WIDTH QUESTIONS FOR FINITE SIMPLE GROUPS

MARTIN W. LIEBECK

Department of Mathematics, Imperial College, London SW7 2BZ, UK
Email: m.liebeck@imperial.ac.uk

Abstract

Let G be a finite group generated by a collection $\mathcal{S}$ of subsets of G. Define the width of G with respect to $\mathcal{S}$ to be the minimal integer n such that G is equal to the union of a product of n subsets in $\mathcal{S}$, together with all subproducts. For example, when $\mathcal{S}$ consists of a single subset, the width is just the diameter of the Cayley graph of G with respect to this subset. This article contains a discussion of a variety of problems concerning the width of simple groups, mainly in the following cases: (1) the case where $\mathcal{S}$ consists of a single subset; (2) the case where $\mathcal{S}$ is closed under conjugation. There are many examples of special interest. Particular emphasis is given to recent results and problems concerning the "word width" of simple groups – namely, the width in the case where $\mathcal{S}$ consists of all values in G of a fixed word map. Also discussed are combinatorial interpretations of some width problems, such as the estimation of diameters of orbital graphs.

1 Introduction

Let G be a finite group, and suppose $\mathcal{S}$ is a collection of subsets of G such that G is generated by their union. Every element $g \in G$ has an expression $g = t_1 \dots t_k$ where $t_i \in T_i \in \mathcal{S}$. Hence it is possible to write G as the union of a product $T_1 \cdots T_d := \{t_1 \dots t_d : t_i \in T_i\}$, together with all subproducts $T_{i_1} \cdots T_{i_k}$ $(i_1 < \cdots < i_k)$, where each $T_i \in \mathcal{S}$ and repeats are allowed among the T_i. We define the *width* of G with respect to $\mathcal{S}$ to be the minimal such positive integer d, and denote this by $\mathrm{width}(G, \mathcal{S})$.

In this article, we consider the problem of finding, or bounding, the width of finite groups in various cases of interest, mainly when G is a finite non-abelian simple or almost simple group. We remind the reader that the finite non-abelian simple groups are the alternating groups of degree at least 5, the simple groups of Lie type over finite fields, and the 26 sporadic groups; and an *almost simple* group is a group G such that $S \lhd G \leq \mathrm{Aut}(S)$ for some non-abelian simple group S. For brevity in the text below, whenever we say a group G is simple, we mean that G is a finite non-abelian simple group.

Examples Here are two contrasting examples of such width problems.

1. Let $G = S_n$, the symmetric group of degree n, and let $\mathcal{S} = \{T\}$, where T is the set of all transpositions. Then $\mathrm{width}(G, \mathcal{S})$ is the minimal value of d such that $S_n = T^d \cup T^{d-1} \cup \cdots \cup \{1\}$ (where $T^k := \{t_1 \dots t_k : t_i \in T\}$). Since every

permutation can be expressed as a product of at most $n-1$ transpositions, and such an expression for an n-cycle requires precisely this number, the width in this example is $n-1$.

2. Again let $G = S_n$, but this time let $\mathcal{S} = \{\langle t_1\rangle, \ldots, \langle t_k\rangle\}$, where $t_1, \ldots, t_k$ are all the transpositions in G (and $k = \binom{n}{2}$). Here the width problem is more subtle than in the previous example: $\mathrm{width}(G, \mathcal{S})$ is the minimal value of d for which we can write $S_n = \langle t_{i_1}\rangle \cdots \langle t_{i_d}\rangle$ (repeats allowed). Notice that the right hand side has at most 2^d elements while the left has $n!$, so the width d must be at least the order of $n \log n$. The question of whether the width in this example does have this order of magnitude is not so easy; we shall give the answer in Section 3.2 (see the proof of Theorem 3.9).

All the width questions we shall discuss in these lectures are of one of the two types in the above examples:

(a) the case where $\mathcal{S}$ consists of a single generating subset S of G

(b) the case where $\mathcal{S}$ consists of a conjugacy class of subsets of G: that is,

$$\mathcal{S} = \{A^g : g \in G\}$$

for some subset A of G.

In case (a), the width is just the diameter of the Cayley graph of G with respect to S. We shall discuss recent developments on this topic for simple groups in the next section. There are many interesting questions arising from case (b), and these will be the focus of the remaining sections.

2 Width, Cayley graphs and orbital graphs

Let G be a finite group with a generating set S which is symmetric – that is, closed under taking inverses – and does not contain the identity. The *Cayley graph* $\Gamma(G, S)$ is defined to be the graph with vertex set G and edges $\{g, gs\}$ for all $g \in G, s \in S$. It is connected and regular of valency $|S|$, and G acts regularly on $\Gamma(G, S)$ by left multiplication. Because of the transitive action of G, the diameter of $\Gamma(G, S)$, denoted by $\mathrm{diam}(G, S)$, is equal to the maximum distance between the identity element and any $g \in G$, and so $\mathrm{diam}(G, S) = \max\{l(g) : g \in G\}$, where $l(g)$ is the length of the shortest expression for g as a product of elements of S. If $d = \mathrm{diam}(G, S)$, then d is minimal such that $G = S^d \cup S^{d-1} \cup \cdots \cup \{1\}$, and hence

$$\mathrm{diam}(G, S) = \mathrm{width}(G, \{S\}).$$

Also $|G| \leq \sum_{r=0}^{d} |S|^r < |S|^{d+1}$. Hence

$$\mathrm{diam}(G, S) > \frac{\log |G|}{\log |S|} - 1. \tag{1}$$

Examples

1. Let $G = C_n = \langle x \rangle$, a cyclic group of order n, and let $S = \{x, x^{-1}\}$. Then $\Gamma(G, S)$ is an n-gon. So $\mathrm{diam}(G, S)$ is $[\frac{n}{2}]$, whereas $\frac{\log |G|}{\log |S|}$ is $\frac{\log n}{\log 2}$.
2. Let $G = S_n$ and S be the set of all transpositions. Here $\mathrm{diam}(G, S)$ is $n - 1$, while $\frac{\log |G|}{\log |S|}$ is roughly $\frac{n}{2}$.
3. Let $G = S_n$ and $S = \{(1\,2), (1\,2\cdots n)^{\pm 1}\}$. In this case $\mathrm{diam}(G, S)$ is roughly n^2, while $\frac{\log |G|}{\log |S|}$ is of the order of $n \log n$. The same orders of magnitude apply to a similar generating set for A_n consisting of a 3-cycle and an n- or $(n-1)$-cycle and their inverses.
4. Let $G = SL_n(q)$ and S be the set of transvections. Then $\mathrm{diam}(G, S) \approx n$ and $\frac{\log |G|}{\log |S|} \approx \frac{n}{2}$.
5. Let $G = SL_n(p)$ (p prime) and $S = \{x^{\pm 1}, y^{\pm 1}\}$ where

$$x = \begin{pmatrix} 1 & 1 & & & \\ & 1 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & 1 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ & & & \cdot & \\ & & & & \cdot \\ \pm 1 & & & & \end{pmatrix}$$

   Then $\frac{\log |G|}{\log |S|} \sim n^2 \log p$, and also $\mathrm{diam}(G, S) \sim n^2 \log p$.

All the above examples are elementary except the last, where the fact that $\mathrm{diam}(G, S) \leq Cn^2 \log p$ for some constant C is a result of Kassabov and Riley [32].

2.1 Babai's Conjecture

Define $\mathrm{diam}(G)$ to be the maximum of $\mathrm{diam}(G, S)$ over all generating sets S. The main conjecture in the field is due to Babai, and appears as Conjecture 1.7 in [6]:

Babai's Conjecture *There is a constant c such that* $\mathrm{diam}(G) < (\log |G|)^c$ *for any non-abelian finite simple group G.*

It can be seen from Example 3 above that c must be at least 2 for the conjecture to hold.

There have been spectacular recent developments on Babai's conjecture, both for groups of Lie type and for alternating groups. We shall discuss these separately.

