PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 146, Number 4, April 2018, Pages 1529–1536
http://dx.doi.org/10.1090/proc/13867
Article electronically published on December 7, 2017

PROBABILISTICALLY NILPOTENT GROUPS

ANER SHALEV

(Communicated by Pham Huu Tiep)

ABSTRACT. We show that, for a finitely generated residually finite group Γ, the word $[x_1, \ldots, x_k]$ is a probabilistic identity of Γ if and only if Γ has a finite index subgroup which is nilpotent of class less than k.

Related results, generalizations and problems are also discussed.

1. INTRODUCTION

A well-known result of Peter Neumann [N] shows that a finite group G in which the probability that two random elements commute is at least $\epsilon > 0$ is bounded-by-abelian-by-bounded; this means that there are normal subgroups N, K of G such that $K \leq N$, N/K is abelian, and both $|G/N|$ and $|K|$ are bounded above by some function of ϵ.

The probability that two elements commute received considerable attention over the years; see for instance [G], [J], [LP], [GR], [GS], [H], [NY], [E]. However, the natural extension to longer commutators and the probability of them being 1 remained unexplored. Here we shed some light on this problem, providing some results, characterizations and directions for further investigations.

Neumann's result, as well as a similar result of Mann on groups with many involutions [M1], can be viewed in the wider context of the theory of word maps (see for instance the survey paper [S] and the references therein) and the notion of probabilistic identities.

A word $w = w(x_1, \ldots, x_k)$ is an element of the free group F_k on $x_1, \ldots, x_k$. Given a group G, the word w induces a word map $w_G : G^k \to G$ induced by substitution. We denote the image of this word map by $w(G)$.

If G is finite, then w induces a probability distribution $P_{G,w}$ on G, given by

$$P_{G,w}(g) = |w_G^{-1}(g)|/|G|^k.$$

A similar distribution is defined on profinite groups G, using their normalizer Haar measure.

This distribution has been studied extensively in recent years, with particular emphasis on the case where G is a finite simple group; see [DPSS], [GS], [LS1], [LS2], [LS3], [LS4], [B]. Here we focus on general finite groups and residually finite groups, and the proofs of our results do not use the Classification of finite simple group.

Received by the editors June 8, 2017 and, in revised form, June 20, 2017.
2010 *Mathematics Subject Classification.* Primary 20E26; Secondary 20P05.
The author was partially supported by ERC advanced grant 247034, ISF grant 1117/13 and the Vinik Chair of mathematics which he holds.

Let Γ be a residually finite group. Recall that a word $w \neq 1$ is said to be a *probabilistic identity* of Γ if there exists $\epsilon > 0$ such that for each finite quotient $H = \Gamma/\Delta$ of Γ we have $P_{H,w}(1) \geq \epsilon$. This amounts to saying that, in the profinite completion $G = \widehat{\Gamma}$ of Γ, we have $P_{G,w}(1) > 0$. We shall also be interested in finite and profinite groups G in which $P_{G,w}(g) \geq \epsilon > 0$ for some element $g \in G$.

While Neumann's result deals with the commutator word $[x_1, x_2]$ of length two, here we consider commutator words of arbitrary length. Define inductively $w_1 = x_1$ and $w_{k+1} = [w_k, x_{k+1}]$. Thus w_k is the left normed commutator $[x_1, \ldots, x_k]$.

Our main result characterizes finitely generated residually finite groups in which such a word is a probabilistic identity. Clearly, if Γ has a nilpotent normal subgroup of finite index m and class $< k$, and $G = \widehat{\Gamma}$, then $P_{G,w_k}(1) \geq m^{-k} > 0$, so w_k is a probabilistic identity of Γ. It turns out that the converse is also true.

Theorem 1.1. *Let Γ be a finitely generated residually finite group, and let k be a positive integer. Then the word $[x_1, \ldots, x_k]$ is a probabilistic identity of Γ if and only if Γ has a finite index normal subgroup Δ which is nilpotent of class less than k.*

In fact our proof shows more, namely: if Γ (or its profinite completion) is generated by d elements, and for some fixed $\epsilon > 0$ and every finite quotient H of Γ there exists $h \in H$ such that $P_{H,w_k}(h) \geq \epsilon$, then the index $|\Gamma : \Delta|$ of the nilpotent subgroup Δ above divides $n!^{n!^d}$, where $n = \lfloor k/\epsilon \rfloor$.

