

GROUP LAWS IMPLYING VIRTUAL NILPOTENCE

R. G. BURNS and YURI MEDVEDEV

(Received 31 October 2001; revised 24 February 2002)

Communicated by R. B. Howlett

Abstract

If $w \equiv 1$ is a group law implying virtual nilpotence in every finitely generated metabelian group satisfying it, then it implies virtual nilpotence for the finitely generated groups of a large class \mathcal{S} of groups including all residually or locally soluble-or-finite groups. In fact the groups of \mathcal{S} satisfying such a law are all nilpotent-by-finite exponent where the nilpotency class and exponent in question are both bounded above in terms of the length of w alone. This yields a dichotomy for words. Finally, if the law $w \equiv 1$ satisfies a certain additional condition—obtaining in particular for any monoidal or Engel law—then the conclusion extends to the much larger class consisting of all ‘locally graded’ groups.

2000 *Mathematics subject classification*: primary 20F19, 20E10, 20F45.

1. Introduction

Let F denote the free group on $X = \{x_1, x_2, \dots\}$. A law (or identical relation) of a group G is an identity $w \equiv 1$ where w is a word from F , valid under every substitution $X \rightarrow G$.

The following two types of group laws are of particular interest: A *positive* (or *monoidal*) law of a group G is a nontrivial identity of the form $u \equiv v$ (or $uv^{-1} \equiv 1$) where u, v are positive words in F , that is, do not involve any of the inverses of the x_i . The *n-Engel* law is the law $[\dots[[x, y], y], \dots, y] \equiv 1$, where y appears n times, $x := x_1, y := x_2$, and $[a, b] := a^{-1}b^{-1}ab$, the *commutator* of elements a, b of an arbitrary group G . (We shall use the ‘left-normed’ convention, defining inductively $[a_1, \dots, a_n] := [[a_1, \dots, a_{n-1}], a_n]$. The *n-Engel* law then takes the form $[x, y, y, \dots, y] \equiv 1$, which is often further abbreviated to $[x, {}_n y] \equiv 1$.)

It is well known that a nilpotent-by-finite exponent group, that is, one with a nilpotent normal subgroup with quotient of finite exponent, satisfies a monoidal law ([11, 12]), and it is immediate that a nilpotent group satisfies the n -Engel law for some n . The converse of neither of these statements is true in general: there exists a 2-generator group which is not nilpotent-by-finite exponent yet satisfies a monoidal law [14], and it is relatively easy to find a group (albeit infinitely generated) which is n -Engel for some n yet not nilpotent (see for example [8, page 132] or [17, page 362, Exercise 1]).

We are interested here in the much-studied question as to the extent to which, that is, for how wide a class of groups, some versions of the converse statements do after all hold, and also in characterizing those laws which behave in this respect like monoidal and Engel laws, that is, imply ‘virtual nilpotence’ for the same large class of groups.

Our results are as follows. We write \mathcal{S} for the class of groups obtained from the class of all groups that are soluble-by-(locally finite of finite exponent) by closing under the operators L and R , where for any group-theoretic class \mathcal{X} , $L\mathcal{X}$ denotes the class of all groups locally in \mathcal{X} and $R\mathcal{X}$ the class of groups residually in \mathcal{X} (see [16, Section 1.1]). (Thus \mathcal{S} contains in particular all residually finite or soluble groups, and all locally finite or soluble groups.) We also write $\hat{\mathfrak{B}}_e$ for the variety consisting of all locally finite groups of exponent dividing e (the hat distinguishing this variety from the variety \mathfrak{B}_e of all groups of exponent dividing e), and, as usual, \mathfrak{N}_c for the variety of all nilpotent groups of class $\leq c$. (Note that the fact that the class $\hat{\mathfrak{B}}_e$ is actually a variety, is a consequence of Zelmanov’s solution of the restricted Burnside problem [24, 23].) Our main result gives a necessary and sufficient condition for a law $w \equiv 1$ to single out from the class \mathcal{S} just the nilpotent-by-finite exponent groups, furnishing in addition the information that the nilpotency class and exponent in question are bounded above solely in terms of the length of w .

THEOREM A (Compare [1, Theorem]). *Let w be any word in F with the property that every finitely generated metabelian group satisfying the law $w \equiv 1$ is necessarily nilpotent-by-finite. Then there exist positive integers $c = c(N)$, $e = e(N)$ depending only on the length N of w such that the groups in the class \mathcal{S} satisfying $w \equiv 1$ are just those in the product variety $\mathfrak{N}_c \hat{\mathfrak{B}}_e$.*

That a monoidal law (in the form $uv^{-1} \equiv 1$, where u, v are positive words) satisfies the hypothesis of Theorem A will emerge in the course of the proof. (This also follows from for instance Point [15].) That the n -Engel law does so follows from Gruenberg’s result [5] that a finitely generated soluble Engel group is nilpotent, and therefore satisfies a monoidal law.

Theorem A is stronger than related results (for example of Point [15]) in that the nilpotency class c figuring in the conclusion, as well as the exponent e , depend only on the length of w . It is this that allows the conclusion to be extended to the full class

\mathcal{S} , as opposed to just finitely generated soluble-by-finite groups, and in Theorem B below, where an additional condition is imposed on w , to the very large class of ‘locally graded’ groups.

As a consequence of Theorem A and a result of Groves [4], we infer the following

DICHOTOMY THEOREM (for words of F). *Let $w \in F$ be an arbitrary word. Exactly one of the following two possibilities occurs:*

- (i) every group in \mathcal{S} satisfying the law $w \equiv 1$ is nilpotent-by-finite exponent;
- (ii) for some $n \geq 2$, $w \equiv 1$ is a law in the restricted wreath product $C_n \wr C$, where C_n, C denote the cyclic group of order n and the infinite cyclic group respectively.

(To see this, suppose that for every prime p , $w \equiv 1$ is *not* a law in $C_p \wr C$, that is, that (ii) fails to hold. Since $C_p \wr C$ generates the product variety $\mathfrak{A}_p \mathfrak{A}$ of the variety of all abelian groups of exponent p by the variety of all abelian groups (see for example [13, Corollary 22.44]), it follows that for any soluble (in particular) variety \mathfrak{V} satisfying the law $w \equiv 1$, we have $\mathfrak{A}_p \mathfrak{A} \not\subseteq \mathfrak{V}$ for all primes p , whence by the dichotomy established by Groves [4, Theorem A (ii)], $\mathfrak{V} \subseteq \mathfrak{N}_{c_1} \hat{\mathfrak{B}}_{e_1}$ for some c_1, e_1 . Hence certainly every metabelian group satisfying $w \equiv 1$ is nilpotent-by-finite exponent, so that w satisfies the assumption of Theorem A, and thence the alternative statement (i).)