2.1.1 Groups of Lie type

For a long time, even $SL_2(p)$ (p prime) was a mystery as far as proving Babai's conjecture was concerned. Probably the first small (symmetric) generating set one thinks of for this group is

$$S = \{ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\pm 1}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\pm 1} \}.$$

Babai's conjecture asserts that $\mathrm{diam}(G,S) < (\log p)^c$ for these generators. Surely this must be easy?

In fact it is not at all easy, and was proved by the following beautiful but indirect method (see [51]). First observe that the matrices in S, when regarded as integer matrices, generate $SL_2(\mathbb{Z})$. Now let $\Gamma(p)$ denote the congruence subgroup which is the kernel of the natural map from $SL_2(\mathbb{Z}) \to SL_2(p)$. If $\mathbb{H}$ is the upper half plane and $X(p)$ denotes the Riemann surface $\Gamma(p)\backslash\mathbb{H}$, denote by $\lambda_1(X(p))$ the smallest eigenvalue for the Laplacian on $X(p)$. A theorem of Selberg [61] gives $\lambda_1(X(p)) \geq \frac{3}{16}$ for all p, and this can be used to show that the Cayley graphs $\{\Gamma_p = \Gamma(SL_2(p), S) : p \text{ prime}\}$ have their second largest eigenvalues bounded away from the valency, and hence that they form a family of *expander graphs*. This means that there is an *expansion* constant $c > 0$, independent of p, such that for every set A consisting of fewer that half the total number of vertices in Γ_p, we have $|\delta A| > c|A|$, where δA is the boundary of A – that is, the set of vertices not in A that are joined to some vertex in A. From the expansion property it is easy to deduce that Γ_p has logarithmic diameter, so that $\mathrm{diam}(\Gamma(SL_2(p), S) < c \log p$, a strong form of Babai's conjecture.

One can adopt essentially the same method for the generators

$$\{\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}^{\pm 1}, \begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix}^{\pm 1}\}$$

of $SL_2(p)$, since, while these do not generate $SL_2(\mathbb{Z})$, they do generate a subgroup of finite index therein. But what if we replace the 2's in these generators with 3's? Then the matrices generate a subgroup of infinite index in $SL_2(\mathbb{Z})$, and the above method breaks down. This question became known as Lubotzky's 1-2-3 problem, and was not solved until the breakthrough achieved by Helfgott [23]:

Theorem 2.1 *Babai's conjecture holds for $G = SL_2(p)$. That is,*

$$\mathrm{diam}(SL_2(p)) < (\log p)^c,$$

where c is an absolute constant.

Helfgott deduced this from his key proposition: for any generating set S of $G = SL_2(p)$, either $|S^3| > |S|^{1+\epsilon}$, or $S^k = G$, where $\epsilon > 0$ and k do not depend on p. (Later it was observed that one can take $k = 3$ here.) The heart of his proof is to relate the growth of powers of subsets A of G with the growth of the corresponding set of scalars $B = \mathrm{tr}(A) = \{\mathrm{tr}(x) : x \in A\}$ in $\mathbb{F}_p$ under sums and products. By doing this he could tap into the theory of additive combinatorics, using results such as the following, taken from [10]: if B is a subset of $\mathbb{F}_p$ with $p^\delta < |B| < p^{1-\delta}$ for some $\delta > 0$, then $|B \cdot B| + |B + B| > |B|^{1+\epsilon}$, where $\epsilon > 0$ depends only on δ.

Following Helfgott's result, there was a tremendous surge of progress in this area. Many new families of expanders were constructed in [9]. Helfgott himself extended his result to $SL_3(p)$ in [24], and this has now been proved for all groups of Lie type of bounded rank in [11, 58]. As a consequence, we have

Theorem 2.2 *If $G = G(q)$ is a simple group of Lie type of rank r, then* $\mathrm{diam}(G) < (\log |G|)^{c(r)}$ *where $c(r)$ depends only on r.*

Again, the theorem is proved via a growth statement: for any generating set S of $G(q)$, either $|S^3| > |S|^{1+\epsilon}$, or $S^3 = G$, where $\epsilon > 0$ depends only on r. From this one gets a strong version of the previous result which takes the size of the generating set S into account:

Theorem 2.3 *If $G = G(q)$ is a simple group of Lie type of rank r, and S is a generating set of G, then $G = S^d$ for some $d \leq (\frac{\log |G|}{\log |S|})^{c(r)}$, where $c(r)$ depends only on r.*

These results, and particularly their developments into the theory of expanders, have many wonderful and surprising applications. For a survey of these developments and some of the applications, see [53].

Finally, let us remark that Babai's conjecture remains open for groups of Lie type of unbounded rank.

2.1.2 Alternating groups

For the alternating groups A_n, Babai's conjecture is that there is a constant C such that $\mathrm{diam}(A_n) < n^C$. Until very recently, the best bound for $\mathrm{diam}(A_n)$ was that obtained by Babai and Seress in [5], where it was proved that

$$\mathrm{diam}(A_n) < \exp((1 + o(1)) \cdot (n \log n)^{1/2}) = \exp((1 + o(1)) \cdot (\log |A_n|)^{1/2}).$$

Various other partial results appeared at regular intervals, such as that in [3], where it was shown that if the generating set S contains a permutation of degree at most $0.33n$, then $\mathrm{diam}(A_n, S)$ is polynomially bounded. But no real progress was made on Babai's conjecture until a recent breakthrough of Helfgott and Seress [25]:

Theorem 2.4 *We have* $\mathrm{diam}(A_n) \leq \exp(O((\log n)^4 \log \log n))$*, where the implied constant is absolute.*

This does not quite prove Babai's conjecture, but it does prove that $\mathrm{diam}(A_n)$ is "quasipolynomial" (where a quasipolynomial function $f(n)$ is one for which $\log f(n)$ is polynomial in $\log n$), which represents a big step forward. The same paper also gives a bound of the same magnitude for the diameter of any transitive subgroup of S_n.

2.2 Orbital graphs

Here we discuss another class of graphs for which the diameter has an interpretation in terms of width.

Denote by (G, X) a permutation group G on a finite set X. Suppose G is transitive on X, and let $X^{\{2\}}$ denote the set of unordered pairs of elements of X.

For each orbit Δ of G on $X^{\{2\}}$, there is a corresponding *orbital graph* having vertex set X and edge set Δ. These are precisely the non-empty graphs on X for which G acts transitively on edges. A well known criterion of D.G. Higman (see [26, 1.12]) states that G is primitive on X if and only if all of its orbital graphs are connected. For G primitive on X, define $\mathrm{diam}(G, X)$ to be the maximum of the diameters of all the orbital graphs.

The diameters of orbital graphs of primitive groups have an interpretation in terms of width. Indeed, let Δ be an orbit of G on $X^{\{2\}}$ as above, and let $\{x, xg\} \in \Delta$, where $g \in G$. Notice that also $\{x, xg^{-1}\} \in \Delta$. Write $H = G_x$. For each i, the set of vertices at distance i from x in the corresponding orbital graph is contained in

$$\{xg^{\pm 1}h_1g^{\pm 1}h_2 \cdots g^{\pm 1}h_i \,:\, h_i \in H\}.$$

It follows that if we define $w = \mathrm{width}(G, \mathcal{S})$ where $\mathcal{S} = H \cup \{g, g^{-1}\}$, then the diameter of the orbital graph lies between w and $[\frac{1}{2}w]$. (Both extremes are possible: for example the diameter is w when $H = 1$ and G is cyclic of prime order.)

For a positive integer d, denote by $\mathcal{C}_d$ the class of all finite primitive permutation groups (G, X) for which $\mathrm{diam}(G, X) \leq d$. In [42], the following problem is addressed.

Problem 2.5 For each d, describe the class of finite primitive groups $\mathcal{C}_d$.

The motivation in [42] is mainly model-theoretical and stems from the fact that for groups of bounded orbital diameter, primitivity is implied by a first order expressible condition in the language of permutation groups (whereas for permutation groups in general, primitivity is not a first order property). This means, for example, that the primitivity condition extends to ultraproducts.

