with $r \geqslant 1$, then provided that $f'(u) \not\equiv 0 \pmod{p}$ there is some $a \in p^r \mathbb{Z}_p$ with $f(a + u) \equiv 0 \pmod{p^{r+1}}$. Since $\mathbb{Z}_p$ is complete with respect to the filtration by the powers of its maximal ideal $p\mathbb{Z}_p$, the procedure can be iterated to give an element v of $\mathbb{Z}_p$ satisfying $f(v) = 0$. Nicolas Bourbaki proved a more general version for systems of n equations $f_1, \dots, f_n$ in n variables over complete commutative rings R with maximal ideal J: the requirement on the derivative is replaced by a non-degeneracy requirement on the Jacobian matrix, and the iterative step leads from an n-tuple $\mathbf{u}$ satisfying the congruences $f_i(\mathbf{u}) \equiv 0 \pmod{J^r}$ to an n-tuple $\mathbf{a}$ of elements of J^r such that $f_i(\mathbf{a} + \mathbf{u}) \equiv 0 \pmod{J^{r+1}}$ for each i.

In order to prove their Key Theorem, Nikolov and Segal use a rather similar procedure to refine approximate solutions in groups. We follow the account given in **[13]**.

Let G be a group. Given n-tuples $\mathbf{a} = (a_1, \dots, a_n)$, $\mathbf{u} = (u_1, \dots, u_n)$ of elements of G (i.e. elements of $G^{(n)}$) we write $\mathbf{a} \cdot \mathbf{u}$ for the n-tuple $(a_1 u_1, \dots, a_n u_n)$. For a word θ, define $\theta'_{\mathbf{u}}(\mathbf{a})$ by $\theta(\mathbf{a} \cdot \mathbf{u}) = \theta'_{\mathbf{u}}(\mathbf{a}) \theta(\mathbf{u})$. At this stage we allow words to contain constant elements from G, so they are formally elements of the free product of a free

group with G. The function $\theta'_{\mathbf{u}}$ can perhaps be thought of as a type of derivative of θ at the point $\mathbf{u}$.

Now suppose that the hypotheses of the Key Theorem hold. Let $(g_1, \dots, g_m)$ be the sequence $(g_1, \dots, g_d)$ repeated many times. To prove the first assertion, fix $h \in H$, and consider the equation

$$h = \prod_{i=1}^{3} \prod_{j=1}^{m} [u_{ij}, g_j] \cdot \prod_{j=1}^{z} v_j^q = \Phi(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)\Psi(\mathbf{v}). \tag{1}$$

We need to find elements $\mathbf{u}_i = (u_{i1}, \dots, u_{im}) \in H^{(m)}$, $\mathbf{v} = (v_1, \dots, v_z) \in H^{(z)}$ solving this equation. Choose small normal subgroups $K \leqslant N$ of G, contained in H, and assume inductively that (1) can be solved modulo K. Thus

$$\Phi(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)\Psi(\mathbf{v}) = k^{-1}h$$

for some $k \in K$. We want to lift $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{v}$ to obtain a solution of (1), by replacing $\mathbf{u}_i$ by $\mathbf{a}_i.\mathbf{u}_i$ and $\mathbf{v}$ by $\mathbf{b}.\mathbf{v}$, with $\mathbf{a}_i \in N^{(m)}$, $\mathbf{b} \in N^{(z)}$. So we require

$$\Phi(\mathbf{a}_1{\cdot}\mathbf{u}_1, \mathbf{a}_2{\cdot}\mathbf{u}_2, \mathbf{a}_3{\cdot}\mathbf{u}_3)\Psi(\mathbf{b}{\cdot}\mathbf{v}) = h,$$

or equivalently

$$\Phi'_{\mathbf{u}}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)\,(\Psi'_{\mathbf{v}}(\mathbf{b}))^{\Phi(\mathbf{u})^{-1}} = k, \tag{2}$$

where $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$. The element $\Phi'_{\mathbf{u}}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ can be written as a product of commutators $[a_{ij}, g_j]^{\gamma_{ij}(\mathbf{u})}$. The analogue of the non-degeneracy requirement in Hensel's lemma is the following condition:

$$K\langle g_j^{\gamma_{ij}(\mathbf{u})} \mid j = 1, \dots, m\rangle = G \quad \text{for } i = 1, 2, 3.$$