A partially related dichotomy is that of Rosenblatt [18] according to which any group G not containing the free monoid M_2 on two generators is *restrained*, that is, for every pair of elements a, b of G , the subgroup $\langle a \rangle^{(b)}$ is finitely generated (or, equivalently, every finitely generated subgroup H has finitely generated commutator subgroup $[H, H]$). (It follows easily from this that for every prime p , the wreath product $C_p \wr C$ embeds M_2 .) A negative answer to the following question would enable the replacement of the class \mathcal{S} in Theorem A and the Dichotomy Theorem by a much larger class—perhaps even that of ‘locally graded’ groups. (A group is *locally graded* if every non-trivial finitely generated subgroup has a proper subgroup of finite index.) Does there exist a group G having a minimal normal subgroup H which is infinitely generated, locally finite and of finite exponent, with G/H infinite cyclic, such that G obeys a law $w \equiv 1$ satisfying the assumption of Theorem A, that is, not holding in any $C_p \wr C$? It can be shown that such a group could not be restrained, and would therefore have to embed M_2 , in view of Rosenblatt’s result. It would thus furnish an example showing that Theorem A cannot be extended to the class of locally graded groups unless w satisfies some additional condition.

Among the laws $w \equiv 1$ satisfying the hypothesis of Theorem A, there are many—including monoidal laws and the Engel laws—that *do* satisfy a condition allowing the extension of Theorem A to the class of locally graded groups. It is well known (see [2, 3, 5, 9, 15]) that if a group G satisfies a monoidal law or an Engel law then G is *uniformly restrained*, that is, there is a positive integer k such that for all $a, b \in G$

the subgroup $\langle a \rangle^{(b)}$ can be generated by $\leq k$ elements. By Kim and Rhemtulla [9], a finitely generated, uniformly restrained, locally graded group is polycyclic-by-finite. Hence applying Theorem A we obtain immediately

THEOREM B. *Let $w \equiv 1$ be a group law with the following two properties:*

- (i) *every finitely generated metabelian group satisfying $w \equiv 1$ is nilpotent-by-finite;*
- (ii) *every group satisfying $w \equiv 1$ is uniformly restrained.*

Then there exist positive integers $c = c(N)$, $e = e(N)$ depending only on the length N of the word w , such that every locally graded group satisfying the law $w \equiv 1$ lies in $\mathfrak{N}_c \hat{\mathfrak{B}}_e$.

A natural question here is whether the qualifier ‘uniformly’ can be omitted from condition (ii).

COROLLARY 1 (Compare [2, Theorem B]). *For any monoidal law $u \equiv v$, the locally graded groups satisfying that law all lie in $\mathfrak{N}_c \hat{\mathfrak{B}}_e$ for some c , e depending only on the larger of the lengths of u , v .*

The analogue for Engel laws can be improved to the following result (see [3]), which implies, but is more detailed than, the result of Kim and Rhemtulla [9] that locally graded n -Engel groups are locally nilpotent.

COROLLARY 2 (Compare [3, Theorem]). *For each positive integer n there exist positive integers $c = c(n)$, $e = e(n)$ depending only on n such that every locally graded n -Engel group lies in $\mathfrak{N}_c \hat{\mathfrak{B}}_e \cap \hat{\mathfrak{B}}_e \mathfrak{N}_c$.*

REMARKS. 1. The class of locally graded groups, which contains the class \mathcal{C} of [2, 3], is large in — for instance — the sense that it contains nearly all the groups encountered in standard textbooks on group theory (for example [8, 17]), or that it consists of the largest class of groups ‘constructible’ from the class of finite groups by closing under most standard and some not-so-standard group-theoretic operations. (A. Rhemtulla has mentioned to us the possibility that the class of locally graded groups may coincide with that of groups with a generalized series (in one sense or another—see for example [8, page 160], [16, page 9]) with finite factors.)

2. There would seem to be considerable overlap between the main result of the present paper and that of Sarah Black [1] (and also those of [2, 3], which it generalizes). In fact [1] gives in essence a characterization of words w yielding the conclusion of our Theorem A, but in terms of the form of w itself. However the results of [1, 2, 3] depended ultimately on a lemma of Shalev [20, Lemma 3.2, Case 2] which is not quite valid as stated. (For the details, see the Remark in Section 2.3.2 below.) Thus

apart from the improvements claimed for the results of the present paper over those of [1, 2, 3], it is also justified by the need to provide arguments based on the corrected version of Shalev's lemma. Nevertheless the arguments we use, although mostly self-contained, are to a considerable extent adapted from those of Shalev [20], as well as [1, 2, 15].

The layout of the proof of Theorem A is as follows: In Section 2 we establish the finite case by means of a succession of reductions. The proof of the theorem is then completed in Section 3.

We are grateful to Olga Macedońska for helpful comments.

2. Proof of Theorem A for finite groups

Thus in this section we shall prove that if $w \equiv 1$ is a law with the property that every finitely generated metabelian group satisfying it is nilpotent-by-finite, then there exist positive integers $c = c(N)$, $e = e(N)$ depending only on the length N of w such that for every *finite* group G satisfying $w \equiv 1$ we have G^e (the subgroup generated by all e -th powers of elements of G) nilpotent of class $\leq c$.

Write F_2 for the free group of rank 2 generated by x, y . We shall need the following

LEMMA 1. *Let w be as above (that is, as in Theorem A). Then the law $w \equiv 1$ has as a consequence a 2-variable law of the form*

$$(1) \quad x(x^{s_1})^{y^{t_1}} \cdots (x^{s_k})^{y^{t_k}} = c,$$

where $k \geq 1$, the s_i and t_i are nonzero integers with $0 > t_1 > \cdots > t_k$, and c is a product of commutators in elements of the form $x^{y^\alpha} := y^{-\alpha} x y^\alpha$, $\alpha \in \mathbb{Z}$.

PROOF. Denote by W_2 the verbal subgroup of F_2 determined by w . Since $\hat{F}_2 := F_2/F_2^{(2)}$ W_2 is a finitely generated metabelian group satisfying the law $w \equiv 1$, it follows from our basic assumption concerning w that \hat{F}_2 is nilpotent-by-finite, and therefore ([11, 12]) satisfies a non-trivial monoidal (here actually semigroup) law

$$(2) \quad u \equiv v,$$

where u, v are non-trivial positive words in x and y of equal exponent sums on each of x and y . We may, by interchanging the variables x and y and cancelling, if necessary, suppose that u begins with x and v with y , and that u and v end in different symbols from $\{x, y\}$. Furthermore if u does not have xy as initial segment, then by applying the change of variables $x \rightarrow xy, y \rightarrow y$, we can ensure that it does.