In [42], the above problem is solved "asymptotically"; as discussed in detail in [42], this leads to the solution of a number of related model-theoretic problems, such as the description of primitive infinite ultraproducts of finite permutation groups, and of primitive ω-saturated pseudofinite permutation groups.

We present part of the main result of [42] in Theorem 2.6 below, which describes the classes of *simple* groups in $\mathcal{C}_d$. This time, unlike the previous section, there is a satisfactory result for groups of unbounded rank.

In order to state the theorem we need to define some terminology. We say that the primitive group (G, X) with G simple is a *standard t-action* if one of the following holds:

(a) $G = A_n$ and $X = I^{\{t\}}$, the set of t-subsets of $I = \{1, ..., n\}$

(b) $G = Cl_n(q)$, a classical group with natural module $V = V_n(q)$ of dimension n over $\mathbb{F}_q$, and X is an orbit of subspaces of dimension or codimension t in V; the subspaces are arbitrary if $G = PSL_n(q)$, and otherwise are totally singular, non-degenerate, or, if G is orthogonal and q is even, non-singular 1-spaces (in which case $t = 1$)

(c) $G = Sp_{2m}(q)$, q is even, and a point stabilizer in G is $O^{\pm}_{2m}(q)$ (here we take $t = 1$).

If $G(q)$ is a simple group of Lie type over $\mathbb{F}_q$, then a *subfield subgroup* is a group $G(q_0)$ embedded naturally in $G(q)$, where $\mathbb{F}_{q_0}$ is a subfield of $\mathbb{F}_q$. For convenience in the statement below we define the *rank* of an alternating group A_n to be n.

We say that a class $\mathcal{C}$ of finite primitive permutation groups is *bounded* if $\mathcal{C} \subseteq \mathcal{C}_d$ for some d. All bounds implicit in the statement below are in terms of d, where $\mathcal{C} \subseteq \mathcal{C}_d$.

Theorem 2.6 *Let $\mathcal{C}$ be an infinite class of finite simple primitive permutation groups, and suppose $\mathcal{C}$ is bounded.*

(i) *If $\mathcal{C}$ consists of simple groups of unbounded ranks, then the groups in $\mathcal{C}$ of sufficiently large rank are alternating or classical groups in standard t-actions, where t is bounded.*

(ii) *If $\mathcal{C}$ consists of simple groups G of bounded rank, then point stabilizers G_x have unbounded orders; moreover, if $G = G(q)$, of Lie type over $\mathbb{F}_q$, and G_x is a subfield subgroup $G(q_0)$, then $|\mathbb{F}_q : \mathbb{F}_{q_0}|$ is bounded.*

Conversely, any class of simple primitive groups satisfying the conclusions of (i) *or* (ii) *is bounded.*

One of the most interesting parts of this result is the converse statement for part (ii): if $\mathcal{C}$ is a class consisting of simple primitive permutation groups of Lie type of bounded Lie rank with unbounded point stabilizers (and also satisfying the given condition on subfields), then $\mathcal{C}$ is a bounded class. For example, if $\mathcal{C}$ consists of the groups $E_8(q)$ (q varying) acting on the coset space $E_8(q)/H(q)$ for some maximal subgroup $H(q)$ arising from a maximal connected subgroup $H(K)$ of the simple algebraic group $E_8(K)$, where $K = \bar{\mathbb{F}}_q$ (for example $H(K) = D_8(K)$ or $A_1(K)$ – see [47]), then the diameters of all the orbital graphs are bounded by an absolute constant. In fact this now follows fom Theorem 2.3, but a direct proof using a substantial amount of model theory can be found in [42].

It would be interesting to have a more explicit solution to Problem 2.5, for example for some small values of d. Work is under way on this.

3 Conjugacy width

We now turn to a discussion of the width of simple groups G with respect to a conjugacy class of subsets – that is, $\mathrm{width}(G, \mathcal{S})$ where $\mathcal{S} = \{A^g : g \in G\}$ for some subset A of G which we take to be of size at least 2. The following lemma shows that in this case no subproducts are required in the definition of width.

Lemma 3.1 *If $A \subseteq G$ with $|A| \geq 2$, and $\mathcal{S} = \{A^g : g \in G\}$, then*

$$\mathrm{width}(G, \mathcal{S}) = \min\{n : G = A^{g_1} \cdots A^{g_n}, g_i \in G\}.$$

Proof This is clear if $1 \in A$. If not, let $a \in A$, set $B = a^{-1}A$, and observe that G is a product of n conjugates of A if and only if it is a product of n conjugates of B. ∎

Examples When G is simple there are many interesting cases to consider. Here are some examples. In the first four, $\mathcal{S}$ consists of a single normal subset of G (i.e. a subset closed under conjugation), so we are back in the Cayley graph case of the previous section.

1. $\mathcal{S} = \{I(G)\}$, where $I(G)$ is the set of involutions in G: here $\text{width}(G, \mathcal{S})$ is the minimal n such that every element of G is a product of n involutions.
2. $\mathcal{S} = \{C(G)\}$, where $C(G) = \{[x, y] : x, y \in G\}$ is the set of commutators in G: here $\text{width}(G, \mathcal{S})$ is often called the *commutator width* of G.
3. $\mathcal{S} = \{P_k(G)\}$, where $k \geq 2$ and $P_k(G) = \{x^k : x \in G\}$ is the set of k^{th} powers in G.
4. (Generalizing Examples 2,3): $\mathcal{S} = \{w(G)\}$, where $w = w(x_1, \ldots, x_k)$ is a fixed word in the free group F_k of rank k and $w(G) = \{w(g_1, \ldots, g_k) : g_i \in G\}$.
5. $G = S_n$ and $\mathcal{S} = \{\langle t_1 \rangle, \ldots, \langle t_k \rangle\}$, where $t_1, \ldots t_k$ are all the transpositions in G (and $k = \binom{n}{2}$), as in Example 2 in Section 1.
6. $\mathcal{S} =$ the set of Sylow p-subgroups of G, where p is a prime dividing $|G|$.

Clearly if $\mathcal{S} = \{A^g : g \in G\}$ as above, then $\text{width}(G, \mathcal{S}) \geq \log |G| / \log |A|$. In [43] the following conjecture was posed.

Conjecture 3.2 *There is an absolute constant c such that for any finite non-abelian simple group G and any subset $A \subseteq G$ with $|A| \geq 2$, we have*

$$\text{width}(G, \mathcal{S}) \leq c \frac{\log |G|}{\log |A|},$$

where $\mathcal{S} = \{A^g : g \in G\}$.

This conjecture has been proved in a number of special cases, as we shall describe below, but it is open in general.

3.1 Normal subsets

In the case where $\mathcal{S}$ consists of a single normal subset of G, Conjecture 3.2 was proved in [50]:

Theorem 3.3 *There is an absolute constant $k > 0$ such that for any finite non-abelian simple group G, and any non-identity normal subset $S \subseteq G$, we have $G = S^n$ for all $n \geq k \log |G| / \log |S|$.*

In particular the diameter of the Cayley graph $\Gamma(G, S)$ is at most $k \frac{\log |G|}{\log |S|}$, so this proves Babai's conjecture in this case in a strong form.

The *covering number* of a finite simple group G is the minimal positive integer n such that $C^n = G$ for all conjugacy classes C of G (see [2]). Theorem 3.3 implies an upper bound for the covering number which is linear in the rank of G; further such bounds can be found in [14, 39], and the precise covering number of $PSL_n(q)$

for $n \geq 3, q \geq 4$ is shown to be n in [40]. However Theorem 3.3 carries much more information than these bounds, since it takes into account the size of the class.

Let us now examine the implications of Theorem 3.3 for Examples 1–4 above.