Since a coset identity is a probabilistic identity we immediately obtain the following.

Corollary 1.2. *Let Γ be a finitely generated residually finite group, and let k be a positive integer. Suppose there exist a finite index subgroup Δ of Γ, and elements $g_1, \ldots, g_k \in \Gamma$ satisfying $[g_1\Delta, \ldots, g_k\Delta] = 1$. Then there exists a finite index subgroup Δ_0 of Γ such that $\gamma_k(\Delta_0) = 1$.*

Here and throughout this paper $\gamma_k(G)$ denotes the kth term of the lower central series of a group G. Theorem 1.1 follows from an effective result on finite groups, which is of independent interest. To state it we need some notation.

For groups G and H we define

$$G(H) = \bigcap_{\phi: G \to H} \ker \phi,$$

namely, the intersection of all kernels of homomorphisms from G to H. Then $G(H) \lhd G$, $G/G(H) \leq H^l$ for some l (which is finite if G is finite). Moreover, $G/G(H)$ satisfies all the identities of H, namely it lies in the variety generated by H.

The case $H = S_n$, the symmetric group of degree n, will play a role below. For a positive integer n define

$$L(n) = lcm(1, 2, \ldots, n),$$

where *lcm* stands for the least common multiple. Clearly, $L(n)$ is the exponent of S_n. It is well known (and easy to verify using the Prime Number Theorem) that $L(n) = e^{(1+o(1))n}$. There are considerably shorter identities for S_n, but the length of its shortest identity is still unknown.

Theorem 1.3. *Fix a positive integer k and a real number $\epsilon > 0$.*

Let $w_k = [x_1, \ldots, x_k]$. Let G be a finite group and suppose that for some $g \in G$ we have $P_{G,w_k}(g) \geq \epsilon$. Set $n = \lfloor k/\epsilon \rfloor$ and let $N = G(S_n)$. Then N is nilpotent of class less than k.

Furthermore, $G/N \leq S_n^l$ for some l, and it satisfies all the identities of S_n. In particular, G/N has exponent dividing $L(\lfloor k/\epsilon \rfloor)$.

The following is an immediate consequence for residually finite groups which are not necessarily finitely generated.

Corollary 1.4. *If w_k is a probabilistic identity of a residually finite group Γ, then Γ is an extension of a nilpotent group of class less than k by a group of finite exponent.*

In [LS3] the following problem is posed.

Problem 1.5. Do all finitely generated residually finite groups Γ which satisfy a probabilistic identity w satisfy an identity?

This seems to be a rather challenging problem. Till recently the only non-trivial cases where a positive answer was known were $w = [x_1, x_2]$ and $w = x_1^2$.

In [LS3] an affirmative answer to Problem 1.5 is given for all words w, provided the group Γ is linear.

In [LS4] it is shown that, if a residually finite group Γ (not necessarily finitely generated) satisfies a probabilistic identity, then the non-abelian upper composition factors of Γ have bounded size. This leads to solutions of problems from [DPSS] and [B].

The next result provides a positive answer to Problem 1.5 for additional words w; it also deals with groups which are not finitely generated.

Corollary 1.6. *Every residually finite group which satisfies the probabilistic identity w_k satisfies an identity.*

Indeed, Corollary 1.4 shows that our group satisfies the identity $[x_1^c, \ldots, x_k^c]$ for some positive integer c.

Definition 1.7. A word $w \in F_k$ is said to be *good* if for any real number $\epsilon > 0$ there exists a word $1 \neq v \in F_m$ (for some m) depending only on w and ϵ such that, if G is a finite group satisfying $P_{G,w}(g) \geq \epsilon$ for some $g \in G$, then v is an identity of G.

For example, results from [N], [M1] and [M2] imply that the words $[x_1, x_2]$ and x_1^2 are good. It is easy to see that, if w is a good word, and v is any word disjoint from w (namely, their sets of variables are disjoint), then the words wv and vw are also good.

Theorem 1.3 above shows that the words $[x_1, \ldots, x_k]$ are good for all k. In fact, we can generalize the latter result as follows.

Proposition 1.8. *Let $w(x_1, \ldots, x_k)$ be a word, and let*

$$w'(x_1, \ldots, x_{k+1}) = [w(x_1, \ldots, x_k), x_{k+1}].$$

Suppose w is good. Then so is w'.

Induction on k immediately yields the following.