Assuming this done, we lift the law (2) to F_2/W_2 , obtaining a relation in F_2/W_2 of the form

$$(3) \quad \bar{u}\bar{v}^{-1} = \bar{d},$$

where $d \in F_2^{(2)}$, and the bars indicate elements in F_2/W_2 . By virtue of the relative freeness of F_2/W_2 , the relation (3) will then define a law ($uv^{-1} \equiv d$) in any group satisfying $w \equiv 1$, that is, represents a 2-variable consequence of $w \equiv 1$.

Note that the word uv^{-1} is reduced as written (since u, v end in different symbols from $\{x, y\}$) and has exponent sums 0 on each of x, y . We now rewrite this word as a product of elements of the form $(x^s)^{y^t}$, starting on the left. Since u has xy as initial segment the first such element will be simply x . It is not difficult to see that since v^{-1} ends in y^{-1} , all other factors of the required form $(x^s)^{y^t}$ will have $t < 0$. (To illustrate, consider $u = xy^2x^2y^3x$ and $v = y^2x^2yx^2y^2$. Here

$$uv^{-1} = xy^2x^2y^3xy^{-2}x^{-2}y^{-1}x^{-2}y^{-2} = x(x^2)^{y^{-2}}x^{y^{-5}}(x^{-2})^{y^{-3}}(x^{-2})^{y^{-2}}.)$$

Then by collecting the factors $(x^s)^{y^t}$ of uv^{-1} modulo $[\langle x \rangle^{F_2}, \langle x \rangle^{F_2}]$, the law $uv^{-1} \equiv d$ implied by (3) becomes $uv^{-1} = x(x^{s_1})^{y^{t_1}} \cdots (x^{s_k})^{y^{t_k}} \gamma \equiv d$, where $k > 0, 0 > t_1 > \cdots > t_k$, the s_i are all nonzero, and $\gamma \in [\langle x \rangle^{F_2}, \langle x \rangle^{F_2}]$. Since $[\langle x \rangle^{F_2}, \langle x \rangle^{F_2}] > F_2^{(2)}$, we have $d\gamma^{-1} =: c \in [\langle x \rangle^{F_2}, \langle x \rangle^{F_2}]$, and we have arrived at a consequence of the law $w \equiv 1$ of the desired form (1). \square

COROLLARY 3. *For each $N > 0$ there exist positive integers $L = L(N), R = R(N)$ depending only on N , such that each word w of length $\leq N$ satisfying the assumption of Theorem A has a 2-variable consequence of the form (1), namely*

$$(4) \quad x(x^{s_1})^{y^{t_1}} \cdots (x^{s_k})^{y^{t_k}} \equiv c,$$

where $k > 0$, the s_i, t_i are nonzero integers, $0 > t_1 > \cdots > t_k$, c is a product of commutators in elements of the form x^{y^α} , $\alpha \in \mathbb{Z}$, and also $|t_k| \leq L$ and $|\alpha| \leq R$ for all integers α such that $x^{\pm y^\alpha}$ figures as an entry in a commutator occurring as a factor of c .

(This is immediate from Lemma 1 and the fact that there are only finitely many such words w of length $\leq N$ (in, say, x_1, \dots, x_N).

We now begin the proof of Theorem A for arbitrary finite G satisfying a law $w \equiv 1$ obeying the assumption of that theorem. The proof involves a sequence of reductions.

2.1. Reduction to the finite soluble case. We shall show that there exists a positive integer e_1 depending only on the length N of w such that G^{e_1} is soluble.

Let A/B be any non-abelian chief factor of the finite group G . (If there are no such chief factors then G is soluble and we may take $e_1 = 1$.) It is well known that A/B

decomposes as a direct product $S_1 \times S_2 \times \dots \times S_l$ of finite non-abelian simple groups all isomorphic to the same simple group S , and that each element g of G acts by conjugation on A/B in such a way as to permute the S_i ; write π_g for the permutation of $\{S_1, \dots, S_l\}$ induced by conjugation by g .

Write $c = c(x, y)$ for the right-hand side of the law (4), a consequence of $w \equiv 1$. We first show that for each $g \in G$ if any orbit under the above action of g on $\{S_1, \dots, S_l\}$ has size exceeding $2R$, then for every member S_r of that orbit we have

$$c(hB, gB) = 1 \text{ in } A/B \text{ for all } hB \in S_r.$$

This is seen as follows. Suppose $c(x, y) \neq 1$ (otherwise there is nothing to prove). Let $x^{y^{\alpha_1}}, x^{y^{\alpha_2}}$ be distinct entries in (possibly coinciding) commutator factors of c . We cannot have $S_r^{\alpha_1} = S_r^{\alpha_2}$, since this would imply that $S_r^{\alpha_1 - \alpha_2} = S_r$, and the orbit containing S_r would then have size at most $|\alpha_1 - \alpha_2| \leq 2R$ (see Corollary 3), contrary to our assumption. Hence $S_r^{\alpha_1} = S_i \neq S_j = S_r^{\alpha_2}$, yielding $[(hB)^{\alpha_1}, (hB)^{\alpha_2}] = 1$ for all $hB \in S_r$, that is, all elements of the form $(hB)^{\alpha}$ where $hB \in S_r$ and x^{y^α} is an entry in c , commute, so that $c(hB, gB) = 1$ (in G/B) for all $g \in G, hB \in S_r$, as claimed.

Suppose now that $g \in G$ and $r \in \{1, \dots, l\}$ are such that $c(hB, gB) = 1$ for all $hB \in S_r$. Since (4) is a law in G , it follows that then

$$(5) \quad h(h^{s_1})^{g^{t_1}} \dots (h^{s_k})^{g^{t_k}} \equiv 1 \text{ mod } B \text{ for all } hB \in S_r,$$

where, as in (4), $k \geq 1, 0 > t_1 > \dots > t_k$, and the s_i are all nonzero. Choose $hB \in S_r$ nontrivial (that is, $hB \neq B$); then since the cosets $(h^{s_i})^{g^{t_i}} B$ belong to various of the direct factors S_1, \dots, S_l of A/B , in order for (5) to hold we must have for at least one $j \in \{1, \dots, k\}$ that also $(h^{s_j})^{g^{t_j}} B$ is a (non-trivial) element of S_r , so that $S_r^{g^{t_j}} = S_r$ for this j . Hence in this situation the size of the orbit containing S_r under the action of π_g is bounded above by $|t_k|$, and therefore by the number L of Corollary 3.