This is incorporated into the induction hypothesis: we want to solve (2) subject to the additional condition

$$\langle g_j^{\gamma_{ij}(\mathbf{a}\cdot\mathbf{u})} \mid j = 1, \dots, m\rangle = G \quad \text{for } i = 1, 2, 3. \tag{3}$$

In the cases requiring consideration, N is a normal subgroup minimal with respect to satisfying $N = [N, G] \neq 1$. Two cases arise: either (a) N is nilpotent of class at most 2, or (b) $N/Z(N)$ is a direct power of a finite simple group S. Entirely different methods are required for the two cases. The proof of that the necessary solutions can be found is a *tour de force.*

In case (a), it is proved that (2) has many solutions with $\mathbf{b} = \mathbf{1}$ and separately that (3) holds for many triples $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) \in N^{(3m)}$. The solution sets are shown to be so large that their intersection must be non-empty, and consequently the equations (2), (3) have a simultaneous solution.

In case (b), two solution sets are again shown to be large, but the procedure splits, depending on the size of S. The proofs depend on the theorems from **[15]**, discussed in the next section.

11. TWISTED COMMUTATORS IN SIMPLE GROUPS

In order to unravel the complexities of the multiplication of group elements in finite groups having normal subgroups N with $N/Z(N)$ a direct power of a non-abelian simple group, Nikolov and Segal **[15]** have to prove some substantial new results about finite simple groups.

Let G be a group. For $x, y \in G$ and $\alpha, \beta \in \operatorname{Aut} G$, define the *twisted commutator* $T_{\alpha,\beta}(x,y) = x^{-1}y^{-1}x^{\alpha}y^{\beta}$, and let $T_{\alpha,\beta}(G,G)$ be the set of such elements. Thus in particular the derived group G' is the group generated by $T_{1,1}$. A group G is *quasisimple* if $G = G'$ and $G/Z(G)$ is simple.

THEOREM 11.1. — *There is a constant D such that if G is a finite quasisimple group and $\alpha_1, \ldots, \alpha_D, \beta_1, \ldots, \beta_D \in \operatorname{Aut} G$ then $G = \prod T_{\alpha_i,\beta_i}$.*

THEOREM 11.2. — *Given $q > 0$ there exist numbers $C(q)$, $M(q)$ with the following property.*

Suppose that G is a finite quasisimple group such that $|G/Z(G)| > C(q)$, let $\beta_1, \ldots, \beta_{M(q)} \in \operatorname{Aut} G$ and suppose that $q_1, \ldots, q_{M(q)}$ divide $M(q)$. Then there are inner automorphisms $\alpha_1, \ldots, \alpha_{M(q)}$ of G such that $G = \prod [G, (\alpha_i\beta_i)^{q_i}]$.

Theorem 11.1 is a corollary of the case of Theorem 11.2 in which $q = 1$. The authors describe these two results as 'far-reaching generalizations' of Theorem 2.3 (b) about ordinary commutators. Let us explain why these results are harder to prove than Theorem 2.3 (b). This theorem asserts that there is some constant c such that every element of every non-abelian finite simple group is a product of c commutators. The proof is simplified by an observation made in Section 2 above. Let $(S_i)_{i\in\mathbb{N}}$ be a family of finite Chevalley groups of the same type ${}^{\varepsilon}X_r$. Then any ultraproduct $\prod S_i/\mathscr{U}$ is isomorphic to a Chevalley group ${}^{\varepsilon}X_r(F)$ over a possibly infinite field F; in particular it is simple and each of its elements is a (finite) product of commutators. It follows that it is impossible to find elements $g_i \in S_i$ such that g_i cannot be expressed as a product of fewer than i commutators for each i. This leads easily to a proof that there is a bound to the widths of the word $[x,y]$ in the groups S_i.

The above argument essentially reduces the proof of Theorem 2.3 (b) to the consideration of finite simple classical groups. A reduction of this type seems unavailable in the context of Theorems 11.1 and 11.2, and instead Nikolov and Segal have to carry out a long and detailed examination of the structure of the simple groups of Lie type.