Corollary 1.9. *The words $[x_1^2, x_2, \ldots, x_k]$ $(k \geq 2)$ are good.*

In fact we also obtain a related structure theorem as follows.

Theorem 1.10. *(i) Let G be a finite group, let $w = [x_1^2, x_2, \dots, x_k]$, and suppose $P_{G,w}(g) \geq \epsilon > 0$ for some $g \in G$. Then there exists $n = n(k, \epsilon)$ depending only on k and ϵ, such that $G(S_n)$ is nilpotent of class at most k.*

(ii) If w above is a probabilistic identity of a finitely generated residually finite group Γ, then Γ has a finite index subgroup which is nilpotent of class at most k.

2. Proofs

In this section we prove Proposition 1.8, Theorem 1.3, Theorem 1.1 and Theorem 1.10, which in turn imply the other results stated in the Introduction.

Lemma 2.1. *Let $w(x_1, \dots, x_k)$ be a word, and let*

$$w'(x_1, \dots, x_{k+1}) = [w(x_1, \dots, x_k), x_{k+1}].$$

Let G be a finite group, and suppose $P_{G,w'}(g) \geq \epsilon > 0$ for some $g \in G$. Choose $g_1, \dots, g_k \in G$ uniformly and independently. Then, for every $0 < \delta < \epsilon$,

$$Prob(|G : C_G(w(g_1, \dots, g_k))| < 1/\delta) > \epsilon - \delta.$$

Proof. Choose $g_{k+1} \in G$ also uniformly and independently. Then

$$Prob([w(g_1, \dots, g_k), g_{k+1}] = g) = P_{G,w'}(g) \geq \epsilon.$$

Given $g_1, \dots, g_k, g \in G$, the number of elements $g_{k+1} \in G$ satisfying

$$[w(g_1, \dots, g_k), g_{k+1}] = g$$

is at most $|C_G(w(g_1, \dots, g_k))|$. This yields

$$\epsilon |G|^{k+1} \leq \sum_{g_1, \dots, g_k \in G} |C_G(w(g_1, \dots, g_k))|.$$

Let

$$p = Prob(|C_G(w(g_1, \dots, g_k))| > \delta |G|) = Prob(|G : C_G(w(g_1, \dots, g_k))| < 1/\delta).$$

Then we obtain

$$\epsilon \leq p + (1-p)\delta,$$

so $p \geq (\epsilon - \delta)/(1 - \delta) > \epsilon - \delta$, as required. □

Proposition 2.2. *Let G, k, w, w', ϵ be as in Lemma 2.1. Suppose $P_{G,w'}(g) \geq \epsilon$ for some $g \in G$ and let $0 < \delta < \epsilon$. Let $N = G(S_n)$, where $n = \lfloor 1/\delta \rfloor$, and set $M = C_G(N)$. Then $P_{G/M,w}(1) > \epsilon - \delta$.*

Proof. Using Lemma 2.1 we obtain

$$Prob(|G : C_G(w(g_1, \dots, g_k))| < 1/\delta) > \epsilon - \delta.$$

Clearly, if $|G : C_G(w(g_1, \dots, g_k))| < 1/\delta$, then the permutation representation of G on the cosets of $C_G(w(g_1, \dots, g_k))$ gives rise to a homomorphism $\phi : G \to S_n$ satisfying $G(S_n) \leq \ker(\phi) \leq C_G(w(g_1, \dots, g_k))$. This implies that

$$Prob(N \leq C_G(w(g_1, \dots, g_k))) > \epsilon - \delta.$$

Since $M = C_G(N)$ we obtain

$$Prob(w(g_1, \dots, g_k) \in M) > \epsilon - \delta,$$

so $P_{G/M,w}(1) > \epsilon - \delta$, as required. □

We now prove Proposition 1.8.

Proof. Recall that $w \in F_k$ is a good word and $w' = [w, x_{k+1}]$.

To show that w' is good, suppose $P_{G,w'}(g) \geq \epsilon > 0$ for some $g \in G$. Set $\delta = \epsilon/2$, $n = \lfloor 1/\delta \rfloor = \lfloor 2/\epsilon \rfloor$ and apply Proposition 2.2. We obtain $P_{G/M,w}(1) > \epsilon/2$, where $M = C_G(G(S_n))$. Since w is good there is a word $1 \neq v \in F_m$ depending on w and ϵ such that $v(G) \leq M$.