3.1.1 Involutions

As in Example 1 above, let $S = I(G)$, the set of involutions in G. To get a feeling for how big $\frac{\log |G|}{\log |S|}$ is, consider $G = PSL_{2m}(q)$ with q odd, m even, and let $t \in G$ be the involution which is the image modulo scalars of the matrix $\mathrm{diag}(I_m, -I_m)$. Then the size of the conjugacy class t^G is roughly $|GL_{2m}(q) : GL_m(q) \times GL_m(q)|$, which is approximately q^{4m^2}/q^{2m^2}, and so $|t^G|$ is of the order of $|G|^{1/2}$. Therefore $\log |G| / \log |S|$ is about 2 in this case. It can be shown that there is an absolute constant $c > 0$ such that $|I(G)| > c|G|^{1/2}$ for all finite simple groups G (see [49, 4.2,4.3]). Hence Theorem 3.3 implies the following.

Corollary 3.4 *There is an absolute constant N such that every element of every finite non-abelian simple group is a product of N involutions.*

It would be quite interesting to know the minimal value of N. It is certainly more than 2: groups in which every element is a product of two involutions are known as *strongly real* groups, and the strongly real simple groups have been classified (see [64, 59]).

3.1.2 Images of word maps

As in Example 4 above, let $w = w(x_1, \ldots, x_k)$ be a fixed non-identity word in the free group F_k of rank k and for a group G define $w(G) = \{w(g_1, \ldots, g_k) : g_i \in G\}$. Let us consider the implications of Theorem 3.3 in the case where G is simple and $S = w(G)$.

We need information about the size of the set $w(G)$. This can be 1 for some simple groups G – for example if $w = x_1^k$ and the exponent of G divides k. The first question to consider is whether there could be a word w for which $w(G) = \{1\}$ for *all* (finite non-abelian) simple groups G. The answer is no: for suppose w is a non-identity word such that $w(SL_2(p)) = \{1\}$ for all primes p. Let ϕ_p be the natural map $SL_2(\mathbb{Z}) \to SL_2(p)$. Then $\bigcap_p \mathrm{Ker}(\phi_p) = 1$, hence also $w(SL_2(\mathbb{Z})) = 1$. However $SL_2(\mathbb{Z})$ contains a free subgroup of rank 2, so this is impossible. Since many simple groups of Lie type over $\mathbb{F}_p$ contain $SL_2(p)$, the assertion follows.

In fact a much stronger assertion about the nontriviality of $w(G)$ for simple groups G holds, as proved in [30]:

Theorem 3.5 *Given any nontrivial word w, there is a constant N_w depending only on w, such that $w(G) \neq \{1\}$ for all simple groups G of order greater than N_w.*

For simple groups of order greater than N_w, how large is $w(G)$? The following gives a weak lower bound. Better bounds will be discussed in Section 4.

Lemma 3.6 *For any non-identity word w, there is a constant $\delta_w > 0$ such that $|w(G)| > |G|^{\delta_w}$ for all simple groups G of order greater than N_w.*

Proof Consider first $G = A_n$. Choose $k = k(w)$ minimal such that $w(A_k) \neq 1$, and let $1 \neq a \in w(A_k)$. Take n to be large in terms of k. If $r = [\frac{n}{k}]$, then G contains a subgroup $H \cong (A_k)^r$. Let $x \in H$ be the image under this isomorphism of the element $(a, \ldots, a) \in A_k^r$. Then $x \in w(H)$ and x moves at least $3r$ points in $\{1, \ldots, n\}$. Now the conjugacy class x^G is contained in $w(G)$, and an elementary calculation shows that $|x^G|$ is at least of the order of $|G|^{1/2k}$, which gives the conclusion in this case.

The case where $G = Cl_n(q)$, a classical group of unbounded dimension n over a finite field $\mathbb{F}_q$, is similar, using a subgroup H of the form $(Cl_k(q))^r$ in the above argument. And when G is a group of Lie type of bounded rank, the fact that any nontrivial conjugacy class has size at least q gives the result. ∎

As before, Theorem 3.3 implies the following consequence.

Corollary 3.7 *Let w be a nontrivial word. Then there is a constant $c = c(w)$ such that for any simple group G of order greater than N_w, we have $G = w(G)^c$ (that is, every element of G is a product of c elements of $w(G)$).*

We shall discuss some recent vast improvements of this result in Section 4.

3.1.3 Remarks on the proof of Theorem 3.3

The proof in [50] is quite technical, but it may be instructive to illustrate two of the main steps with the following example. Let $G = PSL_n(q)$ with $n \geq 3$ and let $C = x^G$, where

$$x = \begin{pmatrix} J_k & \\ & I_{n-k} \end{pmatrix},$$

J_k being the $k \times k$ Jordan block matrix with 1's on and directly above the diagonal and 0's elsewhere. Assume also that n is large compared to k. The centralizer of x can be found in [48, 7.1], and it follows that $|C|$ is roughly $q^{(k-1)(2n-k)}$. Hence $\frac{\log |G|}{\log |C|}$ is of the order of $\frac{n}{2(k-1)}$.

The first step in the proof is the elementary but useful observation that

$$\begin{pmatrix} I_{k-1} & & \\ & J_k & \\ & & I_{n-2k+1} \end{pmatrix} \begin{pmatrix} J_k & \\ & I_{n-k} \end{pmatrix} = \begin{pmatrix} J_{2k-1} & \\ & I_{n-2k+1} \end{pmatrix}.$$

Applying this repeatedly, we can obtain the matrix J_n as a product of approximately $\frac{n}{k-1}$ conjugates of x; in other words, $J_n \in C^{n/(k-1)}$. Set $y := J_n$.

The second step is to apply some character theory of the group G. The following observation essentially goes back to Frobenius, and applies to conjugacy classes in

arbitrary finite groups: for $g \in G$, and an integer $l \geq 2$, the number of ways of writing g as a product of l conjugates of y is

$$\frac{|y^G|^l}{|G|} \sum_{\chi \in Irr(G)} \frac{\chi(y)^l \chi(g^{-1})}{\chi(1)^{l-1}}, \tag{2}$$

where $\mathrm{Irr}(G)$ denotes the set of irreducible characters of G. At this point we apply some basic facts about the irreducible characters χ of $G = PSL_n(q)$:

(a) $|\chi(y)| \leq |C_G(y)|^{1/2} = |C_G(J_n)|^{1/2} \leq q^{n/2}$;

(b) for $\chi \neq 1_G$, the degree $\chi(1) \geq q^{n-1} - 1$;

(c) $|\mathrm{Irr}(G)| < q^{n-1} + 3q^{n-2}$.

Indeed, (a) is trivial, (b) follows from [33] and (c) from [16, 3.6]. Let Σ denote the sum in (2). The contribution to Σ of the trivial character $\chi = 1_G$ is 1. Hence using (a)–(c), we see that

$$|\Sigma| \geq 1 - \frac{(q^{n-1} + 3q^{n-2})q^{nl/2}}{(q^{n-1} - 1)^{l-2}}.$$

Assuming that $n \geq 10$, it follows that $\Sigma \neq 0$ provided $l \geq 7$. Hence $G = (y^G)^7$ under this assumption. Since $y = J_n \in C^{n/(k-1)}$, we therefore have

$$G = (y^G)^7 = C^{7n/(k-1)}.$$

The conclusion of Theorem 3.3 follows in this case.

3.1.4 Commutators

Applying Corollary 3.7 to the commutator word, it follows that every element of every finite simple group is a product of a bounded number of commutators. In fact a much stronger result is true:

Theorem 3.8 (The Ore Conjecture) *Every element of every finite simple group is a commutator.*

This conjecture emerged from a 1951 paper of Ore [56], after which many partial results were obtained, notably those of Thompson [63] for special linear groups, and of Ellers and Gordeev [13] proving the result for groups of Lie type over sufficiently large fields $\mathbb{F}_q$ ($q \geq 8$ suffices). The proof was finally completed in [44]. This was largely based on character theory, via an elementary classical result, again due to Frobenius, that for an element g of a finite group G, the number of solutions $(x, y) \in G \times G$ to the equation $g = [x, y]$ is equal to

$$|G| \sum_{\chi \in Irr(G)} \frac{\chi(g)}{\chi(1)}.$$

Thus g is a commutator if and only if this sum is nonzero. The aim is to show that for G simple, the term coming from the trivial character (namely 1) is greater than the sum of moduli the remaining terms, in other words that

$$\sum_{\chi \neq 1_G} \frac{|\chi(g)|}{\chi(1)} < 1. \tag{3}$$

Here is a sketch of the proof from [44] of Theorem 3.8 for the family of symplectic groups $G = Sp_{2n}(2)$. The argument proceeds by induction. The base cases for the induction are $Sp_{2n}(2)$ with $n \leq 6$, and these were handled computationally; of course $Sp_2(2)$ and $Sp_4(2)$ are non-perfect, so Theorem 3.8 does not apply to them.