12. APPLICATIONS TO MODEL THEORY

The results reported above have given fresh impetus to the area of group theory directly concerned with word values. However, in addition to this area and profinite group theory, there is at least one other important area where the results and methods may be expected to be applied, namely the model theory of groups. This is essentially because model theory, like profinite group theory, copes well with the formation of finite products of good sets, but not generally with unions of ascending chains of such products. We end by describing two applications to the model theory of groups, one of them concerning profinite groups.

We consider properties of groups that can be formulated within the first-order language of group theory. Thus we are concerned with finite *formulae*, involving variables representing group elements, the symbols $=$, $\cdot$, $^{-1}$ and 1 representing equality, group multiplication, inverse and identity, the logical symbols $\rightarrow$ ('implies'), $\wedge$ ('and'), $\vee$ ('or'), $\neg$ ('not'), and quantifiers ranging over all elements of the group. A first-order *sentence* is a formula in which each variable is preceded by a quantifier. We omit the symbol $\cdot$ representing multiplication, and for comprehensibility we introduce abbreviations such as $\neq$, x^2, $[x, y]$, and $\bigwedge_{1 \leqslant i < j \leqslant m}$, with the obvious meanings, and abbreviations such as $\mathbf{x}$ for finite sequences of variables.

For example, let w be a word in a free group of rank r, and for each integer n consider the formula

$$\psi_{w,n}(x) : \exists \mathbf{y}_1 \ldots \exists \mathbf{y}_n \Big(\bigvee_{\varepsilon_1, \ldots, \varepsilon_n \in \{\pm 1\}} x = w(\mathbf{y}_1)^{\varepsilon_1} \ldots w(\mathbf{y}_n)^{\varepsilon_n} \Big)$$

and the sentence

$$\theta_{w,n} : \forall \mathbf{x}_1 \ldots \forall \mathbf{x}_{n+1} \Big(\bigwedge_{\varepsilon_1, \ldots, \varepsilon_{n+1} \in \{\pm 1\}} \psi_{w,n}(w(\mathbf{x}_1)^{\varepsilon_1} \ldots w(\mathbf{x}_{n+1})^{\varepsilon_{n+1}}) \Big).$$

Here each $\mathbf{x}_i$ and each $\mathbf{y}_i$ is understood to be a sequence of r variables. The formula $\psi_{w,n}$ asserts that x is a product of n w-values; $\theta_{w,n}$ asserts that products of $n+1$ w-values can be expressed as products of n w-values, and so it asserts that w has width at most n. In the language of first-order group theory, Theorem 7.1 states that if G is a d-generator profinite group and w is a d-finite word then $w(G)$ is characterized as the set of elements satisfying $\psi_{w,f}$, where the integer f depends on d, w alone.

In [7], Jarden and Lubotzky established the following result.

THEOREM 12.1. — *If G, H are profinite groups, at least one of which is finitely generated, and if G, H satisfy the same first-order sentences (regarded as abstract groups) then G, H are isomorphic.*

Let us prove this result. A well-known elementary result (see for example [**25**, Corollary 4.2.5]) asserts that two profinite groups, at least one of which is finitely generated, are isomorphic if and only if they have the same classes of finite images. Therefore it suffices to prove that G, H have the same finite images. Suppose that G is a d-generator group. Since every finite group F satisfies $w(F) = 1$ for some d-locally finite word w by Lemma 7.2, it will suffice to prove that $G/w(G)$ and $H/w(H)$ are isomorphic for all such words w. Fix w, define $\psi_{w,n}(x)$ and $\theta_{w,n}$ as above, and take n large enough for G to satisfy $\theta_{w,n}$. Therefore H satisfies $\theta_{w,n}$, and in both groups the verbal subgroup corresponding to w consists of the elements x for which $\psi_{w,n}(x)$ holds.

Now $G/w(G)$ is finite; let its elements be $\bar{g}_1, \dots, \bar{g}_m$ and define $\mu(i,j)$ for $1 \leqslant i, j \leqslant m$ by $\bar{g}_i \bar{g}_j = \bar{g}_{\mu(i,j)}$. Thus G satisfies the sentence

$$\exists x_1 \dots \exists x_m \forall g \Big(\bigwedge_{1 \leqslant i,j \leqslant m} \psi_{w,n}(x_i x_j x_{\mu(i,j)}^{-1}) \Big) \wedge \Big(\bigvee_i \psi_{w,n}(g x_i^{-1}) \Big) \wedge \Big(\bigwedge_{i \neq j} \neg \psi_{w,n}(x_i x_j^{-1}) \Big);$$

suitable elements x_i are preimages in G of the elements $\bar{g}_i$. Therefore H also satisfies this sentence. The images in $H/w(H)$ of elements x_i whose existence is asserted are distinct, comprise all elements of $H/w(H)$ and multiply in exactly the same way as the elements $\bar{g}_i$ in $G/w(G)$. Therefore $G/w(G)$ and $H/w(H)$ are isomorphic, as required.