Since $g^{L(n)} \in G(S_n)$ for all $g \in G$, it follows that $v' = [v, x_{m+1}^{L(n)}]$ is an identity of G, which depends on w' and ϵ. Therefore w' is good. □

Since x_1 is a good word, it now follows from Proposition 1.8 by induction on k that all commutator words $[x_1, \dots, x_k]$ are good. We can now also prove the more refined Theorem 1.3.

Proof. We prove, by induction on $k \geq 1$, that, under the assumptions of the theorem, for $n = \lfloor k/\epsilon \rfloor$ and $N = G(S_n)$, we have $\gamma_k(N) = 1$. The other statements of the theorem follow immediately.

If $k = 1$, then $|G|^{-1} = P_{G,x_1}(g) \geq \epsilon$ for some $g \in G$. This yields $|G| \leq 1/\epsilon$. Let $n = \lfloor 1/\epsilon \rfloor$. Then $N = G(S_n) = 1$, which yields the induction base.

Now, suppose the theorem holds for k and we prove it for $k+1$. We assume $P_{G,w_{k+1}}(g) \geq \epsilon$ and let $n = \lfloor (k+1)/\epsilon \rfloor$, $N = G(S_n)$ and $M = C_G(N)$.

Using Proposition 2.2 with $w = w_k, w' = w_{k+1}$ and $\delta = \epsilon/(k+1)$ we obtain

$$P_{G/M,w_k}(1) > k\epsilon/(k+1).$$

By induction hypothesis this implies that $(G/M)(S_{\lfloor k/(k\epsilon/(k+1)) \rfloor})$ is nilpotent of class less than k. Since $k/((k\epsilon/(k+1)) = (k+1)/\epsilon$ we see that $(G/M)(S_n)$ is nilpotent of class less than k. Clearly $(G/M)(S_n) \geq G(S_n)M/M$, and this yields

$$\gamma_k(G(S_n)) \leq M.$$

Therefore

$$\gamma_{k+1}(N) = [\gamma_k(N), N] \leq [M, N] = 1.$$

This completes the proof. □

We can now prove Theorem 1.1. This result follows easily from Corollary 1.4 using Zelmanov's solution to the Restricted Burnside Problem, which, for general exponents, also relies on the Classification of finite simple groups. However, we are able to provide an elementary proof of Theorem 1.1 which avoids these very deep results.

Proof. It suffices to show that, if G is a d-generated finite group satisfying $P_{G,w_k}(1) \geq \epsilon$, then G has a normal subgroup N which is nilpotent of class less than k, such that $|G/N|$ is bounded above in terms of d, k, ϵ only.

Using Theorem 1.3 and its notation, $N = G(S_n)$ is nilpotent of class less than k. We also have $G/N \leq S_n^l$ for some l. Thus G/N is a d-generated group lying in the variety generated by S_n.

A classical result of B. H. Neumann [Ne, 14.3] states that, for every finite group H and a positive integer d, the free d-generated group in the variety generated by H is finite of order dividing $|H|^{|H|^d}$. This implies that

$$|G/N| \leq n!^{n!^d}.$$

Since $n = \lfloor k/\epsilon \rfloor$, $|G/N|$ is bounded above in terms of d, k and ϵ.

This completes the proof of Theorem 1.1. □

Note that the conclusion of Theorem 1.1 holds under the weaker assumption that, for every finite quotient H of Γ there exists an element $h \in H$ such that $P_{H,w_k}(h) \geq \epsilon > 0$. Indeed this follows from Theorem 1.3 as above, replacing $P_{G,w_k}(1)$ by $P_{G,w_k}(g)$.

Finally, we prove Theorem 1.10.

Proof. We prove part (i) of the theorem by induction on k, starting with $k = 1$ and $w = x_1^2$.

By Proposition 5 of [M2], if $P_{G,x_1^2}(g) \geq \epsilon > 0$, then $P_{G,[x_1,x_2]}(1) \geq \epsilon^2$.

By Theorem 1.3 this implies that $G(S_n)$ is abelian, where $n = \lfloor 2/\epsilon^2 \rfloor$.

Now suppose the result holds for $k \geq 1$ and let us prove it for $k+1$. Set $w = [x_1^2, x_2, \ldots, x_k]$ and $w' = [w, x_{k+1}]$.

