ANNALS OF MATHEMATICS

Word maps, conjugacy classes, and a noncommutative Waring-type theorem

By Aner Shalev

SECOND SERIES, VOL. 170, NO. 3

November, 2009

ANMAAH

Annals of Mathematics, **170** (2009), 1383–1416

Word maps, conjugacy classes, and a noncommutative Waring-type theorem

By ANER SHALEV

Abstract

Let $w = w(x_1, \ldots, x_d) \neq 1$ be a nontrivial group word. We show that if G is a sufficiently large finite simple group, then every element $g \in G$ can be expressed as a product of three values of w in G. This improves many known results for powers, commutators, as well as a theorem on general words obtained in [19]. The proof relies on probabilistic ideas, algebraic geometry, and character theory. Our methods, which apply the 'zeta function' $\zeta_G(s) = \sum_{\chi \in \mathrm{Irr}\, G} \chi(1)^{-s}$, give rise to various additional results of independent interest, including applications to conjectures of Ore and Thompson.

Partially supported by ISF and BSF grants.

1. **Main result**

A classical result in Number Theory, which goes back to Lagrange, states that every positive integer is a sum of four squares. Results for some larger powers were obtained, culminating in Hilbert's celebrated solution to Waring's Problem, showing that every positive integer is a sum of $f(k)$ kth powers, where f is a suitable function (see, for instance, [26]).

Are there analogs of this phenomenon for interesting nonabelian groups, such as symmetric groups, or Chevalley groups? We are interested in situations where every group element can be expressed as a short product of elements in the image of a given word map.

To make this precise, let $w = w(x_1, \ldots, x_d)$ be a nontrivial group word, namely a nonidentity element of the free group F_d on $x_1, \ldots, x_d$. Then we may write $w = x_{i_1}^{n_1} x_{i_2}^{n_2} \cdots x_{i_k}^{n_k}$ where $i_j \in \{1, \ldots, d\}$, n_j are integers, and we may assume further that w is reduced. Let G be a group. For $g_1, \ldots, g_d \in G$ we write

$$w(g_1, \ldots, g_d) = g_{i_1}^{n_1} g_{i_2}^{n_2} \cdots g_{i_k}^{n_k} \in G.$$

Let

$$w(G) = \{w(g_1, \ldots, g_d) : g_1, \ldots, g_d \in G\}$$

be the set of values of w in G. For subsets $A, B \subseteq G$ let $AB = \{ab \mid a \in A, b \in B\}$ and $A^k = \{a_1 \cdots a_k \mid a_i \in G\}$.

Fix a nontrivial group word w and let G be a finite simple group. If G is large enough then it follows from Jones [15] that $w(G) \neq \{1\}$ (namely w is not an identity in G). Can we then find a constant c (which may depend on w but not on G) such that $w(G)^c = G$? This is equivalent to the verbal subgroup of the Cartesian product of all finite simple groups generated by values of w being a closed subgroup.

Various instances of this problem were considered in the past decade or two. For $w(x_1, x_2) = [x_1, x_2] = x_1^{-1} x_2^{-1} x_1 x_2$, the commutator word, it was shown by Wilson [34] in 1994, using methods of mathematical logic, that indeed $w(G)^c = G$ for some absolute constant c. In 1996–7 Martinez and Zelmanov [25], and independently Saxl and Wilson [31], solved the problem for the power word $w = x_1^k$. It follows from their result that every element of a large enough finite simple group is a product of $f(k)$ kth powers.

Arbitrary words are dealt with by Liebeck and myself in [19]. Indeed we show in **??** there that for every word w there is a positive integer $c = c(w)$ such that if $w(G) \neq \{1\}$ then $w(G)^c = G$. The purpose of this paper is to obtain a much stronger result. We show that, for large G, the constant c above does not depend on w, and is in fact a surprisingly small number.

THEOREM 1.1. *Let $w \neq 1$ be a group word. Then there exists a positive integer $N = N(w)$ such that for every finite simple group G with $|G| \geq N(w)$ we have $w(G)^3 = G$.*

In fact for finite simple groups G of Lie type we prove even more: if w_1, w_2, w_3 are any three nontrivial group words, then $w_1(G)w_2(G)w_3(G) = G$ provided $|G|$ is large enough (see Theorem 6.6 below). These results are new even for the power words, which have been studied extensively in connection to Burnside type problems.

Commutator words have also been studied extensively in these contexts, also for profinite groups. Around 30 years ago Serre showed that in a finitely generated pro-p group G, every element of the commutator subgroup G' is a bounded product of commutators. This has been extended by many authors, culminating in Segal's proof of a similar result for finitely generated prosolvable groups [32], and in the proof of Nikolov and Segal [28], [29] for finitely generated profinite groups in general. See also Nikolov [27] and the references therein for the concept of commutator width and related positive and negative results.

We note that Theorem 1.1 does not hold for finite groups G in general, even not in the sense of $w(G)^{f(w,n)}$ coinciding with the verbal subgroup generated by w, where f depends on the word w and the minimal number of generators n of G. However, this latter statement does hold for certain (so called locally finite) words w, as established by Nikolov and Segal in [28] and [29]. This fact then enables them to show that any finite index subgroup of a finitely generated profinite group is open. It is still unknown whether every product of kth powers in a n-generated finite group can be expressed as a product of $f(k, n)$ kth powers.

Let us now describe the strategy of the proof of Theorem 1.1. The main tools involved in the proof are algebraic geometry, character theory, and probabilistic ingredients.

In the first stage of the proof we rely on a result of Borel [2] that a word map is a dominant map at the level of simple algebraic groups, and consequences proved by Larsen in [17] showing that word maps have large image in finite simple groups. Combining these results with some extra-arguments we are able to show that if w is a nontrivial group word, and G is a large finite simple group, then $w(G)$ contains a 'large' conjugacy class C_w of G (for example, a conjugacy class of a regular semisimple element in some cases).

In the next stage of the proof we use a probabilistic approach. Given a group word $w \neq 1$ we consider the large class C_w found inside $w(G)$, and study the random variable $y = y_1 y_2 y_3$ where $y_i \in C_w$ are randomly chosen (with uniform distribution).

At this point character theory comes into play. To understand its relevance here, let G be a finite group, $g \in G$, and let $C_i = x_i^G$ $(i = 1, \ldots, k)$ be conjugacy

classes. Let $P_{C_1,\dots,C_k}(g)$ denote the probability that $y_1 \cdots y_k = g$ where $y_i \in C_i$ are randomly chosen. It follows from a classical result (see e.g. [33, §7.2], or [1, 10.1, p. 43]) that

$$P_{C_1,\dots,C_k}(g) = |G|^{-1} \sum_{\chi \in \mathrm{Irr}\, G} \frac{\chi(x_1)\cdots\chi(x_k)\chi(g^{-1})}{\chi(1)^{k-1}}. \tag{1}$$

Unfortunately, in most cases we do not have the full character table of G at our disposal, and it is impossible to compute the right-hand side of (1). However, in our case, where $k = 3$ and $C_i = C_w$, the large conjugacy class found inside $w(G)$, we are able to show that the main term in (1) comes from the trivial character $\chi = 1$, and the contribution of the other terms is marginal.

The main tools in showing this are general character theory (see [14]), the Deligne-Lusztig theory of characters of Chevalley groups [23], and the recent work [22] on the 'zeta function'

$$\zeta_G(s) = \sum_{\chi \in \mathrm{Irr}\, G} \chi(1)^{-s}$$

encoding character degrees. More specifically, we use the fact (established in [22]) that there is an absolute constant c such that every finite simple group G has at most cn irreducible representations of degree n, and closely related results on $\zeta_G(s)$.

At this stage it follows that the random variable y defined above is almost uniformly distributed in the l_∞-norm. In particular, for large enough G, y attains all values $g \in G$, and so

$$w(G)^3 \supseteq C_w^3 = G,$$

completing the proof of our main result.

While we focussed above on the proof for groups of Lie type, which is the more challenging task, it is intriguing that even the proof of Theorem 1.1 for alternating groups A_n is not elementary, in that it relies implicitly on algebraic geometry via [17], as well as on [1] and on the Erdős-Turán theory of random permutations. It would be interesting to find out whether a purely combinatorial proof of Theorem 1.1 for alternating groups exists.

2. **Intermediate results**

In the course of the proof of Theorem 1.1 we establish a variety of results of independent interest, which we state in this section. These are related to short products of conjugacy classes in finite Chevalley groups, and the probability distributions they induce.

We use the notation of the previous section. In particular, for conjugacy classes $C_1, C_2, C_3 \subset G$, and for $g \in G$, $P_{C_1,C_2,C_3}(g)$ is the probability that $g = y_1y_2y_3$ where $y_i \in C_i$ is chosen at random (with respect to the uniform distribution on C_i).

Recall that an element x of a finite group G of Lie type is called regular if its centralizer in the corresponding algebraic group $\overline{G}$ has minimal dimension, namely $\mathrm{rank}(\overline{G})$. We say that x is semisimple if its order is not divisible by p, where p is the defining characteristic of G.

By $o(1)$ we mean a function of $|G|$ alone, which tends to zero as $|G| \to \infty$.

THEOREM 2.1. *Let G be a finite simple group of Lie type, and let C_1, C_2, C_3 be conjugacy classes of regular semisimple elements in G. Then*

$$P_{C_1,C_2,C_3}(g) = (1 + o(1))|G|^{-1} \text{ for all } g \in G.$$

This theorem shows that P_{C_1,C_2,C_3} is almost uniform in the l_∞-norm. This seems to be the first result of this kind for groups of Lie type. For alternating groups A_n, and certain so called almost homogeneous classes $C_1, C_2, C_3 \subseteq A_n$, it is shown in Theorem 1.14 of [20], that P_{C_1,C_2,C_3} is almost uniform.

Theorem 2.1 is best possible in the sense that it does not hold for a product of two classes C_1, C_2. Indeed, to begin with, 1 may not lie in C_1C_2. But even when $1 \in C_1C_2$ (so $C_2 = C_1^{-1}$) it is clear that, fixing $x_1 \in C_1$, we have $P_{C_1,C_2}(1) = |C_1|^{-1} = |G|^{-1}|C_G(x_1)|$, which is much larger than $|G|^{-1}$.

Theorem 2.1 has the following immediate consequence.

COROLLARY 2.2. *There exists an absolute constant c such that, if G is a finite simple group of Lie type, and $C_1, C_2, C_3 \subset G$ are conjugacy classes of regular semisimple elements of G, then $C_1C_2C_3 = G$ provided $|G| \geq c$.*

As already noted, this need not hold for two classes C_1, C_2.

It has been shown by Malle, Saxl and Weigel (see Theorem 2.11 of [24]) that if G is a finite simple classical group which is not an orthogonal group in even dimension, then G has a conjugacy class C with $C^3 = G$.

As an application of Corollary 2.2 we obtain the following.

COROLLARY 2.3. *Every large enough finite simple group G has a conjugacy class C such that $C^3 = G$. Moreover, if G is of Lie type, then any conjugacy class C of a regular semisimple element will do.*

For some particular classes of regular semisimple elements, namely those in split tori, it has been shown by Ellers and Gordeev that $C_1C_2 \supseteq G \setminus \{1\}$ (see for example Theorem 1 in [5]). However, the behavior of C_1C_2 for classes C_i of regular semisimple elements outside split tori is much less understood (see Gow [11] for interesting partial information).

Let $G_r(q)$ denote a finite simple group of Lie type of rank r over the field with q elements. Here r is defined to be the rank of the ambient simple algebraic

group, unless we deal with Lie types ${}^2B_2, {}^2G_2$ or ${}^2F_4$, in which case $r = 1, 1, 2$ respectively.