We have thus established the following assertion: *For each $g \in G$, every orbit of the action of g by conjugation on $\{S_1, \dots, S_l\}$, the set of simple direct factors of an arbitrary non-abelian chief factor A/B of G , has size bounded above by $M := \max\{L, 2R\}$, which depends only on the length N of w .*

It follows immediately that for every simple direct factor S of every non-abelian chief factor of G we have

$$(6) \quad G^{M^l} \leq N_G(S).$$

By Jones [7] only the variety of all groups contains infinitely many pairwise non-isomorphic finite non-abelian simple groups. Hence in particular there are only a finite number of pairwise non-isomorphic groups that can occur as composition factors of groups satisfying any nontrivial law $\hat{w} \equiv 1$ of length $\leq N$ having the same property as $w \equiv 1$. Taking K to be the largest of the orders of these simple groups, it follows

that for every simple direct factor S of every non-abelian chief factor of G we have that $|\text{Aut}(S)|$ divides $K!$. From this and (6) we infer immediately that for any finite group G satisfying $w \equiv 1$ we have $G^{M!K!} \leq C_G(A/B)$ for every non-abelian chief factor A/B of G , whence

$$(7) \quad G^{M!K!} \leq \bigcap C_G(A/B),$$

where the intersection is taken over all non-abelian chief factors A/B of G . It is not difficult to show that the intersection in (7) is a soluble subgroup of G , whence the desired conclusion with $e_1 := M!K!$.

2.2. Reduction to the finite nilpotent case. By the previous reduction, we may now assume that our finite group G satisfying the law $w \equiv 1$ is soluble. We shall now show that there is a positive integer e_2 depending only on the length N of w such that G^{e_2} is nilpotent.

Let $G = G_0 \triangleright G_1 \triangleright \dots \triangleright G_r = \{1\}$ be a chief series for G . Since G is finite soluble, each chief factor G_i/G_{i+1} is finite elementary abelian. Let g be any element of G ; then conjugation by g induces an automorphism of each chief factor G_i/G_{i+1} . In the usual way we may consider G_i/G_{i+1} as an (additively written) vector space V_i over \mathbb{Z}_p (where p is the exponent of G_i/G_{i+1}) on which conjugation by g induces an invertible linear transformation T_g . Now the subgroup H/G_{i+1} of G/G_{i+1} generated by gG_{i+1} together with G_i/G_{i+1} satisfies the law $w \equiv 1$ since G does, and therefore also the 2-variable law (4). Replacing in that law y by $\bar{g}^l := g^l G_{i+1}$, for each $l = 1, 2, \dots$ in turn, and x by $\bar{h} := hG_{i+1}$ for any $h \in G_i$, we obtain

$$(8) \quad \bar{h}(\bar{h}^{s_1})^{\bar{g}^{l_1}} \dots (\bar{h}^{s_k})^{\bar{g}^{l_k}} = 1 \quad \text{for all } h \in G_i, \text{ and each } l = 1, 2, \dots,$$

since the right-hand side $c(x, y)$ of the law (4) becomes $c(\bar{h}, \bar{g}^l) = 1$. In the additive notation of the vector space V_i , equation (8) translates into

$$(I + s_1 T_g^{l_1} + s_2 T_g^{l_2} + \dots + s_k T_g^{l_k}) V_i = \{0\}.$$

Thus the linear transformations T_g^{-l} , $l = 1, 2, \dots$, all satisfy the non-constant polynomial

$$(9) \quad f(z) := 1 + s_1 z^{-l_1} + s_2 z^{-l_2} + \dots + s_k z^{-l_k} \in \mathbb{Z}_p[z].$$

At this juncture we need the following

LEMMA 2 (Compare [19, Lemma 3.3]). *Let K be any field, and $f(z) \in K[z]$ any polynomial over K of degree $d > 0$ with nonzero constant term. If I is the ideal in $K[z]$ generated by $f(z), f(z^2), \dots, f(z^{d+1})$, then there exist positive integers q, m bounded above by a function of d only, such that $(z^m - 1)^q \in I$.*

PROOF. Note that the ring $K[z]$ is a principal ideal domain and so of course also enjoys unique factorization. Let $g(z)$ be a generator of I , and let p_1, \dots, p_t be the distinct prime polynomials occurring as factors of $g(z)$. Observe that since $f(z)$ has non-zero constant term, so also do all of the p_i , and also that since $\deg g \leq \deg f = d$, we have $\deg p_i \leq d$ for all i , and $t \leq d$. Write $P_i := (p_i)$, the ideal generated by $p_i, i = 1, \dots, t$, and consider the fields $F_i := K[z]/P_i$. Regarded as having its coefficients in F_i , the polynomial $f(u)$ has roots $z + P_i, z^2 + P_i, \dots, z^{d+1} + P_i$, since $P_i \supseteq I$. Since $\deg f = d$, it follows that $z^{r_i} \equiv z^{s_i} \pmod{P_i}$ for some $0 < r_i < s_i \leq d + 1$. Now $z \notin P_i$ since the p_i all have non-zero constant term. Hence $z^{m_i} \equiv 1 \pmod{P_i}$ where $0 < m_i := s_i - r_i \leq d$. Writing $m := m_1 m_2 \cdots m_t$, we infer that

$$z^m \equiv 1 \pmod{p_1 p_2 \cdots p_t}.$$

(Note that since $m_i \leq d, t \leq d$, we have $m \leq d^d$.) Let q be largest such that p_i^q is a factor of $g(z)$ for some i . (Observe that since $\deg g \leq d$, we have $q \leq d$.) Then $(p_1 \cdots p_t)^q \in (g(z)) = I$, whence $(z^m - 1)^q \in I$, as required. \square

We now return to our reduction argument. By the lemma just proved, applied to the polynomial $f(z)$ in (9), since T_g^{-1} satisfies $f(z) = 0, f(z^2) = 0, \dots$, we infer the existence of positive integers m, q bounded above by a function of t_k only, such that

$$(10) \quad (T_g^{-m} - 1)^q = 0.$$

In terms of the multiplication in G , since g is arbitrary this translates into:

$$[h, \underbrace{g^m, \dots, g^m}_q] \in G_{i+1} \text{ for all } h \in G_i, g \in G.$$

Since i also was arbitrary, we have

$$[h_1, \underbrace{g^m, \dots, g^m}_q] \in G_{i+2} \text{ for all } h_1 \in G_{i+1}, g \in G.$$

In particular putting $h_1 = [h, \underbrace{g^m, \dots, g^m}_q]$, we obtain

$$[h, \underbrace{g^m, \dots, g^m}_{2q}] \in G_{i+2},$$

and so on. Thus by induction we conclude that for all $g, h \in G$,

$$[h, \underbrace{g^m, \dots, g^m}_{rq}] = 1,$$

that is, for every $g \in G$, g^m is an Engel element (*right Engel*, in the terminology of [6]) of G . By [6, Lemma 6.13] a finite soluble group generated by Engel elements is nilpotent; hence G^m is nilpotent.

Since m is bounded above by a function of $|t_k|$ only, by arguing as in Corollary 3 we can show that m is bounded above by a number B say, depending only on N , the length of w , and we may then take $e_2 := B!$.