Let $g \in G$, and write g in block-diagonal form

$$g = \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & X_k \end{pmatrix} \in Sp_{2n_1}(2) \times \cdots \times Sp_{2n_k}(2) < G,$$

where $\sum n_i = n$, this decomposition being as refined as possible. If each X_i is a commutator in $Sp_{2n_i}(2)$ then g is a commutator in G. Hence induction gives the conclusion except when either

(1) $k = 1$, or

(2) one of the factors $Sp_{2n_i}(2)$ is $Sp_2(2)$ or $Sp_4(2)$.

We call g *unbreakable* if (1) or (2) holds for every such block-diagonal decomposition of g. Thus to prove the theorem for this case it suffices to show that every unbreakable element g of $G = Sp_{2n}(2)$ with $n \geq 7$ is a commutator.

The first step is to prove that the unbreakable element g has small centralizer, namely

$$|C_G(g)| < 2^{2n+15}.$$

For example, if g is unipotent its unbreakability means that it can have few Jordan blocks, and the possiblities for the centralizers of such elements are given by [48, Chapter 7].

Next, a result of Guralnick and Tiep [21] shows that there is a collection $\mathcal{W}$ of 5 irreducible characters of G such that

(i) $\chi(1) \geq \frac{1}{6}(2^n - 1)(2^n - 2)$ for $\chi \in \mathcal{W}$, and

(ii) $\chi(1) \geq 2^{4n-7}$ for $1 \neq \chi \in \mathrm{Irr}(G) \setminus \mathcal{W}$.

Set

$$\Sigma_1(g) = \sum_{\chi \in \mathcal{W}} \frac{|\chi(g)|}{\chi(1)}, \quad \Sigma_2(g) = \sum_{1 \neq \chi \in \mathrm{Irr}(G) \setminus \mathcal{W}} \frac{|\chi(g)|}{\chi(1)}.$$

Letting $k(G)$ denote the number of conjugacy classes of G, it follows from [16, 3.13] that $k(G) \leq (15.2) \cdot 2^n$. Also $\sum_{\chi \in Irr(G)} |\chi(g)|^2 = |C_G(g)|$ by the orthogonailty relations, from which the Cauchy-Schwartz inequality implies that

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

Plugging all this into the expression defining $\Sigma_2(g)$, we obtain

$$\Sigma_2(g) < \frac{\sqrt{15.2} \cdot 2^{n/2} \cdot |C_G(g)|^{1/2}}{2^{4n-7}} < \frac{\sqrt{15.2} \cdot 2^{n/2} \cdot 2^{n+7.5}}{2^{4n-7}} < 0.5.$$

Bounding $\Sigma_1(g)$ depends on some detailed analysis of the values $\chi(g)$ for the characters $\chi \in \mathcal{W}$, from which one shows that $\Sigma_1(g) < 0.2$.

Hence $\Sigma_1(g) + \Sigma_2(g) < 0.7$, which implies that (3) holds, and hence g is a commutator, as required.

This example gives the flavour of the proof of Theorem 3.8, but it must be said that other families of classical groups over small fields do not yield so easily as this. Indeed the unitary groups presented too many technical obstacles for us to handle them in this fashion, and we used a completely different method for these.

3.2 Bounded subsets

Conjecture 3.2 has been proved for bounded subsets in [43, Theorem 3]:

Theorem 3.9 *There is an absolute constant c such that if G is a finite non-abelian simple group, and A is any subset of G of size at least 2, then G is a product of N conjugates of A for some $N \leq c \log |G|$.*

We shall sketch a proof of this result for alternating groups, and refer the reader to [43] for the rest of the proof. Suppose then that $G = A_n$.

First we claim that, in proving the conjecture for a subset A, we may assume that $1 \in A$. Indeed, let $a \in A$ and $B = a^{-1}A$. Then $1 \in B$, and if G is a product of N conjugates of B then it is also a product of N conjugates of A. Secondly, we claim we may assume there exists $x \neq 1$ such that $1, x, x^{-1} \in A$. Indeed, suppose $1 \in A$ and let $x \in A$ be a non-identity element (whose existence follows from the assumption $|A| \geq 2$). Then $1, x, x^2 \in A^2$, hence $x^{-1}, 1, x \in x^{-1}A^2$. Assuming the conjecture holds for sets containing $x^{-1}, 1, x$ we deduce that G is a product of say $N \leq c \log |G| / \log |A^2| \leq c \log |G| / \log |A|$ conjugates of $x^{-1}A^2$, hence it is a product of N conjugates of A^2, so G is a product of $2N \leq 2c \log |G| / \log |A|$ conjugates of A.

So assume that $1, x, x^{-1} \in A \subseteq G$ for some $x \neq 1$. It is easy to choose a 3-cycle $y \in A_n$ such that $[x, y] \neq 1$ has support of size at most 5. Let $C = x^{A_n}$, the conjugacy class of x. Since $[x, y] = x^{-1}x^y \in C^{-1}C$, we see that $C^{-1}C$ contains either a 3-cycle, a 5-cycle or a double transposition. In all cases we deduce that $(C^{-1}C)^2$ contains all double transpositions in A_n. Since $x, x^{-1} \in A$, some product of 4 conjugates of A contains $\{1, t\}$ for a double transposition $t \in A_n$.

At this point a straightforward argument shows that it is sufficient to establish the result for the subset $\{1, \tau\}$ of S_{n-2}, where τ is a transpostion – in other words, that S_{n-2} is a product of $cn \log n$ conjugates of $T := \{1, \tau\}$ (this is Example 2 in Section 1).

This is not as obvious as it might seem. The key to it is a lemma of Abert [1, Lemma 4]: for positive integers a, b, we have $S_{ab} = ABA$, where A is a conjugate

of the natural subgroup $(S_a)^b$ and B is a conjugate of $(S_b)^a$. For notational convenience, replace $n-2$ by n, and let 2^l be the largest power of 2 that is less than or equal to n. Then $\frac{n}{2} < 2^l \leq n$. Repeated application of Abert's lemma shows that S_{2^l} is a product of $2l-1$ conjugates of $(S_2)^{2^{l-1}}$, hence of $(2l-1)2^{l-1}$ conjugates of T. Since it is routine to see that for $\frac{n}{2} < k \leq n$, S_n is a product of at most 8 conjugates of S_k, it follows that S_n is a product of at most $(2l-1)2^{l+2}$ conjugates of T, and the conclusion follows.

3.3 Bounded rank

Conjecture 3.2 has also been proved for simple groups of Lie type of bounded rank, in [18, Theorem 1.3]:

Theorem 3.10 *Fix a positive integer r. There exists a constant $c = c(r)$ such that if G is a finite simple group of Lie type of rank r and A is a subset of G of size at least 2, then G is a product of N conjugates of A for some $N \leq c \log |G| / \log |A|$.*

It is possible to get some of the way towards this result quite quickly, as follows. Firstly, as observed in the sketch proof of Theorem 3.9 above, we can assume that $1 \in A$. Next, by a result in [22], for $1 \neq x \in A$, there are $m \leq 8(2r+1)$ conjugates of x that generate G; call them $x^{g_1}, \ldots, x^{g_m}$. Write $S = A^{g_1} \cdots A^{g_m}$. Then S generates G, so by the Product Theorem 2.3, $G = S^d$ for some $d \leq (\frac{\log |G|}{\log |S|})^{c(r)}$, and hence G is a product of $(\frac{\log |G|}{\log |S|})^{c_1(r)}$ conjugates of A.