Note that if $x \in G = G_r(q)$ is regular then $|C_G(x)| \sim q^r$ (up to multiplicative constants). In fact our methods enable us to extend Theorem 2.1 for a much larger family of conjugacy classes, provided G has large rank. This extension is essential for a whole range of applications (including our main result).

THEOREM 2.4. *For every $\varepsilon > 0$ there is a number $r_1(\varepsilon)$ such that if $G = G_r(q)$ where $r \geq r_1(\varepsilon)$, $C_1, C_2, C_3 \subset G$ are conjugacy classes of elements $x_1, x_2, x_3 \in G$ satisfying*

$$|C_G(x_1)| \cdot |C_G(x_2)| \cdot |C_G(x_3)| \leq q^{(4-\varepsilon)r},$$

Then

$$P_{C_1,C_2,C_3}(g) = (1 + o(1))|G|^{-1} \text{ for all } g \in G.$$

Remark. In fact our proof shows that the conclusion of Theorem 2.4 holds under the somewhat weaker assumption that

$$|C_G(x_1)|\,|C_G(x_2)|\,|C_G(x_3)|/q^{4r-6} \to 0,$$

and that r is larger than some absolute constant.

COROLLARY 2.5. *For every $\varepsilon > 0$ there is a number $r_2(\varepsilon)$ such that if $G = G_r(q)$ where $r \geq r_2(\varepsilon)$, $C_1, C_2, C_3 \subset G$ are conjugacy classes of elements $x_1, x_2, x_3 \in G$ satisfying*

$$|C_G(x_1)| \cdot |C_G(x_2)| \cdot |C_G(x_3)| \leq q^{(4-\varepsilon)r},$$

Then

$$C_1C_2C_3 = G.$$

In particular, if C is a conjugacy class of an element $x \in G$ satisfying $|C_G(x)| \leq q^{(4/3-\varepsilon)r}$, then $C^3 = G$, provided $r \geq r_3(\varepsilon)$.

Remark. The conclusion of Corollary 2.5 holds whenever $r \geq c$ and

$$|C_G(x_1)|\,|C_G(x_2)|\,|C_G(x_3)| \leq \delta \cdot q^{4r-6},$$

where $c, \delta > 0$ are certain absolute constants.

Our proof of Theorems 2.1 and 2.4 relies heavily on character theory.

Let us now consider conjugacy classes of randomly chosen elements. It is shown in Theorem 1.12 of [19] that there exists an absolute constant c, such that if G is a finite simple group, and $x \in G$ is chosen at random, then we have $(x^G)^c = G$ with probability tending to 1 as $|G| \to \infty$. Here we are able to improve this, showing that $c = 3$ will do. The exponent 3 is best possible here.

THEOREM 2.6. *Let G be a finite simple group, and let $x \in G$ be chosen at random. Then the probability that $(x^G)^3 = G$ tends to 1 as $|G| \to \infty$.*

Note that in [3] it is shown that a random conjugacy class C of the alternating group A_n satisfies $C^4 = A_n$ with probability tending to 1.

For more background on products of conjugacy classes in finite simple groups, see [1], [19], and the references therein.

Results 2.3, 2.5 and 2.6 above all show that the cube of large (or random) conjugacy classes in G is the whole of G. Our next result deals with smaller conjugacy classes, and shows that their cubes are also large in some sense.

THEOREM 2.7. *For every $\delta > 0$ there is $\varepsilon > 0$ such that, if G is a finite simple group, and $C \subset G$ is a conjugacy class of size at most $|G|^{1-\delta}$, then*

$$|C^3| \geq |C|^{1+\varepsilon}.$$

In fact for groups of Lie type of bounded rank we obtain an expansion result for C^2 (see Proposition 10.4 below).

Finally, we apply our methods to study longstanding conjectures by Ore and by Thompson. Ore conjectured that every element of a finite simple group is a commutator [30]. Thompson conjectured that every finite simple group G has a conjugacy class C such that $C^2 = G$.

Both conjectures have been proved for alternating groups, and for groups of Lie type over fields with more than 8 elements [5]. A full proof of these conjectures still seems out of reach.[1] Of course, Theorem 1.1 shows that in every large finite simple group all elements are products of three commutators. Using properties of the function $\zeta_G(s)$ established in [22] we improve this as follows.

THEOREM 2.8. *Let G be a finite simple group, and let $x_1, x_2, x_3, x_4 \in G$ be randomly chosen. Then, for every $g \in G$,*

$$\mathrm{Prob}([x_1, x_2][x_3, x_4] = g) = (1 + o(1))|G|^{-1}.$$

In particular, there exists an absolute constant c such that if G is a finite simple group of order at least c then every element of G is a product of two commutators.

We also show that almost all elements of a finite simple group G satisfy the conjectures of Ore and Thompson in the following sense:

THEOREM 2.9. *Let G be a finite simple group.*

(i) *Let* $\mathrm{Com}(G)$ *denote the set of all commutators in G. Then*

$$|\mathrm{Com}(G)|/|G| \to 1 \text{ as } |G| \to \infty.$$

Moreover, there is an absolute constant c such that, if $G = G_r(q)$, then

$$|\mathrm{Com}(G)| \geq (1 - cq^{-2r})|G|.$$

[1] *Footnote added in proofs*: The Ore conjecture has recently been proved by Liebeck, O'Brien, Tiep and myself.

(ii) *There exists a conjugacy class C_G in G such that*

$$|C_G^2|/|G| \to 1 \text{ as } |G| \to \infty.$$

Moreover, there is an absolute constant c such that, if $G = G_r(q)$, then

$$|C_G^2| \geq (1 - cq^{-r})|G|.$$

Our methods also yield results for longer commutators. For $d > 2$ the d-fold commutator $[x_1, \ldots, x_d]$ is defined by induction as $[[x_1, \ldots, x_{d-1}], x_d]$. We study the property of being a d-fold commutator *for all d simultaneously*.

THEOREM 2.10. *Let G be a finite simple group.*

(i) *The probability that a randomly chosen element $g \in G$ is a d-fold commutator for all $d \geq 2$ tends to 1 as $|G| \to \infty$.*

(ii) *There exists an absolute constant c such that if $|G| \geq c$ then every element of G is a product of two elements of G which are d-fold commutators for all $d \geq 2$.*

This paper is organized as follows. In Section 3 we prove the main results for alternating groups. Section 4 deals with characters of finite simple groups of Lie type, establishing a technical result (Theorem 4.1) which is a main tool in this paper. In Section 5 we apply this result and prove results 2.1-2.6 for groups of Lie type. Corollaries 2.3 and 2.5 are then applied in the proof of Theorem 1.1, which is carried out in Section 6. In Section 7 we establish 'expansion' for powers of arbitrary conjugacy classes, and prove Theorem 2.7. In Section 8 we examine probability distributions induced by group words w and their products using a noncommutative Fourier transform, and deduce that they are almost uniform in some cases. Section 9 is devoted to the proof of Theorems 2.8-2.10. Finally, in Section 10, some open problems and examples are presented.

I am grateful to Roman Bezrukavnikov and Michael Larsen for interesting discussions, and to the referee for valuable comments.

Notation. For a group G and $x \in G$ we let x^G denote the conjugacy class of x in G. Let $k(G)$ denote the number of conjugacy classes of G. A normal subset of G is a subset closed under conjugation by the elements of G (namely a union of conjugacy classes). By $G_r(q)$ we denote a finite simple group of Lie type, of rank r over the field with q elements. $\operatorname{Irr} G$ stands for the set of irreducible complex characters of G (so $|\operatorname{Irr} G| = k(G)$). For $g \in G$ we set $R(g) = \max_{1 \neq \chi \in \operatorname{Irr} G} \chi(g)/\chi(1)$.

If $C_1, \ldots, C_k$ are conjugacy classes in G, we denote by $N_{C_1,\ldots,C_k}(g)$ the number of solutions to the equation $y_1 \cdots y_k = g$ where $y_i \in C_i$. We let $P_{C_1,\ldots,C_k}$ denote the corresponding probability distribution on G, so that $P_{C_1,\ldots,C_k}(g) = N_{C_1,\ldots,C_k}(g)/(|C_1| \cdots |C_k|)$ is the probability that $y_1 \cdots y_k = g$ when $y_i \in C$ are randomly chosen.

Given a group word $w = w(x_1, \cdots, x_d)$ and a finite group G we let $w(G)$ denote the image of the word map from G^d to G induced by w. This is clearly a normal subset. Moreover, $w(G)$ is characteristic (invariant under automorphisms of G). We denote by $N_{w,G}(g)$ the number of solutions to the equation $w(x_1, \ldots, x_d) = g$ in G. We let $P_{w,G}$ be the corresponding probability distribution on G, so that, for $g \in G$, $P_{w,G}(g) = N_{w,G}(g)/|G|^d$.

We say that words u, v are disjoint if their subsets of variables are disjoint. For a word w and $k \geq 1$ we let w^k denote a product of k disjoint copies of w (each with its own set of variables).

Let U_G denote the uniform distribution on G. Given distributions P, Q on G, we define their l_∞-distance by

$$\|P - Q\|_\infty = |G| \cdot \max_{g \in G} |P(g) - Q(g)|.$$

Finally, following [22] we set

$$\zeta_G(s) = \sum_{\chi \in \mathrm{Irr}\, G} \chi(1)^{-s},$$

where G is a finite group and $s > 0$. This finite analogue of Witten's 'zeta function' encoding representation degrees of compact Lie groups [35] plays a major role in this paper.

3. **Alternating groups**

In this section we prove Theorems 1.1 and 2.6 for alternating groups A_n. The proof of Theorem 2.7 for A_n is included in Section 7.

We denote the number of cycles (including 1-cycles) in a permutation $\sigma \in S_n$ by $\mathrm{cyc}(\sigma)$.

LEMMA 3.1. *Let C be a conjugacy class in S_n and let $\sigma \in C$. Suppose $\mathrm{cyc}(\sigma) < n/2$. Then $C^3 = \sigma A_n$.*

Proof. By a result of Dvir ([1, Ch. 3, p. 219, Th. 10.2)] if $\mathrm{cyc}(\sigma) \leq (n+1)/2$, and σ is not a fixed-point-free involution, then $C^3 = \sigma A_n$. Our assumption on $\mathrm{cyc}(\sigma)$ implies the two conditions above, and so the conclusion follows. □

LEMMA 3.2. *Let $w \neq 1$ be a group word, and let $\varepsilon > 0$. There is a function f such that if $n > f(w, \varepsilon)$ then $w(A_n)$ contains an S_n-conjugacy class C of an element σ with $\mathrm{cyc}(\sigma) < n^\varepsilon$.*

Proof. We apply Larsen's paper [17]. It follows from the proof of Proposition 8 there, that for every $\delta > 0$ and large enough n, the set of values $w(A_n)$ of w contains an element with at most $n^\delta \log n$ cycles. Taking $\delta = \varepsilon/2$ and n large we obtain $\sigma \in w(A_n)$ with $\mathrm{cyc}(\sigma) \leq n^\varepsilon$.

Since the set $w(A_n)$ is characteristic, it contains the S_n-conjugacy class of σ. The result follows. □

We conclude that, for large n (say $n > N(w)$), the set $w(A_n)$ contains an S_n-class C of an element with less than $n/2$ cycles, and hence $C^3 = A_n$ and so $w(A_n)^3 = A_n$. This proves Theorem 1.1 for alternating groups.

We conclude this section by proving Theorem 2.6 for alternating groups.

LEMMA 3.3. *Let $\sigma \in A_n$ be chosen at random.*

(i) *The probability that* $\mathrm{cyc}(\sigma) < 2\log n$ *tends to* 1 *as* $n \to \infty$.

(ii) *The probability that* $\sigma^{A_n} = \sigma^{S_n}$ *tends to* 1 *as* $n \to \infty$.