2.3. The finite nilpotent case. In view of the above reductions, we may now assume that our finite group G satisfying the law $w \equiv 1$ is nilpotent; in fact, since a finite nilpotent group is the direct product of its Sylow subgroups, we may assume that G is a finite p -group for some prime p .

The proof of Theorem A in this case splits into two subcases.

2.3.1. Subcase: p sufficiently large. This subcase concerns primes $p \geq T$ where $T = T(N)$ is a positive integer depending only on N , to be defined in the course of our argument. We begin with

LEMMA 3. *As before let F denote the free group with free basis $\{x_1, x_2, \dots\}$. A law $w \equiv 1$ satisfying the assumption of Theorem A has as a consequence a law of the form*

$$(11) \quad [x_1^{\hat{e}}, \dots, x_{\hat{c}}^{\hat{e}}] \equiv \gamma_1 \gamma_2 \cdots \gamma_s,$$

where each γ_i is of the form $[\mu_i, \nu_i]$ with μ_i, ν_i commutators in $x_1, \dots, x_{\hat{c}}$ of weights ≥ 2 , and where for every $j = 1, \dots, \hat{c}$, either x_j or x_j^{-1} is an entry in μ_i or ν_i (so that each γ_i has x_j or x_j^{-1} as an entry for all $j = i, \dots, \hat{c}$).

PROOF. Let W denote the verbal subgroup of F generated by w . Then $\bar{F} := F/WF^{(2)}$ is a metabelian group satisfying the law $w \equiv 1$. Since the wreath product $C_{p_1} \wr C$ of a cyclic group of prime order p_1 with an infinite cycle, is finitely generated and metabelian but not nilpotent-by-finite, it cannot satisfy the law $w \equiv 1$. Hence in view of the fact that $C_{p_1} \wr C$ generates the product variety $\mathfrak{A}_{p_1} \mathfrak{A}$, we infer that $\mathfrak{A}_{p_1} \mathfrak{A} \not\subseteq \text{var } \bar{F}$. Hence by [4, Theorem A (ii)] $\text{var } \bar{F}$ is contained in $\mathfrak{N}_{\hat{c}} \hat{\mathfrak{B}}_{\hat{e}}$ for some \hat{c}, \hat{e} , that is, \bar{F} satisfies the law $[x_1^{\hat{e}}, \dots, x_{\hat{c}}^{\hat{e}}] \equiv 1$. Lifting this to F , we have

$$(12) \quad [x_1^{\hat{e}}, \dots, x_{\hat{c}}^{\hat{e}}] = w_1 d \quad \text{in } F,$$

for some $w_1 \in W, d \in F^{(2)}$. Since $d \in F^{(2)}$ it may be expressed as a product $d_1 \cdots d_l$ of commutators of the form $d_i = [\alpha_i, \beta_i]$ where α_i, β_i are commutators of weights ≥ 2 with all entries from among $x_1^{\pm 1}, \dots, x_{\hat{c}}^{\pm 1}$. (If any d_i had entries $x_l^{\pm 1}$ with $l > \hat{c}$, we could eliminate these d_i by applying to (12) the substitution $x_i \rightarrow x_i$ ($i = 1, \dots, \hat{c}$), $x_l \rightarrow 1$ ($l > \hat{c}$).

We now show that there is in fact an equation of the form (12) in F with d a product of commutators (as before) each of which involves *all* of $x_1^{\pm 1}, \dots, x_{\hat{c}}^{\pm 1}$ as entries. Suppose for instance that some of the d_i in $d = d_1 \cdots d_t$ do not involve x_1 . Setting $x_1 \rightarrow 1$ in (12) then yields $1 = w_2 d_{i_1} \cdots d_{i_r}$ in F , where w_2 results from w_1 , and the d_{i_j} are, in the same order as before, just those commutator factors of d not involving $x_1^{\pm 1}$. Hence $d_{i_1} \cdots d_{i_r} \in W$, whence the original right-hand side of (12) can be rewritten as

$$\begin{aligned} w_1 d &= w_1 d_{i_1} \cdots d_{i_r} (d_{i_1} d_{i_2} \cdots d_{i_r})^{-1} d_1 \cdots d_t \\ &= w_3 (d_{i_1} d_{i_2} \cdots d_{i_r})^{-1} d_1 \cdots d_t, \end{aligned}$$

where $w_3 := w_1 d_{i_1} \cdots d_{i_r} \in W$. It is now easy to see that $(d_{i_1} \cdots d_{i_r})^{-1} d_1 \cdots d_t$ can be rewritten as a product of commutators of the required form all involving $x_1^{\pm 1}$. One now continues by treating x_2 in the same way, and so on. □

We now apply this lemma to the situation of our finite p -group G (with p ‘sufficiently large’, satisfying the law $w \equiv 1$). By this lemma $w \equiv 1$ has a consequence of the form

$$[x_1^{\hat{c}}, \dots, x_{\hat{c}}^{\hat{c}}] \equiv \rho,$$

where $\rho \in F^{(2)}$ and is a product of commutators in $x_1, \dots, x_{\hat{c}}$, each involving all of $x_1, \dots, x_{\hat{c}}$. By collecting those of these commutators involving each of $x_1, \dots, x_{\hat{c}}$ exactly once, we obtain $\rho = \rho_1 \rho_2$ where now ρ_1 is a product of commutators from $F^{(2)}$ involving each x_i exactly once, and ρ_2 is a product of commutators from $F^{(2)}$ of weights $> \hat{c}$. Hence on making the substitution $x_1 \rightarrow x, x_2, \dots, x_{\hat{c}} \rightarrow y$, the word $\rho_1(x_1, \dots, x_{\hat{c}})$ becomes $\rho_1(x, y, \dots, y) = 1$, since each commutator factor of $\rho_1(x_1, \dots, x_{\hat{c}})$ is of the form $[\mu, \nu]$ where μ, ν are commutators in disjoint sets of elements from $\{x_1, \dots, x_{\hat{c}}\}$. Hence we have a law of the following form holding in G :

$$(13) \quad [x^{\hat{c}}, \underbrace{y^{\hat{c}}, \dots, y^{\hat{c}}}_{\hat{c}-1}] \equiv \rho_2(x, y, \dots, y),$$

where ρ_2 is a product of commutators from $F^{(2)}$ in x and y of weights $> \hat{c}$.

We now impose a preliminary lower bound on p , namely:

$$p > \hat{e}.$$

From this assumption and the law (13) which our p -group G satisfies, it follows that G satisfies a law of the form

$$(14) \quad [x, \underbrace{y, \dots, y}_{\hat{c}-1}] \equiv \rho_3(x, y),$$

where ρ_3 is a product of commutators in x and y (from $F^{(2)}$) of weights $> \hat{c}$.