□

A *universal* sentence is a first-order sentence of the form $\forall x_1 \dots \forall x_m \, \varphi(x_1, \dots, x_m)$ where φ is quantifier-free, and a $\forall\exists$ sentence is a sentence of the form

$$\forall x_1 \dots \forall x_m \, \exists y_1 \dots \exists y_n \, \psi(x_1, \dots, x_m, y_1, \dots, y_n)$$

where ψ is quantifier-free. The second paragraph of the above argument simply shows that a group satisfying the same first-order sentences as a finite group $\overline{G}$ is necessarily isomorphic to $\overline{G}$. An easy modification establishes the proof of the well-known stronger assertion that if $\overline{G}$, $\overline{H}$ are two structures satisfying the same universal sentences and if one of them is finite then they are isomorphic. Using this we may strengthen Theorem 12.1 as follows: *if G, H are profinite groups, at least one of which is finitely generated, and if G, H satisfy the same $\forall\exists$ sentences (regarded as abstract groups) then G, H are isomorphic.*

Various properties of groups that do not immediately appear to be first-order properties can nevertheless be expressed by a first-order sentence. For example, a group G is nilpotent of class at most 2 if its derived group G' lies in the centre. The simplest sentence expressing this, $(\forall x \in G')(\forall y \in G)\, [x,y] = 1$, is neither first-order nor an abbreviation of a first-order sentence, since G' consists of products of commutators with arbitrarily large numbers of factors. However the sentence

$(\forall x)(\forall y)(\forall z)\,[[x,y],z] = 1$ is first-order and characterizes nilpotent groups of class at most 2.

More remarkable examples arise within the context of finite groups. For example, although there is no first-order sentence characterizing the cyclic groups, or the abelian simple groups, Felgner proved that there is a sentence that is satisfied by a finite group G if and only if G is a non-abelian simple group (see [**24**, Lemma 2.5 ff.]). In [**26**] it was proved that a finite group is soluble if and only if it satisfies the sentence

$$\rho : (\forall g \forall x_1 \dots \forall x_{56} \forall y_1 \dots \forall y_{56})(g = 1 \vee g \neq [g^{x_1}, g^{y_1}] \dots [g^{x_{56}}, g^{y_{56}}]).$$

This sentence states that no non-trivial element g is a product of 56 commutators of pairs of conjugates of g. It is a triviality that a finite group is soluble if and only if no element $g \neq 1$ is a product of elements $[g^x, g^y]$ with an arbitrary number of factors, since these products comprise the derived group of the smallest normal subgroup containing g.

It is also of interest to determine which important subsets of groups are determined by first-order formulae. For example, the sentence $\zeta(x) : (\forall y)\,[x,y] = 1$ characterizes elements of the centre, and, as we saw above, certain sentences of the form $\psi_{w,n}$ characterize the elements of verbal subgroups in some groups.

The (*soluble*) *radical* $R(G)$ of a finite group G is the largest soluble normal subgroup of G. There have been many results characterizing its elements in various ways. Their proofs are generally much harder than related results characterizing soluble groups. In [**27**] it was proved that there is a first-order formula $\sigma(x)$ such that an element g of a finite group G lies in $R(G)$ if and only if $\sigma(g)$ holds in G. The proof depends in a crucial way on a version of the Key Theorem, Theorem 9.3. The point illustrated by all such results is that, although the subgroup of a group generated by a set S contains products of elements of $S \cup S^{-1}$ of lengths that are in principle unbounded, results describing circumstances in which bounds exist are likely to have striking consequences.

REFERENCES

[1] R. G. Burns & Y. Medvedev – "Analytic relatively free pro-p groups", *J. Group Theory* **7** (2004), p. 533–541.

[2] G. A. Fernandéz-Alcober & M. Morigi – "Outer commutator words are uniformly concise", *J. London Math. Soc.* (2010), doi:10.1112/jlms/jdq047.