Proof. The proof relies on the Erdős-Turán Theory of random permutations. By [6, (2.2), p. 176], for any $\varepsilon > 0$, the probability that $\sigma \in S_n$ satisfies

$$(1-\varepsilon)\log n < \mathrm{cyc}(\sigma) < (1+\varepsilon)\log n$$

tends to 1 as $n \to \infty$. Of course a similar result follows for A_n. This proves part (i).

For part (ii) we use Theorem VI of [7], showing that the probability that $\sigma \in S_n$ has no cycles of length $a_1, \ldots, a_s$ is at most $(a_1^{-1} + \ldots + a_s^{-1})^{-1}$. Letting a_i be all the even numbers up to n, we see that the probability that $\sigma \in S_n$ has only cycles of odd length tends to 0 as $n \to \infty$. A similar result follows immediately for A_n. We deduce that the probability that $\sigma^{A_n} = \sigma^{S_n}$ tends to 1 as $n \to \infty$. □

It follows from Lemma 3.3 that for almost all $\sigma \in A_n$ we have $\mathrm{cyc}(\sigma) < n/2$ and $\sigma^{A_n} = \sigma^{S_n}$. This fact, combined with Lemma 3.1, show that, for random $\sigma \in A_n$, the probability that $(\sigma^{A_n})^3 = A_n$ tends to 1 as $n \to \infty$.

Theorem 2.6 is proved for A_n.

4. **Character theoretic preparations**

Our proofs for groups of Lie type rely heavily on character theory. In this section we set the required machinery.

For a group G and elements $x_1, \ldots, x_k \in G$, define

$$E(G, x_1, \ldots, x_k) = \sum_{1 \neq \chi \in \mathrm{Irr}\, G} \frac{|\prod_{i=1}^k \chi(x_i)|}{\chi(1)}.$$

The main result of this section is the following.

THEOREM 4.1. *Let G be a finite simple group of Lie type.*

(i) *Suppose $x_1, x_2, x_3 \in G$ are regular semisimple elements and $G \neq \mathrm{PSL}_2(q)$. Then $E(G, x_1, x_2, x_3) \to 0$ as $|G| \to \infty$.*

(ii) *For each $\varepsilon > 0$ there is $r_1(\varepsilon)$ such that, if $G = G_r(q)$, $r \geq r_1(\varepsilon)$, $k > 1$, $x_1, \ldots, x_k \in G$, and $\prod_{i=1}^k |C_G(x_i)| \leq q^{(4-\varepsilon)r}$, then $E(G, x_1, \ldots, x_k) \to 0$ as $|G| \to \infty$.*

Remarks. 1. We note that, under the assumptions of part (i), we also have $E(G, x_1, x_2) \to 0$ and $E(G, x_1) \to 0$ (the proof is similar and somewhat easier).

2. Part (ii) above is applicable for $k = 2, 3$ (since by [8] there is an absolute constant $c > 0$ such that $|C_G(x)| \geq cq^r / \log q$ for all $x \in G_r(q)$).

In this paper we mainly use it for $k = 3$, so we shall assume this in the proof below. The case $k = 2$ can be proved in a very similar manner.

3. A version of part (ii) for $k = 1$ is stated and proved at the end of this section (see Proposition 4.7 below).

4. It will be clear from the proof of Theorem 4.1 that that the conclusion of part (ii) holds whenever r exceeds some absolute constant and

$$|C_G(x_1)|\,|C_G(x_2)|\,|C_G(x_3)|/q^{4r-6} \to 0.$$

To prove Theorem 4.1 we need a few results. First we quote Theorem 1.1 of [22], which plays a major role in this paper.

THEOREM 4.2. *Let G be a finite simple group, and for a real number s let $\zeta_G(s) = \sum_{\chi \in \mathrm{Irr}\, G} \chi(1)^{-s}$.*

(i) *If $s > 1$ then $\zeta_G(s) \to 1$ as $|G| \to \infty$.*

(ii) *If $s > 2/3$ and $G \neq \mathrm{PSL}_2(q)$ then $\zeta_G(s) \to 1$ as $|G| \to \infty$.*

We also need the following.

LEMMA 4.3. *There are positive constants c_1, c_2 such that if $G = G_r(q)$ then*

(i) *$\chi(1) \geq c_1 q^r$ for all $1 \neq \chi \in \mathrm{Irr}\, G$, and*

(ii) *$k(G) \leq c_2 q^r$.*

Proof. Part (i) follows from work of Landazuri-Seitz [16] and (ii) from work of Fulman-Guralnick [8] (see also Liebeck-Pyber [18] for the case when r is bounded). □

LEMMA 4.4. *Let $G = G_r(q)$ be a finite simple group of Lie type, and let $x \in G$ be a regular semisimple element. Then there is a number $c = c(r)$, depending on r but not on q, such that*

$$|\chi(x)| \leq c$$

for all $\chi \in \mathrm{Irr}\, G$.

Proof. This follows from the Deligne-Lusztig theory, see [23], and formula 4.26.1 in particular. □

LEMMA 4.5. *Let G be a finite group, and $N > 0$. Then*

$$\sum_{\chi \in \mathrm{Irr}\, G, \chi(1) \geq N} \frac{|\chi(x_1)\chi(x_2)\chi(x_3)|}{\chi(1)} \leq \frac{(|C_G(x_1)|\,|C_G(x_2)|\,|C_G(x_3)|)^{1/2}}{N}.$$

Proof. By the generalized orthogonality relations (see Isaacs [14]) we have

$$\sum_{\chi \in \mathrm{Irr}\, G} |\chi(x_i)|^2 = |C_G(x_i)|.$$

In particular $|\chi(x_1)| \leq |C_G(x_1)|^{1/2}$ for all χ.

Clearly

$$\sum_{\chi \in \mathrm{Irr}\, G, \chi(1) \geq N} \frac{|\chi(x_1)\chi(x_2)\chi(x_3)|}{\chi(1)} \leq N^{-1} |C_G(x_1)|^{1/2} \sum_{\chi \in \mathrm{Irr}\, G} |\chi(x_2)\chi(x_3)|.$$

Now, by Cauchy-Schwarz inequality we have,

$$\begin{aligned} \sum_{\chi} |\chi(x_2)\chi(x_3)| &\leq \left(\sum_{\chi} |\chi(x_2)|^2 \right)^{1/2} \left(\sum_{\chi} |\chi(x_3)|^2 \right)^{1/2} \\ &\leq |C_G(x_2)|^{1/2} |C_G(x_3)|^{1/2}. \end{aligned}$$

The result now follows from the two inequalities above. □

LEMMA 4.6. *Let $G = G_r(q)$ be a finite simple classical group. Then* $\mathrm{Irr}\, G$ *has a subset W of so called Weil characters with the following properties*:

(i) $|W| \leq q + 1$.

(ii) *If $x \in G$ and $|C_G(x)| \leq q^m$ then $|\chi(x)| \leq q^{\sqrt{m+b}}$ where b is some absolute constant.*

(iii) *If $\chi \in \mathrm{Irr}\, G \setminus W$ and $r > 5$ then $\chi(1) \geq cq^{2r-3}$ where $c > 0$ is some absolute constant.*

Proof. This follows from the discussion in Section 6 of [21] (see Lemma 6.1 and Lemma 6.2 in particular). □

We note that the set W is sometimes empty, but the lemma still holds in these cases.

Proof of Theorem 4.1. We first prove part (ii) (with $k = 3$). So we assume $G = G_r(q)$ is classical and $x_1, x_2, x_3 \in G$ satisfy

$$\prod_{i=1}^{3} |C_G(x_i)| \leq q^{(4-\varepsilon)r}.$$

Let W be the set of Weil characters of G. Set

$$E_1(G, x_1, x_2, x_3) = \sum_{\chi \in W} \frac{|\chi(x_1)\chi(x_2)\chi(x_3)|}{\chi(1)},$$

and

$$E_2(G, x_1, x_2, x_3) = \sum_{1 \neq \chi \notin W} \frac{|\chi(x_1)\chi(x_2)\chi(x_3)|}{\chi(1)}.$$

Then $E(G, x_1, x_2, x_3) = E_1(G, x_1, x_2, x_3) + E_2(G, x_1, x_2, x_3)$.

By our assumptions $|C_G(x_i)| \leq q^{4r}$ (in fact better bounds can be easily deduced), and so Lemma 4.6(ii) yields $|\chi(x_i)| \leq q^{\sqrt{4r+b}}$ for $i = 1, 2, 3$. We also have $\chi(1) \geq c_1 q^r$ for all nontrivial $\chi \in \operatorname{Irr} G$ by Lemma 4.3(i). Therefore

$$E_1(G, x_1, x_2, x_3) \leq \sum_{\chi \in W} q^{3\sqrt{4r+b}}/\chi(1) \leq |W| q^{3\sqrt{4r+b}}/(c_1 q^r).$$

Since $|W| \leq q + 1$ we obtain

$$E_1(G, x_1, x_2, x_3) \leq c_2 q^{3\sqrt{4r+b}-(r-1)}.$$

It follows that for r sufficiently large (larger than some absolute constant) we have $E_1(G, x_1, x_2, x_3) \to 0$ as $|G| \to \infty$.

Now, if $1 \neq \chi \notin W$, and $r > 5$, Lemma 4.6(iii) yields $\chi(1) \geq N$, where $N = [cq^{2r-3}]$. By Lemma 4.5 we obtain

$$E_2(G, x_1, x_2, x_3) \leq N^{-1}(|C_G(x_1)|\,|C_G(x_2)|\,|C_G(x_3)|)^{1/2}.$$

Using our assumptions on $|C_G(x_i)|$ this yields

$$E_2(G, x_1, x_2, x_3) \leq c_3 q^{-(2r-3)}(q^{(4-\varepsilon)r})^{1/2} = c_3 q^{-(2r-3)+(2-\varepsilon/2)r}.$$

Therefore

$$E_2(G, x_1, x_2, x_3) \leq c_3 q^{-\varepsilon r/2+3}.$$

We conclude that, if $r > 6/\varepsilon$ and q or r tend to infinity, then $E_2(G, x_1, x_2, x_3)$ tends to zero. The proof of part (ii) is complete. Remark 4 following Theorem 4.1 can also be deduced.

Indeed, $E_1(G, x_1, x_2, x_3) \to 0$ provided $r \geq c$, and $E_2(G, x_1, x_2, x_3) \to 0$ provided $q^{-(4r-6)} \prod_{i=1}^{3} |C_G(x_i)| \to 0$.

It remains to prove part (i) of the theorem. Since $|C_G(x_i)| = O(q^r)$ for regular elements x_i, part (i) follows from part (ii) for large rank r. Hence, to prove part (i), we may assume that r is bounded. But then Lemma 4.4 shows that $|\chi(x_i)| \leq c$ for some absolute constant c and for all χ, and so

$$E(G, x_1, x_2, x_3) \leq c^3 \sum_{1 \neq \chi \in \operatorname{Irr} G} \chi(1)^{-1} = c^3(\zeta_G(1) - 1).$$

By Theorem 4.2 above we have $\zeta_G(1) \to 1$ as $|G| \to \infty$, provided $G \neq \mathrm{PSL}_2(q)$. Thus, under the same assumption, $E(G, x_1, x_2, x_3)$ tends to 0. Theorem 4.1 is proved. □

Remark. Note that $G = \mathrm{PSL}_2(q)$ is a genuine exception to (i). Here $E(G, x_1, x_2, x_3)$ may be bounded away from zero (e.g. when $x_1 = x_2 = x_3$). However, $E(G, x_1, x_2, x_3)$ is also bounded above in this case (independently of q). These facts can be easily verified using the well known character table of $\mathrm{PSL}_2(q)$.

For later applications we also need the following variation on Theorem 4.1(ii).