Write $L = L(G)$ for the Lie ring determined in the usual way by the lower central series $G = G_0 > G_1 > \dots > G_k = \{1\}$ of G , that is, with additive structure given by that of the abelian group

$$(G_0/G_1) \oplus (G_1/G_2) \oplus \dots \oplus (G_{k-1}/G_k)$$

(written additively), and the Lie multiplication by the rule $[aG_i, bG_j] := [a, b]G_{i+j+1}$, for $a \in G_{i-1}, b \in G_{j-1}$, extended linearly.

Since in the law (14) the left-hand side has weight \hat{c} while $\text{wt } \rho_3 > \hat{c}$, it follows (as in the proof of [22, Lemma 6], by replacing y by $y_1 \dots y_{\hat{c}-1}$ in (14) and expanding) that $L(G)$ satisfies the ‘linearized Engel law’

$$(15) \quad \sum_{\sigma \in S_{\hat{c}-1}} [x, y_{\sigma(1)}, \dots, y_{\sigma(\hat{c}-1)}] = 0.$$

The argument now continues somewhat as in [2, page 517]. Setting $y = y_1 = \dots = y_{\hat{c}-1}$ in (15) we infer that the subring $(\hat{c} - 1)!L$ satisfies the $(\hat{c} - 1)$ -Engel condition $[x, \hat{c}-1 y] = 0$. Hence by [2, Corollary 4] (a consequence of Zelmanov’s deep results on the Engel condition in Lie algebras—see [21]) there exist positive integers e_3, c_3 depending on \hat{c} such that

$$e_3((\hat{c} - 1)!L)^{e_3} = 0,$$

that is

$$(16) \quad e_3[(\hat{c} - 1)!]^{e_3} L^{e_3} = e_4 L^{e_3} = 0,$$

where we have written $e_4 := e_3[(\hat{c} - 1)!]^{e_3}$. Back in the group G , (16) yields

$$(17) \quad (\gamma_{e_3}(G))^{e_4} \leq \gamma_{e_3+1}(G).$$

We now impose a further lower bound on p , assuming in addition that

$$p > e_4.$$

From this assumption it follows that $(\gamma_{e_3}(G))^{e_4} = \gamma_{e_3}(G)$, and then from (17) we deduce that G has class $\leq c_3$.

Hence provided $p > \max\{\hat{e}, e_4\}$, the p -group G satisfying the law $w \equiv 1$ will have nilpotency class $\leq c_3$ where \hat{e}, e_4, c_3 all depend only on the word w . Much as before, if we now take T to be the maximum of all such \hat{e}, e_4 obtained as above for all the (finitely many) words of length $\leq N$ satisfying the assumption of Theorem A, then $T = T(N)$, that is, depends only on N , and we conclude that:

If G is any p -group with $p > T(N)$ satisfying the law $w \equiv 1$, then there is a positive integer $c_4(N)$ also depending only on N (namely the largest of all the c_3 obtained as above for all words w of length $\leq N$ satisfying the assumption of Theorem A) such that G has nilpotency class $\leq c_4(N)$.

Thus the proof in Subcase 2.3.1 is complete.

2.3.2. *Subcase: $p \leq T(N)$.* We are now reduced to the situation where G is a finite p -group satisfying the law $w \equiv 1$, with $p \leq T(N)$.

We shall need the following two lemmas.

LEMMA 4 (Compare [20, Proof of Lemma 3.2, Case 2]). *Let V be a vector space over \mathbb{Z}_p and T an invertible linear transformation of V whose order s is a power of p . Suppose that there exist integers $m, q > 0$ such that $(T^m - 1)^q = 0$. Then the order s of T satisfies the inequality $s < mpq$.*

PROOF. Write $s = p^r$, $r \geq 1$. (If $s = 1$ the desired inequality holds trivially.) Since T satisfies the polynomial $z^{p^r} - 1 = (z - 1)^{p^r}$ in $\mathbb{Z}_p[z]$, the minimal polynomial of T is a power of $(z - 1)$, say $(z - 1)^l$. From $l \leq p^{r-1}$ it would follow that $T^{p^{r-1}} - 1 = (T - 1)^{p^{r-1}} = 0$, contradicting the assumption that T has order p^r . Hence $p^r \geq l > p^{r-1}$. Since by hypothesis T also satisfies $(z^m - 1)^q = 0$ we must have $mq \geq l > p^{r-1}$. Hence $mpq > p^r = s$, as required. \square

REMARK. In Case 2 of [20, Lemma 3.2] slightly more is claimed on the basis of the assumption $(T^m - 1)^q = 0$, namely that $s \leq 2mq$. However the example of $V := \mathbb{Z}_p \oplus \mathbb{Z}_p$, $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, shows this to be false in general, since here T has order p , yet $(T - 1)^2 = 0$, so that taking $m = 1$, $q = 2$, we have $2mq \not\geq p = s$ for primes $p > 3$.

LEMMA 5 (Compare [19, Lemma 3.3]). *Let $f(z)$ be a polynomial over \mathbb{Z} of degree $d > 0$ with non-zero constant term, and let J be the ideal of $\mathbb{Z}[z]$ generated by $f(z), f(z^2), \dots, f(z^{d+1})$. There exist positive integers q, m bounded above by a function of d only, and a positive integer r determined by f , such that $r(z^m - 1)^q \in J$.*

PROOF. The ideal of $\mathbb{Q}[z]$ corresponding to J is $\mathbb{Q} \otimes J =: I$, say. Applying Lemma 2 of Section 2.2 to this ideal, with $K = \mathbb{Q}$, we infer the existence of d -bounded positive integers m, q such that $(z^m - 1)^q \in I$. Hence there exists a positive integer r such that $r(z^m - 1)^q \in J$. (Clearly r —as well as m and q —is determined by the polynomial $f(z)$. However whether r is bounded above by a function of d alone is not clear.) \square

We are now ready for the proof of Theorem A in the situation of a finite p -group G with $p \leq T(N)$ satisfying the law $w \equiv 1$. The argument largely follows that of Shalev [20], with some adjustments.

Let H/K be any elementary abelian normal section of G , that is, $H, K \trianglelefteq G$, $K < H$, H/K an elementary abelian p -group, and let $g \in G$ be arbitrary. As before we denote by T_g the linear transformation of H/K considered as a vector space over \mathbb{Z}_p , induced by conjugation by g . Since G satisfies $w \equiv 1$, by Lemma 2 of Section 2.2

and the argumentation just preceding and following that lemma, there exist positive integers m, q determined by w alone, such that

$$(T_g^m - 1)^q = 0.$$

Applying Lemma 4 above, we infer that the order s ($= p^r$ say) of T_g satisfies the inequality $s < mqp$. Since in the present subcase we are assuming $p \leq T(N)$, it follows that for every $g \in G$ the order s of T_g is bounded above by a number D , say, depending only on the word w . Hence $G^{D!}$ centralizes every elementary abelian normal section of G . It is not difficult to verify (or see [20, Lemma 4.1]) that therefore $P := G^{D!}$ (or $P := G^{2D!}$ if $p = 2$) is a *powerful* p -group, that is, $P^p \geq [P, P]$ (or $P^4 \geq [P, P]$ if $p = 2$).