We assume $P_{G,w'}(g) \geq \epsilon > 0$. Choose $0 < \delta < \epsilon$ (depending on ϵ), and set $m = \lfloor 1/\delta \rfloor$ and $M = C_G(G(S_m))$. Then, by Proposition 2.2 we have $P_{G/M,w}(1) > \epsilon - \delta$. By induction hypothesis there exists $n = n(k, \epsilon - \delta)$ such that $(G/M)(S_n)$ is nilpotent of class at most k. This yields

$$\gamma_{k+1}(G(S_n)) \leq M.$$

Therefore

$$[\gamma_{k+1}(G(S_n)), G(S_m)] \leq [M, N] = 1.$$

Define $n(k+1, \epsilon) = \max(n(k, \epsilon - \delta), \lfloor 1/\delta \rfloor)$. Then it follows that, for

$$n' = n(k+1, \epsilon) = \max(n, m)$$

we have

$$\gamma_{k+2}(G(S_{n'})) = 1,$$

proving part (i).

Part (ii) follows from part (i) as in the proof of Theorem 1.1. □

By choosing δ to be a suitable function of ϵ (so that $m = n$ at each inductive step) one may obtain explicit good bounds on $n(k, \epsilon)$. We leave this for the interested reader.

3. Related problems

We conclude with some natural questions and directions for further research.

Problem 3.1. Characterize residually finite groups in which $[x_1, \ldots, x_k]$ is a probabilistic identity.

In particular, do these groups have a finite index subgroup which is nilpotent of class less than k?

The answer is positive for $k = 2$. Indeed, a result of Lévai and Pyber [LP, 1.1(iii)] shows that a profinite group G in which $[x_1, x_2]$ is a probabilistic identity has an open abelian subgroup, whose finite index need not be bounded in terms of $P_{G,[x_1,x_2]}(1)$. This implies a similar result for residually finite groups.

By Theorem 1.1, the answer to the question above is positive for all k, provided the ambient group Γ (or its profinite completion) is finitely generated. In the general case it follows that, for some n, $\Gamma(S_n)$ (which may have infinite index in Γ) is nilpotent of class less than k. But the reverse implication does not hold, as the product of infinitely many copies of S_n demonstrates.

Recall that a word w is said to be good if for every $\epsilon > 0$ there exists a word $v \neq 1$, depending only on w and ϵ, such that, for every finite group G, if $P_{G,w}(g) \geq \epsilon$ for some $g \in G$, then v is an identity of G.

Problem 3.2. Are all non-identity words good?

This does not seem likely (or provable), so it would be nice to find an example of a word which is not good.

Problem 3.3. Characterize the good words, or at least find more examples of them.

This is particularly interesting for some specific words.

Problem 3.4. Are power words x_1^k good?

Let us say that general commutator words are words constructed from the variables x_k ($k \geq 1$) in finitely many steps in which we pass from previously constructed words w_1, w_2 in disjoint sets of variables to the word $[w_1, w_2]$.

For examples, let $\delta_0(x_1) = x_1$ and for $k \geq 1$ set

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

Thus a group satisfies the identity δ_k if and only if it is solvable of derived length at most k.

Problem 3.5. Are general commutator words good? Are the words δ_k good?

A positive answer would of course follow from a positive answer to the following.

Problem 3.6. Suppose w_1, w_2 are good words in disjoint sets of variables. Does it follow that the word $[w_1, w_2]$ is good?

It would be nice to find analogues of Theorem 1.1 where we replace nilpotency by solvability.

Problem 3.7. Let Γ be a finitely generated residually finite group and suppose the word δ_k is a probabilistic identity of Γ. Does it follow that Γ has a solvable subgroup Δ of finite index? Can we further require that the derived length of Δ is at most k?

Finally, for a finite group G, set

$$Pr_k(G) = P_{G,w_k}(1),$$

the probability that $[g_1, \ldots, g_k] = 1$ in G. Note that $Pr_2(G)$, denoted in the literature by $Pr(G)$ and $cp(G)$, was widely studied by various authors. It was shown in [G] that the maximal value of $Pr(G)$ for G non-abelian is $5/8$.

Problem 3.8. Given $k \geq 3$, find the maximal value of $Pr_k(G)$ for finite groups G satisfying $\gamma_k(G) \neq 1$.

The set $\{Pr(G) : G \text{ a finite group}\}$ also received considerable attention; see [J], [H] and [E]. In the latter paper, Eberhard shows that this set is well ordered by $>$ and that its limit points are all rational.