PROPOSITION 4.7. *Let $G = G_r(q)$. For each $\varepsilon > 0$ there is $r_1(\varepsilon)$ such that, if $r \geq r_1(\varepsilon)$, and $x \in G$ satisfies $|C_G(x)| \leq q^{(3-\varepsilon)r}$, then $E(G,x) \to 0$ as $|G| \to \infty$.*

Proof. Recall that

$$E(G,x) = \sum_{1 \neq \chi \in \mathrm{Irr}\, G} \frac{|\chi(x)|}{\chi(1)}.$$

The proof follows that of 4.1(ii) with some adjustments. First, by Cauchy-Schwarz inequality we have

$$\sum_{\chi \in \mathrm{Irr}\, G} |\chi(x)| \leq (\sum_{\chi} |\chi(x)|^2)^{1/2} |\mathrm{Irr}\, G|^{1/2} = |C_G(x)|^{1/2} k(G)^{1/2}.$$

From this it follows that, given a positive integer N, we have

$$\sum_{\chi \in \mathrm{Irr}\, G, \chi(1) \geq N} \frac{|\chi(x)|}{\chi(1)} \leq \frac{|C_G(x)|^{1/2} k(G)^{1/2}}{N}. \tag{2}$$

Suppose $G = G_r(q)$. We adopt the notation of the proof of 4.1. In particular we set $N = [cq^{2r-3}]$ and write $E(G,x) = E_1(G,x) + E_2(G,x)$, where

$$E_1(G,x) = \sum_{\chi \in W} \frac{|\chi(x)|}{\chi(1)},$$

and

$$E_2(G,x) = \sum_{1 \neq \chi \notin W} \frac{|\chi(x)|}{\chi(1)}.$$

Assuming $|C_G(x)| \leq q^{(3-\varepsilon)r}$ we then have

$$|E_1(G,x)| \leq |W| q^{\sqrt{3r+b}} / (c_1 q^r) \leq c_2 q^{\sqrt{3r+b}-(r-1)},$$

which tends to zero as $|G| \to \infty$, provided r is larger than some absolute constant.

Now, for $1 \neq \chi \notin W$ we have $\chi(1) \geq N$, so using (2) we obtain

$$E_2(G,x) \leq \frac{|C_G(x)|^{1/2} k(G)^{1/2}}{N} \leq \frac{q^{(3-\varepsilon)r/2} (c_2 q^r)^{1/2}}{cq^{2r-3}} \leq c_3 q^{-\varepsilon r/2+3}.$$

Thus $E_2(G,x) \to 0$ if $r > 6/\varepsilon$. The result follows. □

Remarks. 1. It follows from the proof above that

$$E(G,x) = O(q^{\sqrt{3r+b}-(r-1)} + q^{3-3r/2} |C_G(x)|^{1/2}).$$

2. This implies that, if $r \geq c$, and $|C_G(x)|/q^{3r-6} \to 0$, then $E(G,x) \to 0$. We shall also use this version of Proposition 4.7.

5. **Almost uniform distributions, I**

In this section we show how the quantity $E(G, x_1, x_2, x_3)$ estimated in the previous section can be used to study the distributions P_{C_1,C_2,C_3} and their l_∞-distance from the uniform distribution on G. We will show that, for large classes C_i, P_{C_1,C_2,C_3} is almost uniform, and prove Theorems 2.1, 2.4, 2.6, and related corollaries.

Let G be a finite group, and let $x_1, x_2, x_3 \in G$. Set $C_i = x_i^G$. Recall that U_G denotes the uniform distribution on G.

LEMMA 5.1. *With the above notation we have*

$$\|P_{C_1,C_2,C_3} - U_G\|_\infty \le E(G, x_1, x_2, x_3).$$

Proof. Let $P = P_{C_1,C_2,C_3}$. Recall that by (1) we have for $g \in G$

$$P(g) = |G|^{-1} \sum_{\chi \in \mathrm{Irr}\, G} \frac{\chi(x_1)\chi(x_2)\chi(x_3)\chi(g^{-1})}{\chi(1)^2}.$$

Set

$$\Delta(g) = \sum_{1 \ne \chi \in \mathrm{Irr}\, G} \frac{\chi(x_1)\chi(x_2)\chi(x_3)\chi(g^{-1})}{\chi(1)^2}.$$

Then since $|\chi(g^{-1})/\chi(1)| \le 1$ we obtain

$$|\Delta(g)| \le \sum_{1 \ne \chi \in \mathrm{Irr}\, G} \frac{|\chi(x_1)\chi(x_2)\chi(x_3)|}{\chi(1)} = E(G, x_1, x_2, x_3).$$

We also have

$$P(g) = |G|^{-1}(1 + \Delta(g)).$$

Thus

$$|P(g) - |G|^{-1}| \le |G|^{-1}|\Delta(g)| \le |G|^{-1} E(G, x_1, x_2, x_3).$$

Since $\|P - U_G\|_\infty = |G| \max_{g \in G} |P(g) - |G|^{-1}|$ the result follows. □

The next result establishes an almost uniform distribution for the random variable $y_1 y_2 y_3$ where $y_i \in C_i$, for large classes C_i.

THEOREM 5.2. *Let $G = G_r(q)$.*

(i) *If C_1, C_2, C_3 are classes of regular semisimple elements in G, then*

$$\|P_{C_1,C_2,C_3} - U_G\|_\infty \to 0 \text{ as } |G| \to \infty.$$

(ii) *The same holds when $C_i = x_i^G$ $(i = 1, 2, 3)$, $\prod_{i=1}^3 |C_G(x_i)| \le q^{(4-\varepsilon)r}$, and $r \ge r_1(\varepsilon)$.*

Proof. Part (i) follows from Theorem 4.1(i) and Lemma 5.1, provided $G \neq \mathrm{PSL}_2(q)$. Suppose $G = \mathrm{PSL}_2(q)$. For $g \in G$ let $\Delta(g)$ be as in the proof above. Then

$$\|P_{C_1,C_2,C_3} - U_G\|_\infty = \max_{g \in G} \Delta(g),$$

so it suffices to show that $\Delta(g) \to 0$ as $|G| \to \infty$.

Suppose first $g \neq 1$. Recall that $R(g)$ is the maximal character ratio of g over the characters $\chi \neq 1$. Thus $|\chi(g)/\chi(1)| \leq R(g)$ for all $\chi \neq 1$, and

$$|\Delta(g)| \leq E(G, x_1, x_2, x_3) R(g).$$

By the character table of G we have $R(g) \leq 2q^{-1/2}$. Since $E(x_1, x_2, x_3) \leq c$ for some absolute constant c (see remark after the proof of Theorem 4.1), it follows that $\Delta(g) \to 0$ as $q \to \infty$, provided $g \neq 1$.

It remains to show that this also holds for $g = 1$. In this case

$$\Delta(1) = \sum_{\chi \neq 1} \frac{\chi(x_1)\chi(x_2)\chi(x_3)}{\chi(1)},$$

and the fact that $\Delta(1) \to 0$ as $q \to \infty$ can be easily verified using the character table of G. This completes the proof of part (i).

To prove part (ii) we combine Lemma 5.1 with part (ii) of Theorem 4.1. □

Proof of results 2.1–2.5. Theorem 2.1 and 2.4 are parts (i) and (ii) of Theorem 5.2 respectively. The remark after Theorem 2.4 follows from Remark 4 after Theorem 4.1.

Corollaries 2.2, 2.3 and 2.5 also follow. Indeed, we choose G large enough so that $\|P_{C_1,C_2,C_3} - U_G\|_\infty < 1$. Then $P_{C_1,C_2,C_3}(g) > 0$ for all $g \in G$, so $C_1C_2C_3 = G$. The remark following Corollary 2.5 follows from Lemma 5.1 and previous remarks. □

To prove Theorem 2.6 we need the following.

LEMMA 5.3. *Let G be a finite group and let $x \in G$ be chosen at random. Then, for each number $N > 0$, the probability that $|C_G(x)| < N$ is at least $1 - k(G)/N$.*

Proof. Let $S = \{x \in G : |C_G(x)| \geq N\}$. Then for $x \in S$ we have $|x^G| \leq |G|/N$. Now, S is a normal subset, which splits into at most $k(G)$ conjugacy classes. Hence $|S| \leq k(G)|G|/N$. We see that $|C_G(x)| \geq N$ holds with probability at most $k(G)/N$. The result follows. □

COROLLARY 5.4. *Let $G = G_r(q)$, and let $\alpha > 0$. Then the probability that $|C_G(x)| < q^{(1+\alpha)r}$ is at least $1 - cq^{-\alpha r}$ for some absolute constant $c > 0$. In particular, this probability tends to 1 as $|G| \to \infty$.*

Proof. This follows from the lemma above, using the inequality $k(G) \leq c_2 q^r$ (see Lemma 4.3 above). □

We also rely on the following result of Guralnick and Lübeck [12].

LEMMA 5.5. *There exists an absolute constant a such that the number of regular semisimple elements in $G = G_r(q)$ is at least $(1 - aq^{-1})|G|$.*

Hence for $q \to \infty$ most elements of G are regular semisimple (but this is not the case when q is bounded).

Proof of Theorem 2.6. It remains to prove the theorem for $G = G_r(q)$ of Lie type. Let $|G| \to \infty$ and let $x \in G$ be chosen at random.

Suppose first that $r \geq r_2(1/4)$, where r_2 is the function in the first assertion of Corollary 2.5. Then, if $|C_G(x)| < q^{5r/4}$ we have $|C_G(x)|^3 < q^{(4-1/4)r}$, and so Corollary 2.5 shows that $(x^G)^3 = G$.

Combining this with Corollary 5.4 yields

$$\text{Prob}((x^G)^3 = G) \geq \text{Prob}(|C_G(x)| < q^{5r/4}) \geq 1 - cq^{-r/4},$$

which tends to 1 as $|G| \to \infty$.

So suppose now that $r < r_1(1/4)$. In particular, since $|G| \to \infty$ we have $q \to \infty$. By Corollary 2.3

$$\text{Prob}((x^G)^3 = G) \geq \text{Prob}(x \text{ is regular semisimple}) \geq 1 - aq^{-1},$$

where the last inequality is Lemma 5.5. We see that in this case too $\text{Prob}((x^G)^3 = G) \to 1$. The result follows.

In fact our arguments show that for some absolute constants $a, b > 0$ we have

$$\text{Prob}((x^G)^3 = G) \geq 1 - aq^{-br}. \qquad \square$$

6. **Proof of main result**

We now turn to the proof of Theorem 1.1. Fix the word $w \neq 1$, and let $G = G_r(q)$. We first quote Proposition 7 of Larsen [17] whose proof relies on algebraic geometry and the Larsen-Pink method.

LEMMA 6.1. *Given a positive integer r and a group word $w \neq 1$ there is a positive constant $c(r, w)$ depending only on r and w such that*

$$|w(G_r(q))| \geq c(r, w)|G_r(q)|.$$

We can now deduce the following.

LEMMA 6.2. *If r is bounded and q is large enough, then $w(G_r(q))$ contains a regular semisimple element.*

Proof. Combining Lemmas 6.1 and 5.5 we see that, if q is sufficiently large (so that $c(r, w) + 1 - aq^{-1} > 1$), then $w(G)$ must intersect the set of regular semisimple elements of G. $\square$

To handle groups of unbounded rank, we may restrict to classical groups. Here it is convenient to deal with quasi-simple groups which are covers of the finite simple classical groups.

LEMMA 6.3. *Consider the natural embedding* $\phi : \mathrm{SL}_2(q^m) \to \mathrm{SL}_{2m}(q)$. *Then if* $m \geq N(w)$ *there exists* $x \in w(\mathrm{SL}_2(q^m))$ *such that* $\phi(x)$ *is regular semisimple in* $\mathrm{SL}_{2m}(q)$.