Now let $A = H_1/K_1$ be any abelian (not necessarily elementary abelian) normal section of G and again let $g \in G$ be arbitrary. Since the group (gK_1, A) is a cyclic extension of the abelian group A and satisfies the law $w \equiv 1$, it follows as in Section 2.2 (see equation (8)) that

$$a(\alpha_g^{s_1})^{\bar{g}^{l_1}} \cdots (\alpha_g^{s_k})^{\bar{g}^{l_k}} = 1 \quad \text{for all } a \in A, l = 1, 2, \dots,$$

where $\bar{g} := gK_1$. Write α_g for the automorphism of A induced by conjugation by g . Then (compare (9)) the automorphisms $\alpha_g^{-l}, l = 1, 2, \dots$, all satisfy (in $\text{End } A$ with A written additively) the non-constant polynomial of positive degree given by

$$f(z) = 1 + s_1z^{-l_1} + s_2z^{-l_2} + \cdots + s_kz^{-l_k} \in \mathbb{Z}[z].$$

Writing J for the ideal of $\mathbb{Z}[z]$ generated by all polynomials $h(z)$ such that $h(\alpha_g^{-1}) = 0$, we infer from Lemma 5 above the existence of positive integers r, m, q depending on $f(z)$, and therefore ultimately on w only, such that

$$(18) \quad r(\alpha_g^m - 1)^q = 0 \quad \text{in } \text{End } A.$$

Write $e_s := D!m$ and consider $Q := G^{e_s}$. Since $G^{D!m} \leq (G^{D!})^m = P^m$, it follows that $Q \leq P^m$. Since P is powerful, we have by [10, Propositions 1.7 and 4.1.7] that every element of the group P^m , and therefore every element of Q , is of the form g^m for some $g \in P$. Hence for every element $h \in Q$ we have in view of (18)

$$r(\alpha_h - 1)^q = 0 \quad \text{in } \text{End } A.$$

Since A is a p -group it follows that in fact

$$(19) \quad p^s(\alpha_h - 1)^q = 0 \quad \text{in } \text{End } A,$$

where p^s is the largest power of p dividing r .

Consider the series $A > pA > p^2A > \dots > p^tA = 0$ for A (written additively). Since each factor $p^iA/p^{i+1}A$ is elementary abelian and therefore centralized by P and hence by $Q (\leq P)$, we have for all $h \in Q$

$$(\alpha_h - 1)p^iA \leq p^{i+1}A, \quad i = 0, 1, \dots, t - 1.$$

Thus by (19) we have $(\alpha_h - 1)^{q+s} = 0$ in $\text{End } A$ for all $h \in Q$. Translated into multiplicative notation this becomes

$$(20) \quad [A, {}_{q+s}h] = \{1\} \quad \text{for all } h \in Q.$$

Write c' for the nilpotency class of Q , and $c'' := \lceil (c' + 1)/2 \rceil$. Then $\gamma_{c''}(Q)$ is a fully invariant abelian subgroup of $Q = G^{e_5}$ and hence of G , so that by (20)

$$[\gamma_{c''}(Q), {}_{q+s}h] = \{1\} \quad \text{for all } h \in Q.$$

By [20, Proposition D] it follows that

$$[\gamma_{c''}(Q), \underbrace{Q, Q, \dots, Q}_{\beta}] = \{1\}$$

for some $\beta = \beta(q + s, d(Q))$ where $d(Q)$ is the least number of generators of Q . Hence $c' \leq \beta + \lceil (c' + 1)/2 \rceil$, yielding $c' \leq \beta + (c' + 2)/2$, whence $c' \leq 2\beta + 2$.

Now since $Q = G^{e_5}$, the quotient G/Q has exponent bounded above by e_5 . Hence by Zelmanov's solution of the Restricted Burnside Problem [24, 23], $|G/Q|$ is bounded above by a function of e_5 and $d(G)$. Hence $d(Q)$ is also so bounded, whence β and therefore also the class c' of Q are so bounded.

We separate out part of what we have so far proved (for use in a slightly different context below) as a

PROPOSITION (Compare [20, Proposition 5.3]). *Given any $d > 0$, if G is any finite d -generator p -group with $p \leq T(N)$, satisfying the law $w \equiv 1$, then there exist positive integers $e_5 = e_5(N)$ and $c' = c'(N, d)$ (where N is the length of w) such that*

$$\gamma_{c'+1}(G^{e_5}) = \{1\}.$$

(Note that $G^{e_5} =: Q$ above. Note also that from the preceding argument we may immediately infer only the dependence of e_5 and c' on the word w rather than just its length. However by the usual argument (see earlier), involving possible replacement of e_5 and c' by larger numbers, we may in fact assume dependence only on N —as well as d in the case of c' .)

We now return to the powerful p -group P ($:= G^{D^!}$). Let g_1, g_2 be any elements of P . Applying the above proposition to $P_0 := \langle g_1, g_2 \rangle$ we have, since P_0 is 2-generator, that

$$(21) \quad \gamma_{(c(N,2)+1)}(P_0^{e_5}) = \{1\}.$$

Writing $c_5(N) := c'(N, 2)$, it follows from (21) in particular that $[g_1^{e_5}, c_5 g_2^{e_5}] = 1$ for all $g_1, g_2 \in P$, that is, P satisfies the law

$$(22) \quad [x^{e_5}, c_5 y^{e_5}] = 1,$$

where e_5, c_5 depend only on N .

The proof now continues as in [2, page 515], starting from the second displayed formula on that page. (Note that the first displayed formula there is incorrect.) That argument starts with a powerful p -group satisfying the law $[x^{e_1}, c_1 y^{e_1}] \equiv 1$ for some e_1, c_1 , and concludes (at the top of p.518) that there exist numbers e_3, c_3 depending only on e_1, c_1 such that the original powerful p -group belongs to $\mathfrak{N}_{e_3} \hat{\mathfrak{B}}_{e_3}$. Translating this into our present notation, we infer the existence of numbers e_6, c_6 depending only on e_5, c_5 such that $P \in \mathfrak{N}_{c_6} \hat{\mathfrak{B}}_{e_6}$. Then since $P = G^{D^!}$ we have the desired conclusion for G , completing the proof in Subcase 2.3.2, and with that the proof of Theorem A for finite groups. □

3. Completion of the proof of Theorem A

We now show that any finitely generated soluble-by-finite group G satisfying the law $w \equiv 1$ lies in the variety $\mathfrak{N}_c \mathfrak{B}_e$ for some $c = c(N), e = e(N)$ depending only on the length N of w . In view of this dependence of c, e only on N , Theorem A then follows immediately.