Problem 3.9. For $k \geq 3$, study the set $\{Pr_k(G) : G \text{ a finite group}\}$. Are its limit points all rational? Is it well ordered by $>$?

References

[B] A. Bors, Fibers of automorphic word maps and an application to composition factors, arXiv math:1608.00131.

[DPSS] John D. Dixon, László Pyber, Ákos Seress, and Aner Shalev, *Residual properties of free groups and probabilistic methods*, J. Reine Angew. Math. **556** (2003), 159–172, DOI 10.1515/crll.2003.019. MR1971144

[E] Sean Eberhard, *Commuting probabilities of finite groups*, Bull. Lond. Math. Soc. **47** (2015), no. 5, 796–808, DOI 10.1112/blms/bdv050. MR3403962

[GS] Shelly Garion and Aner Shalev, *Commutator maps, measure preservation, and T-systems*, Trans. Amer. Math. Soc. **361** (2009), no. 9, 4631–4651, DOI 10.1090/S0002-9947-09-04575-9. MR2506422

[GR] Robert M. Guralnick and Geoffrey R. Robinson, *On the commuting probability in finite groups*, J. Algebra **300** (2006), no. 2, 509–528, DOI 10.1016/j.jalgebra.2005.09.044. MR2228209

[G] W. H. Gustafson, *What is the probability that two group elements commute?*, Amer. Math. Monthly **80** (1973), 1031–1034, DOI 10.2307/2318778. MR0327901

[H] Peter Hegarty, *Limit points in the range of the commuting probability function on finite groups*, J. Group Theory **16** (2013), no. 2, 235–247, DOI 10.1515/jgt-2012-0040. MR3031872

[J] Keith S. Joseph, *Research Problems: Several Conjectures on Commutativity in Algebraic Structures*, Amer. Math. Monthly **84** (1977), no. 7, 550–551, DOI 10.2307/2320020. MR1538432

[LS1] Michael Larsen and Aner Shalev, *Fibers of word maps and some applications*, J. Algebra **354** (2012), 36–48, DOI 10.1016/j.jalgebra.2011.10.040. MR2879221

[LS2] Michael Larsen and Aner Shalev, *On the distribution of values of certain word maps*, Trans. Amer. Math. Soc. **368** (2016), no. 3, 1647–1661, DOI 10.1090/tran/6389. MR3449221

[LS3] Michael Larsen and Aner Shalev, *A probabilistic Tits alternative and probabilistic identities*, Algebra Number Theory **10** (2016), no. 6, 1359–1371, DOI 10.2140/ant.2016.10.1359. MR3544299

[LS4] M. Larsen and A. Shalev, Words, Hausdorff dimension and randomly free groups, arXiv math:1706.08226.

[LP] L. Lévai and L. Pyber, *Profinite groups with many commuting pairs or involutions*, Arch. Math. (Basel) **75** (2000), no. 1, 1–7, DOI 10.1007/s000130050466. MR1764885

[M1] Avinoam Mann, *Finite groups containing many involutions*, Proc. Amer. Math. Soc. **122** (1994), no. 2, 383–385, DOI 10.2307/2161027. MR1242094

[M2] A. Mann, *Groups satisfying identities with high probability*, Internat. J. Algebra Comput., to appear.

[NY] Rajat Kanti Nath and Manoj Kumar Yadav, *On the probability distribution associated to commutator word map in finite groups*, Internat. J. Algebra Comput. **25** (2015), no. 7, 1107–1124, DOI 10.1142/S0218196715500320. MR3432221

[Ne] B. H. Neumann, *Identical relations in groups. I*, Math. Ann. **114** (1937), no. 1, 506–525, DOI 10.1007/BF01594191. MR1513153

[N] Peter M. Neumann, *Two combinatorial problems in group theory*, Bull. London Math. Soc. **21** (1989), no. 5, 456–458, DOI 10.1112/blms/21.5.456. MR1005821

[S] Aner Shalev, *Some results and problems in the theory of word maps*, Erdös centennial, Bolyai Soc. Math. Stud., vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 611–649, DOI 10.1007/978-3-642-39286-3_22. MR3203613

Einstein Institute of Mathematics, Hebrew University, Givat Ram, Jerusalem 91904, Israel

E-mail address: shalev@math.huji.ac.il