Proof. Using the fact that $|w(\mathrm{SL}_2(q^m))| \geq \delta q^{3m}$ for some fixed $\delta > 0$ and all large m, it follows as in [17] that for every $\varepsilon > 0$ and $m > f(w, \varepsilon)$ there is a regular semisimple element $x \in w(\mathrm{SL}_2(q^m))$ whose order is at least $(q^m)^{1-\varepsilon}$. We use this for $\varepsilon = 1/4$.

Let $y = \phi(x) \in \mathrm{SL}_{2m}(q)$. Then y is semisimple and we claim it is regular. To see this, note that y has a conjugate $z \in \mathrm{SL}_{2m}(q)$ which can be written with Jordan blocks $\lambda_i \in \mathrm{GL}_{d_i}(q)$ on the diagonal. Since y is in the image of $\mathrm{SL}_2(q^m)$, all these Jordan blocks have the same size $d_i = e$, where e divides $2m$. In this case we see that the order of y divides $q^e - 1$.

This yields $e \geq m(1-\varepsilon) = 3m/4$. It follows that $e = 2m$ or $e = m$. In the first case z consists of a single Jordan block, and so is regular. In the second case z consists of two Jordan blocks λ_1, λ_2 which must be distinct (since x is regular in $\mathrm{SL}_2(q^m)$), so again z is regular, and so is y. □

PROPOSITION 6.4. *Let* w *be a nontrivial group word. There is an absolute constant* c *and a number* $r_1(w)$ *such that if* $G = G_r(q)$ *is a finite simple classical group of rank* $r \geq r_1(w)$, *then there is an element* $x \in w(G)$ *satisfying* $|C_G(x)| < q^{r+c}$.

Proof. It suffices to show this for the natural quasi-simple covers G of the simple classical groups $G_r(q)$, since the result for the projective groups $G/Z(G)$ would easily follow.

Note that the method of [17] for dealing with classical groups of large rank uses embeddings of alternating groups and produces elements of centralizer order $O(q^{r^{1+\varepsilon}})$ which is not sufficiently small for our purpose here.

Instead we shall use embeddings of SL_2 over larger fields to get elements with much smaller centralizers.

If $G = \mathrm{SL}_n(q)$ where $n = 2m$ is even, then Lemma 6.3 shows that, for m large, $w(G)$ contains a regular semisimple element. Note that, if $n = 2m + 1$, then the regular semisimple element constructed inside $\mathrm{SL}_{2m}(q) < \mathrm{SL}_{2m+1}(q)$ remains regular in $\mathrm{SL}_{2m+1}(q)$ (having an extra 1 on the diagonal).

For the other types of classical groups we use natural embeddings of $\mathrm{SL}_n(q)$ as follows:

$$\mathrm{SL}_n(q) < \mathrm{Sp}_{2n}(q),$$

$$\mathrm{SL}_n(q) < \mathrm{SO}_{n,n}(q) < \mathrm{SO}_{2n+1}(q) < \mathrm{SO}_{n+2,n}(q),$$

$$\mathrm{SL}_n(q) < \mathrm{SU}_{2n}(q) < \mathrm{SU}_{2n+1}.$$

It is easy to verify that if $x \in \mathrm{SL}_n(q)$ is a regular semisimple element and G is one of the classical groups in which $\mathrm{SL}_n(q)$ is embedded as above, then we have $|C_G(x)| \leq q^{r+c}$ where r is the rank of G and c is a small constant. Thus the regular semisimple element $x \in w(\mathrm{SL}_n(q)) \subseteq w(G)$ satisfies the required conclusion.

The result follows. □

In fact for our purpose it is convenient to use the following weaker version of Proposition 6.4.

COROLLARY 6.5. *Let w be a nontrivial group word. There is a number $r_1(w)$ such that if $G = G_r(q)$ and $r \geq r_1(w)$, then there is an element $x \in w(G)$ satisfying $|C_G(x)| < q^{5r/4}$.*

We can now prove the following.

THEOREM 6.6. *Let $w_1, w_2, w_3 \neq 1$ be group words. Then there exists a number $N = N(w_1, w_2, w_3)$ such that, if G is a finite simple group of Lie type of order at least N, then $w_1(G)w_2(G)w_3(G) = G$.*

Proof. Let r_1, r_2 be the functions appearing in Corollaries 6.5 and 2.5 respectively. Set

$$r_0 = \max\{r_1(w_1), r_1(w_2), r_1(w_3), r_2(1/4)\}.$$

Let $G = G_r(q)$. If $r \geq r_0$ then, by 6.5, there are elements $x_i \in w_i(G)$ such that $|C_G(x_i)| < q^{5r/4}$ for $i = 1, 2, 3$. Corollary 2.5 now yields

$$w_1(G)w_2(G)w_3(G) \supseteq x_1^G x_2^G x_3^G = G.$$

So suppose $r < r_0$. If G is large enough (larger than a constant depending on w_1, w_2, w_3), then, for $i = 1, 2, 3$, $w_i(G)$ contains a regular semisimple element x_i by Lemma 6.2. Thus $w_i(G) \supseteq x_i^G$, and Corollary 2.2 yields $x_1^G x_2^G x_3^G = G$ provided $|G| \geq c$. This implies the required conclusion. □

Proof of Theorem 1.1. This follows immediately from Theorem 6.6. □

7. **Class expansion**

Results proved in Section 5 show that $C^3 = G$ for large conjugacy classes C of finite simple groups G. In this section we turn our attention to smaller classes C, showing that C^3 is usually significantly larger than C. In particular we prove Theorem 2.7.

Our main tool is Theorem 1.1 of [19], which we state below.

THEOREM 7.1. *There exists an absolute constant $a \in \mathbb{N}$ such that if G is a finite simple group, and $C \subseteq G$ is a normal subset of size greater than* 1, *then there exists a positive integer $k \leq a \log|G|/\log|C|$ such that $C^k = G$.*

From this deduce that a bounded power of a normal subset is fairly large.

COROLLARY 7.2. *There exists an absolute constant $b \in \mathbb{N}$ such that for any finite simple group G and a normal subset $C \subseteq G$ we have*

$$|C^b| \geq \min\{|C|^2, |G|\}.$$

Proof. Let a be as in Theorem 7.1. We claim that $b = 4a$ is as required. Let $C \subseteq G$ be a normal subset. We may assume $|C| > 1$ otherwise the conclusion holds trivially. We distinguish between two cases.

Case 1. $|C| \geq |G|^{1/4}$. Then $\log|G|/\log|C| \leq 4$, and so by Theorem 7.1 we have $C^{4a} = G$ so $|C^b| = |G|$.

Case 2. $|C| < |G|^{1/4}$. We show that in this case $|C^b| \geq |C|^2$. Suppose, by contradiction, that $|C^b| < |C|^2$. Choose $k \leq a \log|G|/\log|C|$ such that $C^k = G$. Let m be the upper integral part of k/b. Then

$$|G| = |C^k| \leq |(C^b)^m| \leq |C^b|^m < (|C|^2)^m = |C|^{2m}. \tag{3}$$

Now,

$$2m < 2\left(\frac{k}{b} + 1\right) = \frac{k}{2a} + 2 \leq \frac{a \log|G|/\log|C|}{2a} + 2 = \frac{1}{2}\log|G|/\log|C| + 2.$$

Our assumption on $|C|$ yields $\log|G|/\log|C| > 4$, and so

$$\frac{1}{2}\log|G|/\log|C| + 2 < \log|G|/\log|C|.$$

It follows that $2m < \log|G|/\log|C|$, which means that $|C|^{2m} < |G|$. This contradicts inequality (3) above. The result follows. □

We also need a recent lemma due to Helfgott.

LEMMA 7.3. *Let $b > 2$ be an integer. Let A be a finite subset of a group G. Suppose $|A^b| \geq |A|^{1+\delta}$ for some $\delta > 0$. Then $|A^3| \geq |A|^{1+\varepsilon}$ where $\varepsilon > 0$ depends only on b and δ.*

Proof. This follows easily from Lemma 2.2 of [13] and its proof. □

We can now prove

THEOREM 7.4. *For every $\delta > 0$ there is $\varepsilon > 0$ such that for any finite simple group G and a normal subset $C \subseteq G$ satisfying $|C| \leq |G|^{1-\delta}$ we have*

$$|C^3| \geq |C|^{1+\varepsilon}.$$

Proof. We may assume $\delta < 1$, otherwise the conclusion holds trivially. By Corollary 7.2 we have $|C^b| \geq |C|^2$ or $|C^b| = |G|$. Our assumption on $|C|$ yields $|C|^2 \geq |C|^{1+\delta}$ and $|G| \geq |C|^{1+\delta}$. Hence in any case we have $|C^b| \geq |C|^{1+\delta}$, and the required conclusion now follows from Lemma 7.3. □

Proof of Theorem 2.7. This follows immediately from Theorem 7.4. □

8. **Almost uniform distributions, II**

The idea of the proof of Theorem 1.1 for groups of Lie type was to find a conjugacy class $C \subset w(G)$ with the property that the distribution $P_{C,C,C}$ is almost uniform in the l_∞-norm. However, we may consider directly the distribution $P_{w^3,G}$ associated with a product of three disjoint copies of w and ask whether it itself is almost uniform. We shall show in this section how properties of the function ζ_G are important in this context too.

A natural machinery to examine this question is the so called noncommutative Fourier transform. As is well known, $P_{w,G}$ is a class function on G, and as such it can be expressed uniquely as a linear combination of irreducible characters. For convenience we write

$$P_{w,G} = |G|^{-1} \sum_{\chi \in \mathrm{Irr}\, G} a_{w,\chi} \chi,$$

where $a_{w,\chi} \in \mathbb{C}$ are the so called Fourier coefficients.

Given class functions $f_1, f_2 : G \to \mathbb{C}$ define their convolution by

$$(f_1 * f_2)(g) = \sum_{g_1 g_2 = g} f_1(g_1) f_2(g_2).$$

It is easy to verify, using the generalized orthogonality relations, that, if $f_1 = \sum a_\chi \chi$ and $f_2 = \sum b_\chi \chi$ then

$$f_1 * f_2 = |G| \sum_\chi \frac{a_\chi b_\chi}{\chi(1)} \chi.$$

Now, if u, v are disjoint words, then we have

$$P_{uv,G} = P_{u,G} * P_{v,G}.$$

This yields the basic relation

$$a_{uv,\chi} = \frac{a_{u,\chi} a_{v,\chi}}{\chi(1)}.$$

In particular it follows by induction on k that

$$a_{w^k,\chi} = \frac{(a_{w,\chi})^k}{\chi(1)^{k-1}}, \tag{4}$$

so

$$P_{w^k,G} = |G|^{-1} \sum_{\chi \in \mathrm{Irr}\, G} \frac{(a_{w,\chi})^k}{\chi(1)^{k-1}} \chi.$$

Using an inverse Fourier transform we may reconstruct the Fourier coefficients $a_{w,\chi}$ as follows.

$$a_{w,\chi} = \frac{1}{|G|^d} \sum_{g_1,\ldots,g_d \in G} \chi(w(g_1,\ldots,g_d)^{-1}),$$

which is the average value of the character χ on $w(\bar{g})^{-1}$. In particular we have $a_{w,1} = 1$ for all words w. We note that equation (4) above was derived by Gallagher in [9].

We can now obtain the following.

PROPOSITION 8.1. *With the above notation we have*

(i) $\|P_{w,G} - U_G\|_\infty \leq \sum_{\chi \neq 1} |a_{w,\chi}| \chi(1)$.

(ii) $\|P_{w^k,G} - U_G\|_\infty \leq \sum_{\chi \neq 1} \frac{|a_{w,\chi}|^k}{\chi(1)^{k-2}}$.