Getting rid of the exponent $c_1(r)$ takes a lot more effort, and this is the main content of [18]. Along the way, they prove an interesting growth result for conjugates ([18, 1.4]): for G and A as in the theorem above, either $A^3 = G$ or there exists $g \in G$ such that $|AA^g| > |A|^{1+\epsilon}$, where $\epsilon > 0$ depends only on the rank r.

3.4 Sylow subgroups

The width of simple groups with respect to a class of Sylow p-subgroups has only been addressed in the case of groups of Lie type, where p is the natural characteristic.

Theorem 3.11 *If G is a simple group of Lie type over a field of characteristic p, then G is a product of 5 Sylow p-subgroups.*

This was first proved in [46] with a bound of 25 instead of 5; the improvement to 5 was announced in [4]. The proof in [46] uses the BN-structure of G, and shows that if $U \in \mathrm{Syl}_p(G)$ is the unipotent radical of a Borel subgroup B, and V is the unipotent radical of the opposite Borel, then $G = UVUV \cdots VU$ (25 terms). The reduction to 5 terms was achieved by using what has become known as the "Gowers trick", a very useful tool in the theory of width:

Proposition 3.12 *Let $n > 2$ be an integer and let G be a finite group and let k be the minimal degree of a nontrivial complex character of G. Suppose that $A_i \subseteq G$, $i = 1, 2, \ldots, n$ are such that $\frac{|A_i|}{|G|} \geq k^{-(n-2)/n}$. Then $G = A_1 \cdot A_2 \cdots A_n$.*

This can often be used when G is a group of Lie type, since these have relatively large minimal nontrivial character degrees (see [33]).

This result has an application to the width of finite linear groups. The starting point is an elegant result of Hrushovski and Pillay [27], proved using model theory (and not using the classification of finite simple groups):

Theorem 3.13 *Let p be a prime, n a positive integer, and suppose G is a subgroup of $GL_n(p)$ that is generated by elements of order p. Then $G = \langle x_1 \rangle \langle x_2 \rangle \cdots \langle x_k \rangle$ for some elements x_i of order p, where $k = k(n)$ depends only on n.*

Note that the result is trivial if p is bounded in terms of n. It was generalized as follows in [46]:

Theorem 3.14 *There is a function $f : \mathbb{N} \to \mathbb{N}$ such that the following holds. Let n be a positive integer, p a prime with $p \geq f(n)$, and F a field of characteristic p. If G is a finite subgroup of $GL_n(F)$ generated by elements of order p, then G is a product of 5 of its Sylow p-subgroups.*

Again this is proved without the classification, but using the marvellous theorem of Larsen and Pink [34] as a substitute: if S is a finite simple subgroup of $GL_n(F)$, where F is a field of characteristic p, then either S is of Lie type in characteristic p, or $|S|$ is bounded in terms of n. Bounds for the function $f(n)$ in the above theorem are not addressed in [46], but using the classification Guralnick [19] showed that $f(n) = n + 3$ works; this is best possible, as can be seen from the example of the alternating group $A_p < GL_{p-2}(p)$ (via the action on the fully deleted permutation module for A_p over $\mathbb{F}_p$) – clearly A_p is not a product of a bounded number of its Sylow p-subgroups.

4 Word maps

In this section we develop further the theory of word maps on simple groups, introduced in Section 3.1.2. Let $w = w(x_1, \ldots, x_k)$ be a nontrivial word in the free group F_k of rank k, and for a group G, denote also by $w : G^k \to G$ the *word map* sending $(g_1, \ldots, g_k) \to w(g_1, \ldots, g_k)$ for $g_i \in G$. Write $w(G)$ for the image of this map.

We shall focus on word maps on finite (non-abelian) simple groups G. Recall from Theorem 3.5 that there is a constant N_w such that $w(G) \neq \{1\}$ for simple groups G with $|G| > N_w$.

Questions Here are a few natural questions one might ask about word maps:

1. How large is $w(G)$? Previously we saw in Lemma 3.6 that $|w(G)| > |G|^{\delta_w}$ for some $\delta_w > 0$ depending only on w. Can one do better than this?

2. What is the w-width of G, i.e. the width of G with respect to $w(G)$? We saw in Corollary 3.7 that it is bounded above by a constant $c(w)$. Is it possible to improve this?
3. For $g \in G$, define $P_w(g)$ to be the probability that $w(g_1, \ldots, g_k) = g$ for $g_i \in G$ chosen uniformly at random; so

$$P_w(g) = \frac{|w^{-1}(g)|}{|G|^k}.$$

   What can one say about the probability distribution P_w on G? Is it always close to the uniform distribution? Or are there words w for which P_w is highly non-uniform?
4. Regarding Question 3, consider for example $G = SL_2(p)$ with p prime. The proportion of elements of order p in G is precisely $\frac{1}{p}$, so one cannot design an algorithm in computational group theory that is based on finding an element of order p in G by random search. But can one find a fiendishly clever word w for which $\sum_{g \in C} P_w(g) >> \frac{1}{p}$, where C is the set of elements of order p? Such a word would be very interesting computationally.

4.1 Size

Sometimes $w(G) = G$ for all simple groups G – for example for the commutator word $w = [x_1, x_2]$, by the Ore Conjecture (Theorem 3.8); and sometimes $w(G) \neq G$ – for example for $w = x_1^2$, or any power word $w = x_1^k$ for which $\mathrm{hcf}(k, |G|) \neq 1$. Nevertheless, the following result of Larsen and Shalev [36, 2.1 and 1.10] shows that images of word maps on simple groups are always large:

Theorem 4.1 *Let w be a nontrivial word and r a positive integer. There exist positive constants $N(w)$ and $c(r)$ depending only on w and r respectively, such that the following hold.*

(i) *If G is a simple group of Lie type of rank at most r, then $|w(G)| > c(r)|G|$ provided $|G| > N(w)$.*
(ii) *If G is an alternating group A_n, then $|w(G)| > n^{-4}|G|$ provided $n > N(w)$.*

In fact a result stronger than (i) is proved in [36, 1.12]: one can take $c(r) = cr^{-1}$ for some absolute constant c, provided G is not of type PSL or PSU.

There are some interesting tools used in the proof of the above theorem. For (i), a crucial ingredient is a result of Borel [8], which states that if $G = G(q)$ is of Lie type over $\mathbb{F}_q$, and $\bar{G} = G(\bar{\mathbb{F}}_q)$ is the corresponding simple algebraic group over the algebraic closure $\bar{\mathbb{F}}_q$, then the word map $w : \bar{G}^k \to \bar{G}$ is dominant, which is to say that it has dense image. Further arguments from algebraic geometry are used to deduce part (i).

The proof of part (ii) involves a neat application of the celebrated result of Vinogradov [65] that every sufficiently large odd integer is a sum of three primes. So let n be large, and write $n = p_1 + p_2 + p_3 + 3 + \delta$ with p_i primes and $\delta \in \{0, 1\}$.

The group $L_i := PSL_2(p_i)$ has a 2-transitive action of degree $p_i + 1$, so we can embed $L_1 \times L_2 \times L_3 < A_n$ in a natural way. A by-product of the proof of part (i) is that $w(L_i)$ contains an element x_i of order $\frac{p_i-1}{2}$, and x_i acts in the degree $p_1 + 1$ representation as a product of two cycles of length $\frac{p_i-1}{2}$ and two fixed points. Hence $x := x_1x_2x_3 \in w(A_n)$ has 6 long cycles and 6 or 7 fixed points. Then $|C_{A_n}(x)|$ is of the order of n^6, which shows that $|w(A_n)|$ is at least of the order of $n^{-6}|A_n|$. Improving the exponent to -4 (in fact to $-29/9$ in [36, 1.10]) takes more work.