The argument will follow that of [2] beginning on page 518, adapted to our present situation. (In particular the context in [2, page 518 *et seqq.*] of a generating subsemigroup S of G satisfying a semigroup law will no longer be relevant. However the proof given there of Theorem D to the effect that if S satisfies a semigroup law then that law actually holds in G (and furthermore $G \in \mathfrak{N}_{c(N)} \hat{\mathfrak{B}}_{e(N)}$) remains valid.)

With G , as above, any finitely generated soluble-by-finite group satisfying the law $w \equiv 1$, suppose that m is such that G^m is soluble. Observe first that by a well-known result of Hall (see for example [17, 15.4.1, page 453]) the quotient

$$\bar{G} := G/[\gamma_k(G^m), \gamma_k(G^m)]$$

is residually finite for every k since it is abelian-by-nilpotent-by-finite. Hence there is a chain $\bar{G} \geq \bar{G}_1 \geq \bar{G}_2 \geq \dots$ of normal subgroups of \bar{G} such that $|\bar{G} : \bar{G}_i| < \infty$

and $\bigcap \bar{G}_i = \{1\}$. Since \bar{G}/\bar{G}_i is finite and satisfies the law $w \equiv 1$, we have by the finite case of the theorem, already established, that $\bar{G}/\bar{G}_i \in \mathfrak{N}_c \hat{\mathfrak{B}}_e$ for some $c = c(N)$, $e = e(N)$ depending only on the length N of w . Since \bar{G} is a subcartesian product of the \bar{G}/\bar{G}_i , it follows that $\bar{G} \in \mathfrak{N}_c \hat{\mathfrak{B}}_e$, or, equivalently,

$$(23) \quad \gamma_c(G^e) \leq [\gamma_k(G^m), \gamma_k(G^m)] \quad \text{for all } k.$$

Consider now the descending chain $G = N_0 \geq N_1 \geq N_2 \geq \dots$ of subgroups of finite index in G , defined inductively by

$$N_1 := G^{me}, \quad N_{i+1} := N_i^e, \quad \text{for } i \geq 1.$$

Since the N_i are all finitely generated subgroups of G , the argument leading to (23) applies to each of them, so that

$$\gamma_c(N_i^e) \leq [\gamma_k(N_i), \gamma_k(N_i)] \quad \text{for all } i, k,$$

that is,

$$\gamma_c(N_{i+1}) \leq [\gamma_k(N_i), \gamma_k(N_i)] \quad \text{for all } i, k.$$

Hence writing l for the solubility length of G^m , we have in particular

$$\gamma_c(N_{l+1}) \leq \gamma_c(N_l)^{(1)} \leq \gamma_c(N_{l-1})^{(2)} \leq \dots \leq \gamma_c(N_1)^{(l)} = \{1\},$$

where the equality follows from the fact that $N_1 = G^{me} \leq G^m$ and G^m is soluble of length l . Hence G is a finite extension of a group of class $\leq c$, namely N_{l+1} , and is therefore residually finite, whence we infer, as earlier, that $G \in \mathfrak{N}_c \hat{\mathfrak{B}}_e$. \square

References

- [1] S. Black, 'Which words spell "almost nilpotent" ?', preprint, 1997.
- [2] R. G. Burns, O. Macedońska and Y. Medvedev, 'Groups satisfying semigroup laws, and nilpotent-by-Burnside varieties', *J. Algebra* **195** (1997), 510–525.
- [3] R. G. Burns and Y. Medvedev, 'A note on Engel groups and local nilpotence', *J. Austral. Math. Soc. Ser. A* **64** (1998), 92–100.
- [4] J. R. J. Groves, 'Varieties of soluble groups and a dichotomy of P. Hall', *Bull. Austral. Math. Soc.* **5** (1971), 391–410.
- [5] K. W. Gruenberg, 'Two theorems on Engel groups', *Math. Proc. Cambridge Philos. Soc.* **49** (1953), 377–380.
- [6] B. Huppert, *Endliche Gruppen I* (Springer, Berlin, 1967).
- [7] G. A. Jones, 'Varieties and simple groups', *J. Austral. Math. Soc.* **17** (1974), 163–173.
- [8] M. I. Kargapolov and Ju. I. Merzljakov, *Fundamentals of the theory of groups* (Springer, Berlin, 1979).

- [9] Y. K. Kim and A. H. Rhemtulla, 'Weak maximality condition and polycyclic groups', *Proc. Amer. Math. Soc.* **123** (1995), 711–714.
- [10] A. Lubotzky and A. Mann, 'Powerful p -groups, I. Finite groups', *J. Algebra* **105** (1987), 484–505.
- [11] A. I. Mal'cev, 'Nilpotent semigroups', *Uchen. Zap. Ivanovsk. Ped. Inst.* **4** (1953), 107–111.
- [12] B. H. Neumann and T. Taylor, 'Subsemigroups of nilpotent groups', *Proc. Roy. Soc. Ser. A* **274** (1963), 1–4.
- [13] H. Neumann, *Varieties of groups* (Springer, Berlin, 1967).
- [14] A. Yu. Olshanskiĭ and A. Storozhev, 'A group variety defined by a semigroup law', *J. Austral. Math. Soc. Ser. A* **60** (1996), 255–259.
- [15] F. Point, 'Milnor identities', *Comm. Algebra* **24** (1996), 3725–3744.
- [16] D. J. S. Robinson, *Finiteness conditions and generalized soluble groups. Part I* (Springer, Berlin, 1972).
- [17] ———, *A course in the theory of groups* (Springer, Berlin, 1982).
- [18] J. M. Rosenblatt, 'Invariant measures and growth conditions', *Trans. Amer. Math. Soc.* **193** (1974), 33–53.
- [19] J. F. Semple and A. Shalev, 'Combinatorial conditions in residually finite groups, I', *J. Algebra* **157** (1993), 43–50.
- [20] A. Shalev, 'Combinatorial conditions in residually finite groups, II', *J. Algebra* **157** (1993), 51–62.
- [21] E. I. Zelmanov, 'Engel Lie algebras', *Dokl. Akad. Nauk SSSR* **292** (1987), 265–268.
- [22] ———, 'On some problems of group theory and Lie algebras', *Math. USSR-Sbornik* **66** (1990), 159–168.
- [23] ———, 'The solution of the restricted Burnside problem for 2-groups', *Mat. Sb.* **182** (1991), 568–592.
- [24] ———, 'The solution of the restricted Burnside problem for groups of odd exponent', *Math. USSR-Izv.* **36** (1991), 41–60.

Department of Mathematics and Statistics
York University
Toronto, Ontario
Canada
e-mail: rburns@pascal.math.yorku.ca

Bank of Montreal
Toronto, Ontario
Canada M3J 1P3