Proof. Fix $g \in G$. Then

$$\left|P_w(g) - |G|^{-1}\right| = |G|^{-1} \left| \left(\sum_\chi a_{w,\chi} \chi(g)\right) - 1 \right| = |G|^{-1} \left| \sum_{\chi \neq 1} a_{w,\chi} \chi(g) \right|.$$

Thus

$$|G| \left|P_w(g) - |G|^{-1}\right| \leq \sum_{\chi \neq 1} |a_{w,\chi} \chi(g)| \leq \sum_{\chi \neq 1} |a_{w,\chi}| \chi(1).$$

This proves part (i). Part (ii) follows by applying part (i) to w^k, using the formula $a_{w^k,\chi} = \frac{(a_{w,\chi})^k}{\chi(1)^{k-1}}$. □

COROLLARY 8.2. (i) *Suppose there is a constant $c = c(w)$ such that for all $\chi \in \mathrm{Irr}\, G$ we have $|a_{w,\chi}| \leq c$. Then*

$$\|P_{w^k,G} - U_G\|_\infty \leq c_1(\zeta_G(k-2) - 1),$$

where c_1 depends on w and k.

(ii) *Suppose there are constants $\varepsilon = \varepsilon(w) > 0$ and $c = c(w)$ such that for all $\chi \in \mathrm{Irr}\, G$ we have $|a_{w,\chi}| \leq c\chi(1)^{1-\varepsilon}$. Then*

$$\|P_{w^k,G} - U_G\|_\infty \leq c_2(\zeta_G(\varepsilon k - 2) - 1),$$

where c_2 depends on w and k.

Proof. This follows immediately from part (ii) Proposition 8.1. □

Now, if G is a finite simple group, then recent results on its zeta function $\zeta_G(s)$ come into play, and yield the following.

THEOREM 8.3. *Let G be a finite simple group, and w a group word.*
(i) *Suppose $G \neq \mathrm{PSL}_2(q)$ and $|a_{w,\chi}| \leq c(w)$ for all $\chi \in \mathrm{Irr}\, G$. Then*

$$\| P_{w^3,G} - U_G \|_\infty \to 0 \text{ as } |G| \to \infty.$$

(ii) *Suppose $|a_{w,\chi}| \leq c\,\chi(1)^{1-\varepsilon}$ for all $\chi \in \mathrm{Irr}\, G$, where $c, \varepsilon > 0$ depend only on w. Then there exists a constant k (depending only on w) such that*

$$\| P_{w^k,G} - U_G \|_\infty \to 0 \text{ as } |G| \to \infty.$$

Proof. By Corollary 8.2(i) we have

$$\| P_{w^3,G} - U_G \|_\infty \leq c_1(\zeta_G(1) - 1).$$

By Theorem 4.2 we have $\zeta_G(1) - 1 \to 0$ as $|G| \to \infty$, provided $G \neq \mathrm{PSL}_2(q)$. This proves part (i).

Part (ii) is proved using 4.2(i) and 8.2(ii). Indeed, if $k > 3/\varepsilon$, then $\varepsilon k - 2 > 1$, hence $\zeta_G(\varepsilon k - 2) \to 0$, and so $\| P_{w^k,G} - U_G \|_\infty \to 0$. □

We now apply this general result for some specific words.

COROLLARY 8.4. *Let G be a finite simple group. Suppose $G \neq \mathrm{PSL}_2(q)$. Then*

$$\| P_{x_1^2 x_2^2 x_3^2,G} - U_G \|_\infty \to 0 \text{ as } |G| \to \infty.$$

Proof. Given a character $\chi \in \mathrm{Irr}\, G$ let $i(\chi)$ be its Schur indicator (see e.g. [14]). Recall that $i(\chi) = 0$ if χ is not a real character, and $i(\chi) \in \{-1, 1\}$ if χ is real. By a classical result of Frobenius and Schur we have

$$N_{x_1^2,G}(g) = \sum_{\chi \in \mathrm{Irr}\, G} i(\chi)\chi(g).$$

This yields the Fourier expansion

$$P_{x_1^2,G} = |G|^{-1} \sum_{\chi \in \mathrm{Irr}\, G} i(\chi)\chi, \tag{5}$$

and so $|a_{x_1^2,\chi}| = |i(\chi)| \leq 1$.

It now follows from part (i) of Theorem 8.3 that $\| P_{x_1^2 x_2^2 x_3^2,G} - U_G \|_\infty \to 0$ as $|G| \to \infty$ □

Remarks. 1. We claim that $G = \mathrm{PSL}_2(q)$ is a genuine exception to this result. To analyze this case write

$$P_{x_1^2 x_2^2 x_3^2,G}(g) = |G|^{-1} \sum_{\chi \in \mathrm{Irr}\, G} i(\chi) \frac{\chi(g)}{\chi(1)^2}.$$

For $g \neq 1$ we obtain

$$|G||P_{x_1^2x_2^2x_3^2,G}(g) - |G|^{-1}| \leq \sum_{1 \neq \chi \in \mathrm{Irr}\, G} \frac{|\chi(g)|}{\chi(1)^2} \leq R(g)(\zeta_G(1) - 1).$$

Since $\zeta_G(1)$ is bounded, while $R(g) \leq 2q^{-1/2}$ tends to zero as $q \to \infty$, the right-hand side above tends to zero. Thus nonidentity elements are obtained with probability $(1 + o(1))|G|^{-1}$.

However, the behavior of the identity element in $\mathrm{PSL}_2(q)$ is different. Indeed, it can be verified from the character table of G that

$$\sum_{\chi \in \mathrm{Irr}\, G} \frac{i(\chi)}{\chi(1)} \to h \text{ as } q \to \infty,$$

where $h = 3/2$ if q is odd, and $h = 2$ if q is even. This shows that $P_{x_1^2x_2^2x_3^2,G}(1) = (h + o(1))|G|^{-1}$, and so, as $q \to \infty$ we have

$$\|P_{x_1^2x_2^2x_3^2,G} - U_G\|_\infty \to h - 1,$$

which is $1/2$ or 1.

2. Product of mth powers for $m > 2$ can also be analyzed in some cases. For example, for alternating groups $G = A_n$ we can show that

$$\|P_{x_1^m \cdots x_{m+1}^m,G} - U_G\|_\infty \to 0 \text{ as } |G| \to \infty.$$

Since this paper focusses on groups of Lie type we shall not include a detailed proof.

For groups of fixed Lie type we can show the following.

PROPOSITION 8.5. *Fix a group word $w \neq 1$ and let $G = G_r(q)$ with r fixed and $q \to \infty$. Then there exists $k = k(r)$ depending only on r such that*

$$\|P_{w^k,G} - U_G\|_\infty \to 0 \text{ as } |G| \to \infty.$$

Proof. We use the upper bound

$$P_{w,G}(1) \leq cq^{-1},$$

where $c = c(r, w)$ depends only on r and w. This bound essentially follows from the fact that at the level of algebraic groups $\bar{G}$ the equation $w = 1$ defines a proper subvariety of $\bar{G}^d$, and from counting q-rational points on algebraic varieties. See Section 4 of [4] for the details.

Another tool we use is Gluck's upper bound on the character ratio $R(g)$, which shows that, if $1 \neq g \in G$ and $1 \neq \chi \in \mathrm{Irr}\, G$, then

$$|\chi(g)| \leq aq^{-1/2}\chi(1),$$

where a is some absolute constant (see [10]).

Using these bounds we obtain

$$|a_{w,G}| \le |G|^{-d} \sum_{g_1,\dots,g_d \in G} |\chi(w(g_1,\dots,g_k)^{-1})| \le P_{w,G}(1)\chi(1) + aq^{-1/2}\chi(1).$$

This yields

$$|a_{w,\chi}| \le (cq^{-1} + aq^{-1/2})\chi(1) \le bq^{-1/2}\chi(1),$$

for some constant b. Therefore, by 8.1(ii) we have

$$\|P_{w^k,G} - U_G\|_\infty \le \sum_{\chi \ne 1} \frac{|a_{w,\chi}|^k}{\chi(1)^{k-2}} \le \sum_{\chi \ne 1} \frac{b^k q^{-k/2}\chi(1)^k}{\chi(1)^{k-2}} \le b^k q^{-k/2} \sum_\chi \chi(1)^2.$$

This yields

$$\|P_{w^k,G} - U_G\|_\infty \le b^k q^{-k/2}|G|.$$

Now, we have $|G_r(q)| \le q^{4r^2}$. Therefore, fixing $k > 8r^2$, we see that

$$\|P_{w^k,G} - U_G\|_\infty \to 0 \text{ as } q \to \infty. \qquad \square$$

As we shall see in the next section, the methods described here for general words w work even better for the commutator word.

9. **Conjectures of Ore and Thompson**

In this section we discuss some results related to following longstanding conjectures by Ore and Thompson.

ORE CONJECTURE. *Every element in a finite simple group is a commutator.*

THOMPSON CONJECTURE. *Every finite simple group G has a conjugacy class C such that $C^2 = G$.*

It is easy to see that Thompson conjecture implies Ore conjecture. Both conjectures are known to be true for alternating groups, and for groups of Lie type over fields with more than 8 elements [5].

We show below how results from Section 8, properties of $\zeta_G(s)$, and Corollary 2.5 shed new light on these difficult problems.

Set $w = [x_1, x_2]$. Then w^k denotes a product of k commutators in independent variables. We first the state a well known result.

LEMMA 9.1. *Let w be the commutator word, k a positive integer, and G a finite group. Then*

$$P_{w^k,G} = |G|^{-1} \sum_{\chi \in \mathrm{Irr}(G)} \chi(1)^{-(2k-1)}\chi.$$

Proof. A classical result of Frobenius shows that

$$P_{[x_1,x_2],G} = |G|^{-1} \sum_{\chi \in \mathrm{Irr}\, G} \chi(1)^{-1} \chi. \tag{6}$$

Hence, in the notation of the previous section, we have

$$a_{[x_1,x_2],\chi} = \chi(1)^{-1}.$$

In view of (4) we obtain

$$a_{w^k,G} = \frac{(\chi(1)^{-1})^k}{\chi(1)^{k-1}} = \chi(1)^{-(2k-1)}.$$

The result follows. □

LEMMA 9.2. *With the above notation we have*

$$\| P_{w^k,G} - U_G \|_\infty \leq \zeta_G(2k-2) - 1.$$

Proof. Set

$$\Delta(g) = \sum_{1 \neq \chi \in \mathrm{Irr}(G)} \frac{\chi(g)}{\chi(1)^{2k-1}}.$$

By Lemma 9.1 we have

$$|P_{w^k,G}(g) - |G|^{-1}| \leq |G|^{-1} |\Delta(g)|.$$

Since $|\chi(g)| \leq \chi(1)$ we have

$$|\Delta_g| \leq \sum_{1 \neq \chi \in \mathrm{Irr}(G)} \chi(1)^{-(2k-2)} = \zeta_G(2k-2) - 1.$$

The result follows. □

Proof of Theorem 2.8. We claim that, for $k \geq 2$, and for G simple, $P_{w^k,G}$ is almost uniform in the l_∞-norm. Indeed, by Theorem 1.1 of [22] (see Theorem 4.2 here), $\zeta_G(s) \to 1$ as $|G| \to \infty$, for any fixed $s > 1$. Hence $\zeta_G(2k-2) - 1 \to 0$, and so Lemma 9.2 implies that $\| P_{w^k,G} - U_G \|_\infty \to 0$ as $|G| \to \infty$.

This proves the claim. The case $k = 2$ gives Theorem 2.8. □

In particular, we conclude that every element of a large finite simple group is a product of two commutators.

To prove Theorem 2.9 we need the following.

LEMMA 9.3. *For every finite group G and an element $g \in G$ we have*

$$P_{[x_1,x_2],G}(g) \geq |G|^{-1}(1 - E(G,g)).$$

Consequently, if $E(G,g) < 1$ then g is a commutator in G.