There are some related results that should be mentioned here, which show that if one omits the condition that G is sufficiently large in terms of w in the above theorem, then $w(G)$ can be an arbitrary subset of G subject to the obvious necessary condition that it contains the identity and is invariant under $\mathrm{Aut}(G)$. Indeed, in [52], Lubotzky proves:

Theorem 4.2 *Let G be a finite non-abelian simple group, and let A be a subset of G such that $1 \in A$ and A is invariant under $\mathrm{Aut}(G)$. Then there is a word $w = w(x_1, x_2)$ in the free group of rank 2 such that $w(G) = A$.*

Explicit constructions of such words can be found in [31], and further results of this type in [41].

4.2 Width

Recall that for a word w and a simple group G such that $w(G) \neq 1$, the *w-width* of G is the width of G with respect to $w(G)$. A rather crude bound for w-width was given in Corollary 3.7. Can this be improved?

We pointed out at the beginning of the last section that this width is greater than 1 if w is a power word x_1^k. Hence the following remarkable result, the culmination of several papers of Shalev together with Larsen and Tiep [35, 36, 38, 62], is the best possible one of its kind.

Theorem 4.3 *For any nontrivial word w there is a constant N_w such that $w(G)^2 = G$ for all finite non-abelian simple groups G of order greater than N_w.*

Thus the w-width of all sufficiently large simple groups is at most 2. The proof that it is at most 3, originally a result in [62], was simplified for groups of Lie type in [55] using the Gowers trick (Proposition 3.12). Here is their idea in the bounded rank case. Proposition 3.12 with $n = 3$ implies that if G is a finite group with minimal nontrivial character degree k, and $A \subseteq G$ with $|A| \geq k^{-1/3}|G|$, then $G = A^3$. Letting $G = G(q)$ be a simple group of Lie type of rank r over $\mathbb{F}_q$, we have $k \geq aq^r$ for some positive absolute constant a by [33]. Fixing r, we have $|w(G)| > (aq^r)^{-1/3}|G|$ for sufficiently large q by Theorem 4.1(i), and hence $G = w(G)^3$, giving the claimed result for groups of bounded rank.

The problem of determining w-width was termed the "Waring problem" for simple groups by Shalev, by analogy with the celebrated Waring problem in number theory: this concerns the determination of the function $g : \mathbb{N} \to \mathbb{N}$, where $g(k)$ is

defined to be minimal such that every positive integer is the sum of $g(k)$ k^{th} powers. (So $g(k)$ could be thought of as the additive width of $\mathbb{N}$ with respect to the set of k^{th} powers.)

In direct analogy with Waring's problem, let us consider the width of the power word x_1^k for simple groups G, where $k \geq 2$. By Theorem 4.3, the width is 2 for sufficiently large G. But this is not the case for *all* G – for example the word x_1^{30} is trivial on A_5. For which values of k could the width be 2 for all simple groups G? Clearly not when k is the exponent of a simple group. An obvious family of positive integers that are not equal to the exponent of a simple group are those which are divisible by at most two primes (by Burnside's $p^a q^b$ theorem). For such integers we have the following result from [20]:

Theorem 4.4 *Let p, q be primes and a, b positive integers, and let $N = p^a q^b$. Then the word map $(x, y) \to x^N y^N$ is surjective on all finite (non-abelian) simple groups.*

4.3 Surjective and non-surjective words

If w has width 1 on G (i.e. $w(G) = G$), we call w a surjective word on G. Some words are surjective on *all* groups: these are precisely the words w in the free group F_k such that $w \in x^{e_1} \cdots x_k^{e_k} F_k'$, where $e_1, \ldots, e_k$ are integers with highest common factor 1 (see [60, 3.1.1]).

We have already observed that there are words that are non-surjective on finite simple groups, such as power words x_1^r. On the other hand, there are various special words that have been proved to be surjective on all finite simple groups: these include the commutator word (Theorem 3.8) and the word $x_1^N x_2^N$ for $N = p^a q^b$ (Theorem 4.4).

Could it be that the only words that are non-surjective on large simple groups are power words of the form $w = v^m$ $(m \geq 2)$? An affirmative answer was stated as a conjecture in [7, 7.14]. However it is not the case:

Theorem 4.5 *Define the word*

$$w = x_1^2 [x_1^{-2}, x_2^{-1}]^2 \in F_2.$$

Then the word map $(x, y) \to w(x, y)$ is non-surjective on $PSL_2(p^{2r+1})$ for all non-negative integers r and all odd primes $p \neq 5$ such that $p^2 \not\equiv 1 \bmod 16$ and $p^2 \not\equiv 1 \bmod 5$.

For example, w is non-surjective on $PSL_2(3^{2r+1})$ for all r.

This result was proved in [29], as part of a non-surjectivity theorem for the family of words of the form $x_1^2 [x_1^{-2}, x_2^{-1}]^k$ with $2k + 1$ prime.

Here is a sketch of the proof of Theorem 4.5. Let $G = SL_2(K)$ with K a field. The starting point is the observation, going back to Fricke and Klein (see [15]) that for any word $w = w(x_1, x_2)$, there is a polynomial $P_w(s, t, u)$ such that for all $x, y \in G$,

$$\mathrm{Tr}(w(x, y)) = P_w(\mathrm{Tr}(x), \mathrm{Tr}(y), \mathrm{Tr}(xy)).$$

We call P_w the *trace* polynomial of w. A proof of this fact, providing a constructive method of computing P_w for a given word w, can be found in [57, 2.2]. The method is based on the following identities for traces of 2×2 matrices A, B of determinant 1:

(1) $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$
(2) $\mathrm{Tr}(A^{-1}) = \mathrm{Tr}(A)$
(3) $\mathrm{Tr}(A^2B) = \mathrm{Tr}(A)\mathrm{Tr}(AB) - \mathrm{Tr}(B)$.

As an example, let us compute P_c for the commutator word $c = [x_1, x_2]$. First observe that

$$\begin{aligned}\mathrm{Tr}(x^2y^2) &= \mathrm{Tr}(x)\mathrm{Tr}(xy^2) - \mathrm{Tr}(y^2) \ \ ((\text{by } (3))\\ &= \mathrm{Tr}(x)(\mathrm{Tr}(y)\mathrm{Tr}(yx) - \mathrm{Tr}(x)) - \mathrm{Tr}(y)^2 + 2\\ &= stu - s^2 - t^2 + 2,\end{aligned}$$

where $s = \mathrm{Tr}(x), t = \mathrm{Tr}(y), u = \mathrm{Tr}(xy)$. Hence

$$\begin{aligned}\mathrm{Tr}(x^{-1}y^{-1}xy) &= \mathrm{Tr}((x^{-1}y^{-1})^2yxxy)\\ &= \mathrm{Tr}(x^{-1}y^{-1})\mathrm{Tr}(xy) - \mathrm{Tr}(yxxy) \ \ ((\text{by } (3))\\ &= \mathrm{Tr}(yx)\mathrm{Tr}(xy) - \mathrm{Tr}(x^2y^2) \ \ ((\text{by } (1),(2)).\end{aligned}$$

It follows that $P_c = s^2 + t^2 + u^2 - stu - 2$.

If one plays around with the polynomials P_w for various words w, they do not appear to have any obvious (or non-obvious) nice behaviour. However, for the magic word $w = x_1^2\,[x_1^{-2}, x_2^{-1}]^2$ in Theorem 4.5, the polynomial P_w turns out to have a miraculous property. We compute that

$$P_w = s^{10} - 2s^9tu - 10s^8 + 2s^8t^2 + s^8t^2u^2 + \cdots - 6s^2u^2 - 2,$$

a polynomial with 29 terms, of degree 12. What is this miraculous property?

Claim Let p be a prime with $p \neq 2, 5$, $p^2 \not\equiv 1 \bmod 16$ and $p^2 \not\equiv 1 \bmod 5$, and let $F = \mathbb{F}_{p^{2r+1}}$. Then

$$P_w(s,t,u) \neq 0 \quad \text{for all } s,t,u \in F.$$

It follows from this that for any $x, y \in SL_2(F)$ we have $\mathrm{Tr}(w(x,y)) = P_w(s,t,u) \neq 0$. Hence the image of w contains no matrices of trace 0, and it follows that w is non-surjective on $PSL_2(F)$, proving Theorem 4.5.