[3] B. Hartley – "Subgroups of finite index in profinite groups", *Math. Z.* **168** (1979), p. 71–76.

[4] D. F. HOLT – "Enumerating perfect groups", *J. London Math. Soc.* **39** (1989), p. 67–78.

[5] S. V. IVANOV – "P. Hall's conjecture on the finiteness of verbal subgroups", *Izv. Vyssh. Uchem. Zaved.* **305** (1989), p. 60–70.

[6] A. JAIKIN-ZAPIRAIN – "On the verbal width of finitely generated pro-p groups", *Rev. Mat. Iberoam.* **24** (2008), p. 617–630.

[7] M. JARDEN & A. LUBOTZKY – "Elementary equivalence of profinite groups", *Bull. London Math. Soc.* **40** (2008), p. 887–896.

[8] D. LASCAR – "Autour de la propriété du petit indice", *Proc. London Math. Soc.* **62** (1991), p. 25–53.

[9] M. W. LIEBECK, E. A. O'BRIEN, A. SHALEV & P. H. TIEP – "The Ore conjecture", preprint.

[10] M. W. LIEBECK & A. SHALEV – "Diameters of finite simple groups: sharp bounds and applications", *Ann. of Math.* **154** (2001), p. 383–406.

[11] C. MARTINEZ & E. I. ZELMANOV – "Products of powers in finite simple groups", *Israel J. Math.* **96** (1996), p. 469–479.

[12] N. NIKOLOV – "Strange images of profinite groups", preprint arXiv:0901.0244.

[13] N. NIKOLOV & D. SEGAL – "Finite index subgroups in profinite groups", *C. R. Math. Acad. Sci. Paris* **337** (2003), p. 303–308.

[14] ———, "On finitely generated profinite groups. I. Strong completeness and uniform bounds", *Ann. of Math.* **165** (2007), p. 171–238.

[15] ———, "On finitely generated profinite groups. II. Products in quasisimple groups", *Ann. of Math.* **165** (2007), p. 239–273.

[16] ———, "Powers in finite groups", preprint arXiv:0909.4639.

[17] S. OATES & M. B. POWELL – "Identical relations in finite groups", *J. Algebra* **1** (1964), p. 11–39.

[18] F. POINT – "Ultraproducts and Chevalley groups", *Arch. Math. Logic* **38** (1999), p. 355–372.

[19] V. A. ROMAN'KOV – "The width of verbal subgroups of solvable groups", *Algebra and Logik* **21** (1982), p. 41–49.

[20] J. SAXL & J. S. WILSON – "A note on powers in simple groups", *Math. Proc. Cambridge Philos. Soc.* **122** (1997), p. 91–94.

[21] D. SEGAL – "On verbal subgroups of adelic groups", *J. Algebra* **326** (2011), p. 227–237.

[22] M. G. SMITH & J. S. WILSON – "On subgroups of finite index in compact Hausdorff groups", *Arch. Math. (Basel)* **80** (2003), p. 123–129.

[23] J. S. WILSON – "Two-generator conditions for residually finite groups", *Bull. London Math. Soc.* **23** (1991), p. 239–248.

[24] ———, "On simple pseudofinite groups", *J. London Math. Soc.* **51** (1995), p. 471–490.

[25] ———, *Profinite groups*, London Mathematical Society Monographs, vol. 19, The Clarendon Press, Oxford Univ. Press, 1998.

[26] ———, "Finite axiomatization of finite soluble groups", *J. London Math. Soc.* **74** (2006), p. 566–582.

[27] ———, "First-order characterization of the radical of a finite group", *J. Symbolic Logic* **74** (2009), p. 1429–1435.

[28] E. I. ZEL'MANOV – "Solution of the restricted Burnside problem for groups of odd exponent", *Math. USSR Izvestiya* **36** (1991), p. 41–60.

[29] ———, "Solutions of the restricted Burnside problem for 2-groups", *Mat. Sb.* **182** (1991), p. 568–592; translation in *Math. USSR-Sb* **72** (1992), p. 543–565.

John S. WILSON

Mathematical Institute
Oxford University
24–29 St Giles'
OXFORD OX1 3LB
United Kingdom
E-mail : wilsonjs@maths.ox.ac.uk