Proof. Recall that

$$E(G,g) = \sum_{1 \neq \chi \in \mathrm{Irr}(G)} \frac{|\chi(g)|}{\chi(1)}.$$

By (6) we have

$$P_{[x_1,x_2],G}(g) = |G|^{-1}(1 + \Delta(g)),$$

where

$$\Delta(g) = \sum_{1 \neq \chi \in \mathrm{Irr}(G)} \frac{\chi(g)}{\chi(1)}.$$

Since $|\Delta(g)| \leq E(G,g)$ the result follows. □

We can now show that every element with a small centralizer is a commutator in $G_r(q)$.

THEOREM 9.4. *There exists an absolute constant $\varepsilon > 0$ such that, if $G = G_r(q)$, then every element $g \in G$ satisfying*

$$|C_G(g)| \leq \varepsilon \cdot q^{3r}$$

is a commutator in G.

Proof. We may assume $q \leq 8$. Remark 1, following Proposition 4.7, shows that

$$E(G,g) \leq cq^{\sqrt{3r+b}-(r-1)} + c|C_G(g)|^{1/2}q^{3-3r/2}.$$

Denote the two summands on the right-hand side by A and B respectively. Choose r_0 such that for all $r > r_0$ we have $A < 1/2$. Note that, for some absolute constant $a > 0$ we have $|C_G(x)| \geq aq^r$ for all $x \in G$. Now, choose $\varepsilon > 0$ such that

$$a/\varepsilon > 64^{r_0} \text{ and } \varepsilon < 2^{-14}c^{-2}.$$

Suppose $|C_G(g)| \leq \varepsilon \cdot q^{3r}$. Then $aq^r \leq \varepsilon q^{3r}$ so $a/\varepsilon \leq q^{2r} \leq 64^r$. By the choice of ε this yields $r > r_0$, and so $A < 1/2$.

Next we have

$$B = c|C_G(g)|^{1/2}q^{3-3r/2} \leq c\varepsilon^{1/2}q^3 \leq 64c\varepsilon^{1/2}.$$

By the choice of ε it follows that $B < 1/2$, and so

$$E(G,g) \leq A + B < 1.$$

Thus g is a commutator by Lemma 9.3. □

Proof of Theorem 2.9(i). We may assume $G = G_r(q)$ with $q \leq 8$. By Theorem 9.4 and Corollary 5.4 we have

$$\mathrm{Prob}(g \in G \text{ is a commutator}) \geq \mathrm{Prob}(|C_G(g)| \leq \varepsilon q^{3r}) \geq 1 - c\varepsilon q^{-2r}.$$

The result follows. □

PROPOSITION 9.5. *Let $G = G_r(q)$, let $C_1, C_2 \subset G$ be conjugacy classes, and let $x_i \in C_i$ $(i = 1, 2)$.*

(i) *If x_1, x_2 are regular semisimple then C_1C_2 contains all semisimple elements of G.*

(ii) *If x_1, x_2 are regular semisimple, $\varepsilon > 0$, and $r \geq r(\varepsilon)$, then C_1C_2 contains all elements $g \in G$ satisfying $|C_G(g)| \leq q^{(2-\varepsilon)r}$.*

(iii) *If $|C_G(x_i)| \leq q^{5r/4}$ for $i = 1, 2$, and $r \geq c$, then C_1C_2 contains all elements $g \in G$ satisfying $|C_G(g)| \leq q^{5r/4}$.*

(iv) *If $|C_G(x_1)|\,|C_G(x_2)| \leq q^{(3-\varepsilon)r}$ where $\varepsilon > 0$, and $r \geq r(\varepsilon)$, then C_1C_2 contains all elements $g \in G$ satisfying $|C_G(g)| \leq q^{(1+\varepsilon)r}$.*

Proof. Part (i) is proved in [11]. To prove part (ii) we use Corollary 2.5. Let D be the conjugacy class of g in G, where $|C_G(g)| \leq q^{(2-\varepsilon)r}$. Thus D^{-1} is the class of g^{-1}. We have $|C_G(x_i)| \leq cq^r$, so

$$|C_G(x_1)||C_G(x_2)||C_G(g^{-1})| \leq (cq^r)^2 q^{(2-\varepsilon)r} = c^2 q^{(4-\varepsilon)r}.$$

If $r \geq r_1(\varepsilon)$ then $c^2 q^{(4-\varepsilon)r} \leq q^{(4-\varepsilon/2)r}$. If, in addition, $r \geq r_2(\varepsilon/2)$, where r_2 is as in 2.5, it follows that

$$C_1C_2D^{-1} = G.$$

In particular $1 \in C_1C_2D^{-1}$ and this implies $D \subseteq C_1C_2$, and $g \in C_1C_2$, as required.

Part (iii) is proved in a similar manner, using Corollary 2.5 with $\varepsilon = 1/4$ and $c = r_2(1/4)$. Part (iv) follows again from 2.5. The result follows. □

COROLLARY 9.6. *Let $G = G_r(q)$, $x \in G$, and $C = x^G$.*

(i) *If x is regular semisimple then*

$$|C^2|/|G| \to 1 \text{ if } |G| \to \infty.$$

Moreover, we have

$$|C^2|/|G| \geq 1 - q^{-(1-\varepsilon)r},$$

provided $r \geq r(\varepsilon)$.

(ii) *If $|C_G(x)| \leq q^{5r/4}$ and $r \geq c$ then*

$$|C^2|/|G| \to 1 \text{ if } |G| \to \infty.$$

(iii) *More generally, if $x_1, x_2 \in G$ satisfy*

$$|C_G(x_1)||C_G(x_2)| \leq q^{(3-\varepsilon)r},$$

and $C_i = x_i^G$ $(i = 1, 2)$, then, for $r \geq r(\varepsilon)$ we have

$$|C_1C_2|/|G| \to 1 \text{ as } |G| \to \infty.$$

Proof. By Proposition 9.5(i), C^2 contains all regular semisimple elements of G, and by Lemma 5.5 the number of these elements is at least $(1 - aq^{-1})|G|$.

This shows that

$$|C^2|/|G| \geq 1 - aq^{-1},$$

which proves part (i) when r is bounded (so $q \to \infty$).

So suppose now that $r \to \infty$. In this case we use 9.5(ii) to conclude that, if $r \geq r(\varepsilon/2)$, then C^2 contains all elements with centralizer order at most $q^{(2-\varepsilon/2)}$. Combining this with 5.4 we obtain for r large enough

$$|C^2|/|G| \geq 1 - cq^{-(1-\varepsilon/2)r} \geq 1 - q^{-(1-\varepsilon)r},$$

as required. In particular $|C^2|/|G| \to 1$ as $|G| \to \infty$. Part (i) is proved.

The proof of part (ii) is similar, combining 9.5(iii) with 5.4. Part (iii) follows from 9.5(iv) and 5.4. □

Proof of Theorem 2.9(ii). It suffices to prove the theorem for $G = G_r(q)$ with $q \leq 8$. Thus $r \to \infty$ in our case. Part (i) of Corollary 9.6 now shows the existence of a class $C_G \subset G$ such that $|C_G^2|/|G| \to 1$. Moreover, letting $C = C_G$ be a class of regular semisimple elements, and using the remark following Corollary 2.5, we see that C^2 contains all elements $g \in G$ satisfying

$$|C_G(g)| \leq \delta \cdot q^{2r},$$

where $\delta > 0$ is some absolute constant. Combining this with 5.4 yields

$$|C^2|/|G| \geq 1 - cq^{-r},$$

as required. □

To prove Theorem 2.10 we need the following.

PROPOSITION 9.7. *Let $G = G_r(q)$ and define*

$$S = \{s \in G : |C_G(s)| \leq q^{5r/4}\}.$$

Then, if $r \geq c$, every $s \in S$ can be expressed as $s = [s_1, g_1]$ where $s_1 \in S$ is conjugate to s and $g_1 \in G$. Consequently, all the elements of S are d-fold commutators for all $d \geq 2$.

Proof. Suppose $s \in S$ is given, and let $C = s^G$. Let c be as in 9.5(iii). Then $r \geq c$ implies $C^{-1}C \supseteq S$.

In particular, there is $s_1 \in C$ and $g_1 \in G$ such that $s = s_1^{-1}s_1^{g_1} = [s_1, g_1]$. This proves the first assertion.

Now, in a similar manner we may write $s_1 = [s_2, g_2]$ with $s_2 \in C \subseteq S$, and so $s = [[s_2, g_2], g_1] = [s_2, g_2, g_1]$. Proceeding by induction it follows that s is a d-fold commutator for all d. □

Proof of Theorem 2.10. The result is clear for groups for which Ore conjecture holds, so we may assume $G = G_r(q)$ with $q \leq 8$. Hence, as $|G| \to \infty$ we have $r \to \infty$, so we may assume $r \geq c$. Using Proposition 9.7 and its notation it follows

that the probability that $x \in G$ is a d-fold commutator for all d is at least $|S|/|G|$. By 5.4 we have $|S|/|G| \to 1$, proving part (i) of the theorem.

Part (ii) follows from (i). Indeed, it suffices to take $|G|$ large enough so that $|S|/|G| > 1/2$ to deduce that $S^2 = G$, hence the result. □

10. **Open problems and examples**

It is interesting to find out whether Theorem 1.1 is best possible. We pose the following

Problem 10.1. Let $w \neq 1$ be a group word. Is there a positive integer $N = N(w)$ such that for every finite simple group G with $|G| \geq N(w)$ we have $w(G)^2 = G$?

We have shown that this holds for example for the commutator word.

Additional positive evidence is given below.

PROPOSITION 10.2. *Let $w \neq 1$ be a group word and let G be a finite simple group of Lie type. Then*

$$|w(G)^2|/|G| \to 1 \text{ as } |G| \to \infty.$$

Proof. Let c be the constant appearing in part (ii) of Corollary 9.6, and $r_1(w)$ be as in Corollary 6.5. Let $G = G_r(q)$.

If $r \geq \max\{c, r_1(w)\}$ then by 6.5 there exists $x \in w(G)$ such that $|C_G(x)| \leq q^{5r/4}$. Let $C = x^G$. Then $C \subseteq w(G)$, and by 9.6(ii) we have $|C^2|/|G| \to 1$ as $|G| \to \infty$. Hence $|w(G)^2|/|G| \to 1$.

Suppose now that $r < \max\{c, r_1(w)\}$. Then by Lemma 6.2, if G is large enough there exists $x \in w(G)$ which is regular semisimple. Corollary 9.6(i) shows that $|C^2|/|G| \to 1$ as $|G| \to \infty$. Hence in this case too we have $|w(G)^2|/|G| \to 1$. The result follows. □

In fact, using Proposition 6.4, it follows that for $G = G_r(q)$ with $r \geq r(\varepsilon)$ we have

$$|w(G)^2|/|G| \geq 1 - q^{-(1-\varepsilon)r}.$$

Consider the following related problem on class expansion.

CONJECTURE 10.3. *There exists an absolute constant $\varepsilon > 0$ such that for every finite simple group G and every conjugacy class C of G we have*

$$|C^2| \geq \min\{|C|^{1+\varepsilon}, |G| - \delta\},$$

where $\delta = 0$ if C is real, and $\delta = 1$ otherwise.

Recall that we obtained an expansion result for C^3, which makes the above extension quite plausible. Its importance stems from the fact that it implies Thompson Conjecture and Ore Conjecture up to finitely many exceptions. Indeed, if $\varepsilon \leq 1$ is as in 10.3, then $C^2 = G$ for any real conjugacy C of G of size at least $|G|^{1-\varepsilon/2}$

(since then $|C|^{1+\varepsilon} \geq |G|$). Now, groups of large Lie rank (namely rank $r \geq f(\varepsilon)$) have such conjugacy classes, so Thompson conjecture would follow.

In certain cases we can indeed obtain an expansion result for C^2, though not as strong as in 10.3.

PROPOSITION 10.4. *There is an absolute constant $b > 0$ such that for every finite simple group of Lie type over the field with q elements, and every nonidentity conjugacy class C of G, we have*

$$|C^2| \geq |C| \cdot bq^{1/2}.$$

Consequently, if G has rank r, then

$$|C^2| \geq |C|^{1+\varepsilon},$$

where $\varepsilon > 0$ depends only on r.

Proof. Let $C = x^G$ and $1 \neq g \in G$. Recall that $N_{C,C}(g)$ is the number of solutions to the equation $y_1 y_2 = g$ with $y_i \in C$. By (1) we have

$$N_{C,C}(g) = \frac{|C|^2}{|G|}(1 + \Delta(g)),$$

where

$$\Delta(g) = \sum_{1 \neq \chi \in \operatorname{Irr} G} \frac{\chi(x)^2 \chi(g^{-1})}{\chi(1)}.$$

Clearly

$$|\Delta(g)| \leq \sum_{\chi \neq 1} |\chi(x)|^2 \cdot R(g) = (|C_G(x)| - 1)R(g).$$

We have already noted that $R(g) \leq aq^{-1/2}$ for some absolute constant a (see [10]). This gives

$$N_{C,C}(g) \leq \frac{|C|^2}{|G|}(1 + |C_G(x)|aq^{-1/2}) \leq \frac{|C|^2}{|G|}|C_G(x)|a_1 q^{-1/2} = a_1|C|q^{-1/2},$$

where $a_1 > 0$ is some absolute constant.

Now, since $N_{C,C}(1) \leq |C|$ we clearly have

$$|C|^2 \leq |C| + \sum_{1 \neq g \in C^2} N_{C,C}(g) \leq |C| + |C^2| \cdot a_1|C|q^{-1/2}.$$

This yields $|C| \leq 1 + |C^2| \cdot a_1 q^{-1/2}$, which implies

$$|C^2| \geq (|C| - 1) \cdot a_1^{-1} q^{1/2}.$$

This implies the required conclusion. □

This result provides an affirmative answer to Conjecture 10.3 for groups of bounded rank.

We conclude this paper with a problem regarding certain distributions. The modern approach to Waring's problem (based on Hardy and Littlewood's circle method and other methods) is estimating the number of representations of a number as a sum of a given number of kth powers. In our noncommutative context, the analogous approach is to study the distributions associated with words w and their disjoint powers. For example, can we strengthen Theorem 1.1, and show that not only $w(G)^3 = G$, in fact $P_{w^3,G}$ is almost uniform?

Problem 10.5. Let $w \neq 1$ be a group word, and let G be a finite simple group.

(i) When can we deduce that

$$\|P_{w^3,G} - U_G\|_\infty \to 0 \text{ as } |G| \to \infty?$$

(ii) Show that there exists $k = k(w)$ depending only on w such that we always have

$$\|P_{w^k,G} - U_G\|_\infty \to 0 \text{ as } |G| \to \infty.$$

Note that the answer to part (i) is not always positive. Indeed, $w = x_1^2$ and $G = \mathrm{PSL}_2(q)$ provide a counter-example (though $\|P_{w^3,G} - U_G\|_\infty$ is still bounded in this case). We have shown that part (ii) holds for groups of bounded rank (see 8.5 above).

The methods of Section 8 provide a possible approach to solving Problem 10.5. For example, to prove (ii) it suffices to show that, if $w \neq 1$, then there exist positive constants $\varepsilon = \varepsilon(w)$ and $c = c(w)$ such that for every finite simple group G and every $\chi \in \mathrm{Irr}\, G$ we have

$$a_{w,\chi} \leq c\,\chi(1)^{1-\varepsilon}.$$

Finally, note that while we do not have a counter-example to Problem 10.1, the analogous probabilistic version fails drastically, in the sense that $P_{w^2,G}$ may be highly nonuniform. To see this, let $w = x_1^2$. It follows from (4) and (5) that

$$P_{x_1^2x_2^2,G}(g) = |G|^{-1} \sum_{\chi \in \mathrm{Real}(G)} \chi(g)/\chi(1),$$

where $\mathrm{Real}(G)$ denotes the set of real characters of G. In particular we see that the probability that $x_1^2x_2^2 = 1$ is

$$P_{x_1^2x_2^2,G}(1) = |G|^{-1}|\mathrm{Real}(G)|,$$

which is typically much larger than $|G|^{-1}$. Indeed, for $G = A_n$ or $\mathrm{PSL}_n(q)$ we easily obtain $\mathrm{Real}(G) \to \infty$ as $|G| \to \infty$, and this yields

$$\|P_{x_1^2x_2^2,G} - U_G\|_\infty \geq |\mathrm{Real}(G)| - 1 \to \infty.$$

References

[1] Z. ARAD and M. HERZOG (eds.), *Products of Conjugacy Classes in Groups*, *Lecture Notes in Math.* **1112**, Springer-Verlag, New York, 1985. MR 87h:20001 Zbl 0561.2004

[2] A. BOREL, On free subgroups of semisimple groups, *Enseign. Math.* **29** (1983), 151–164. MR 85c:22009 Zbl 0533.22009

[3] J. L. BRENNER, Covering theorems for FINASIGs. VIII. Almost all conjugacy classes in $\mathcal{A}_n$ have exponent ≤ 4, *J. Austral. Math. Soc.* **25** (1978), 210–214. MR 58 #872 Zbl 0374.20039

[4] J. D. DIXON, L. PYBER, Á. SERESS, and A. SHALEV, Residual properties of free groups and probabilistic methods, *J. reine angew. Math.* (*Crelle's*) **556** (2003), 159–172. MR 2004g:20093 Zbl 1027.20013

[5] E. W. ELLERS and N. GORDEEV, On the conjectures of J. Thompson and O. Ore, *Trans. Amer. Math. Soc.* **350** (1998), 3657–3671. MR 98k:20022 Zbl 0910.20007

[6] P. ERDŐS and P. TURÁN, On some problems of a statistical group-theory. I, *Z. Wahr. Verw. Gebiete* **4** (1965), 175–186. MR 32 #2465 Zbl 0137.25602

[7] ———, On some problems of a statistical group-theory. II, *Acta math. Acad. Sci. Hungar.* **18** (1967), 151–163. MR 34 #7624 Zbl 0189.31302

[8] J. FULMAN and R. M. GURALNICK, Bounds on the number and sizes of conjugacy classes in finite Chevalley groups, preprint.

[9] P. X. GALLAGHER, The generation of the lower central series, *Canadian J. Math.* **17** (1965), 405–410. MR 30 #4826 Zbl 0134.26202

[10] D. GLUCK, Sharper character value estimates for groups of Lie type, *J. Algebra* **174** (1995), 229–266. MR 96m:20021 Zbl 0842.20014

[11] R. GOW, Commutators in finite simple groups of Lie type, *Bull. London Math. Soc.* **32** (2000), 311–315. MR 2001b:20024 Zbl 1021.20012

[12] R. M. GURALNICK and F. LÜBECK, On p-singular elements in Chevalley groups in characteristic p, in *Groups and Computation*, III (*Columbus*, OH, 1999), *Ohio State Univ. Math. Res. Inst. Publ.* **8**, de Gruyter, Berlin, 2001, pp. 169–182. MR 2002d:20074 Zbl 1001.20045

[13] H. A. HELFGOTT, Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, *Ann. of Math.* **167** (2008), 601–623. MR 2415382 Zbl 05578700

[14] I. M. ISAACS, *Character Theory of Finite Groups*, *Pure and Appl. Math.* **69**, Academic Press, New York, 1976. MR 57 #417 Zbl 0337.20005

[15] G. A. JONES, Varieties and simple groups, *J. Austr. Math. Soc.* **17** (1974), 163–173. MR 49 #9081 Zbl 0286.20028

[16] V. LANDAZURI and G. M. SEITZ, On the minimal degrees of projective representations of the finite Chevalley groups, *J. Algebra* **32** (1974), 418–443. MR 50 #13299 Zbl 0325.20008

[17] M. LARSEN, Word maps have large image, *Israel J. Math.* **139** (2004), 149–156. MR 2004k: 20094 Zbl 1130.20310

[18] M. W. LIEBECK and L. PYBER, Upper bounds for the number of conjugacy classes of a finite group, *J. Algebra* **198** (1997), 538–562. MR 99c:20023 Zbl 0892.20017

[19] M. W. LIEBECK and A. SHALEV, Diameters of finite simple groups: sharp bounds and applications, *Ann. of Math.* **154** (2001), 383–406. MR 2002m:20029 Zbl 1003.20014

[20] ———, Fuchsian groups, coverings of Riemann surfaces, subgroup growth, random quotients and random walks, *J. Algebra* **276** (2004), 552–601. MR 2005e:20076 Zbl 1068.20052

[21] M. W. LIEBECK and A. SHALEV, Character degrees and random walks in finite groups of Lie type, *Proc. London Math. Soc.* **90** (2005), 61–86. MR 2006h:20016 Zbl 1077.20020

[22] ———, Fuchsian groups, finite simple groups and representation varieties, *Invent. Math.* **159** (2005), 317–367. MR 2005j:20065 Zbl 1134.20059

[23] G. LUSZTIG, Characters of Reductive Groups over Finite Fields, in *Ann. Math. Studies*, Princeton Univ. Press, Princeton, NJ, 1984. MR 86i:20062 Zbl 0572.20026

[24] G. MALLE, J. SAXL, and T. WEIGEL, Generation of classical groups, *Geom. Dedicata* **49** (1994), 85–116. MR 95c:20068 Zbl 0832.20029

[25] C. MARTINEZ and E. ZELMANOV, Products of powers in finite simple groups, *Israel J. Math.* **96** (1996), 469–479. MR 97k:20050 Zbl 0890.20013

[26] M. B. NATHANSON, *Additive Number Theory*: *The Classical Bases*, *Grad. Texts in Math.* **164**, Springer-Verlag, New York, 1996. MR 97e:11004

[27] N. NIKOLOV, On the commutator width of perfect groups, *Bull. London Math. Soc.* **36** (2004), 30–36. MR 2004m:20055 Zbl 1048.20013

[28] N. NIKOLOV and D. SEGAL, On finitely generated profinite groups. II. Products in quasisimple groups, *Ann. of Math.* **165** (2007), 239–273. MR 2008f:20053 Zbl 1126.20018

[29] ———, On finitely generated profinite groups. II. Products in quasisimple groups, *Ann. of Math.* **165** (2007), 239–273. MR 2008f:20053 Zbl 1126.20018

[30] O. ORE, Some remarks on commutators, *Proc. Amer. Math. Soc.* **2** (1951), 307–314. MR 12,671e Zbl 0043.02402

[31] J. SAXL and J. S. WILSON, A note on powers in simple groups, *Math. Proc. Cambridge Philos. Soc.* **122** (1997), 91–94. MR 98e:20022 Zbl 0890.20014

[32] D. SEGAL, Closed subgroups of profinite groups, *Proc. London Math. Soc.* **81** (2000), 29–54. MR 2001f:20058 Zbl 1030.20017

[33] J.-P. SERRE, *Topics in Galois Theory*, *Res. Notes in Math.* **1**, Jones and Bartlett Publishers, Boston, MA, 1992. MR 94d:12006 Zbl 0746.12001

[34] J. WILSON, First-order group theory, in *Infinite Groups* 1994 (*Ravello*), de Gruyter, Berlin, 1996, pp. 301–314. MR 99f:03045 Zbl 0866.20001

[35] E. WITTEN, On quantum gauge theories in two dimensions, *Comm. Math. Phys.* **141** (1991), 153–209. MR 93i:58164 Zbl 0762.53063

(Received January 24, 2006)
(Revised July 2, 2006)

E-mail address: shalev@math.huji.ac.il
THE HEBREW UNIVERSITY OF JERUSALEM, EINSTEIN INSTITUTE OF MATHEMATICS, 91904 JERUSALEM, ISRAEL