Proof of Claim The claim follows from the following amazing factorization. Letting ζ be a primitive 5^{th} root of unity, P_w factorizes over $\mathbb{Z}[\zeta + \zeta^{-1}]$ as follows:

$$\begin{aligned}P_w(s,t,u) = \ & (s^2 - 2) \ \times\\ & (s^4 - s^3tu + s^2t^2 - 4s^2 + 2 + \zeta + \zeta^{-1}) \ \times\\ & (s^4 - s^3tu + s^2t^2 - 4s^2 + 2 + \zeta^2 + \zeta^{-2}).\end{aligned}$$

Let $s, t, u \in F$. If the first factor $s^2 - 2$ is 0, then F has a square root of 2, which is not the case by the assumption that $p^2 \not\equiv 1 \bmod 16$. And if one of the other

factors is 0, then $\zeta + \zeta^{-1} \in F$, which is also impossible since $p^2 \not\equiv 1 \bmod 5$. Hence $P_w(s,t,u) \neq 0$, proving the claim and the theorem.

One might ask how we came up with the magic word w in Theorem 4.5. The answer is that we computed (by machine) the polynomials P_v for v in a list of representatives of minimal length for certain automorphism classes of words in F_2, generated using [12]. We then tested whether these polynomials were surjective on a selection of small fields. Nothing of interest came up until the length of the representatives reached 14 (which is the length of the magic w). We noticed that P_w was nonzero on the fields $\mathbb{F}_3$ and $\mathbb{F}_{27}$. The rest is history.... It is interesting (to me) to note that although, as I have said, computation played a key role in our discovery of the family of non-surjective words, the final proofs in [29] are completely theoretical and make no use at all of machine computation.

In principle one can try to use the same method to look for non-surjective words on higher rank groups. For example, for a word map w on $G = SL_3(K)$, the trace of $w(x,y)$ for $x, y \in G$ can be expressed as a polynomial in the variables $\mathrm{Tr}(x^{\pm 1})$, $\mathrm{Tr}(y^{\pm 1})$, $\mathrm{Tr}((xy)^{\pm 1})$, $\mathrm{Tr}((x^{-1}y)^{\pm 1})$, $\mathrm{Tr}([x,y])$ (see [28, 4.6]). Again, there is an algorithm for computing these polynomials, so as above one can test for non-surjectivity on small fields in the hope of coming up with promising words. No such promising words have come up in tests so far, and indeed it may be that there are no magic words to be found for higher ranks. In this direction we propose the following conjecture:

Conjecture 4.6 *Let w be a nontrivial word, and assume that w is not a proper power (i.e. there is no word v such that $w = v^m$ with $m \geq 2$). Then there is a constant $r = r(w)$ such that w is surjective on all simple groups of Lie type of rank at least r and all alternating groups of degree at least r.*

4.4 Probability

Recall that for a nontrivial word $w \in F_k$ and a finite group G, we define the probability distribution P_w on G by

$$P_w(g) = \frac{|w^{-1}(g)|}{|G|^k} \quad (g \in G).$$

Let U be the uniform distribution on G (so $U(g) = \frac{1}{|G|}$ for all $g \in G$). For an infinite family $\mathcal{F}$ of groups, we say that the word map w is almost uniform on $\mathcal{F}$ if for groups $G \in \mathcal{F}$ we have

$$||P_w - U||_1 := \sum_{g \in G} |P_w(g) - U(g)| \to 0 \ \text{ as } \ |G| \to \infty.$$

When $\mathcal{F}$ is the finite simple groups, various word maps have been shown to be almost uniform: the commutator word $[x_1, x_2]$ in [17]; and the words $x_1^a x_2^b$ in [37].

Does there exist a word map that is highly non-uniform on a family of simple groups? Currently there is not much evidence for or against this. However as

observed by Macpherson and Tent in [54, 4.10], one can say the following. For a word w and a family $G(q)$ of groups of a fixed Lie type, as $q \to \infty$ the fibres $w^{-1}(g)$ have cardinalities of the order of cq^d with d a non-negative integer, where the number of possibilites for c, d is bounded; the same applies to the cardinality of $w^{-1}(C)$ for a conjugacy class C. It follows, for example, that for a word map $w = w(x_1, \ldots, x_k)$ on the family $PSL_2(p)$ (p prime), as $p \to \infty$ the probability that $w(g_1, \ldots, g_k)$ has order p for random g_i is of the order of $\frac{1}{p^c}$ for $c = 1, 2$ or 3. In particular, it cannot be of an order of magnitude greater than $\frac{1}{p}$, giving a disappointingly negative answer to Question 4 stated at the beginning of this section.

References

[1] M. Abert, Symmetric groups as products of abelian subgroup, *Bull. London Math. Soc.* **34** (2002), 451–456.

[2] Z. Arad, J. Stavi & M. Herzog, Powers and products of conjugacy classes in groups, in: *Products of conjugacy classes in groups*, 6–51, Lecture Notes in Math. **1112**, Springer, Berlin, 1985.

[3] L. Babai, R. Beals & A. Seress, On the diameter of the symmetric group: polynomial bounds, Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 1108–1112, ACM, New York, 2004.

[4] L. Babai, N. Nikolov & L. Pyber, Product growth and mixing in finite groups, In: Proc. 19th Ann. Symp. on Discrete Algorithms (SODA'08), ACM–SIAM 2008, 248–257.

[5] L. Babai & A. Seress, On the diameter of Cayley graphs of the symmetric group, *J. Combin. Theory Ser. A* **49** (1988), 175–179.

[6] L. Babai & A. Seress, On the diameter of permutation groups, *European J. Combin.* **13** (1992), 231–243.

[7] T. Bandman, S. Garion & B. Kunyavskii, Equations in simple matrix groups: algebra, geometry, arithmetic, dynamics, *Cent. Eur. J. Math.* **12** (2014), 175–211.

[8] A. Borel, On free subgroups of semisimple groups, *Enseign. Math.* **29** (1983), 151–164.

[9] J. Bourgain & A. Gamburd, Uniform expansion bounds for Cayley graphs of $SL_2(\mathbb{F}_p)$, *Ann. of Math.* **167** (2008), 625–642.

[10] J. Bourgain, N. Katz & T. Tao, A sum-product estimate in finite fields, and applications, *Geom. Funct. Anal.* **14** (2004), 27–57.

[11] E. Breuillard, B. Green & T. Tao, Approximate subgroups of linear groups, *Geom. Funct. Anal.* **21** (2011), 774–819.

[12] B. Cooper & E. Rowland, Growing words in the free group on two generators, *Illinois J. Math.* **55** (2011), 417–426.

[13] E.W. Ellers & N. Gordeev, On the conjectures of J. Thompson and O. Ore, *Trans. Amer. Math. Soc.* **350** (1998), 3657–3671.

[14] E.E. Ellers, N. Gordeev & M. Herzog, Covering numbers for Chevalley groups, *Israel J. Math.* **111** (1999), 339–372.

[15] R. Fricke & F. Klein, Vorlesungen über die Theorie der Automorphen Funktionen, 1 and 2, Teubner, Leipzig, 1897 and 1912.

[16] J. Fulman & R. Guralnick, Bounds on the number and sizes of conjugacy classes in finite Chevalley groups with applications to derangements, *Trans. Amer. Math. Soc.* **364** (2012), 3023–3070.

[17] S. Garion & A. Shalev, Commutator maps, measure preservation, and T-systems, *Trans. Amer. Math. Soc.* **361** (2009), 4631–4651.

[18] N. Gill, I. Short, L. Pyber & E. Szabó, On the product decomposition conjecture for finite simple groups, *Groups Geom. Dyn.* **7** (2013), 867–882.

[19] R.M. Guralnick, Small representations are completely reducible, *J. Algebra* **220** (1999), 531–541.
[20] R.M. Guralnick, M.W. Liebeck, E.A. O'Brien, A. Shalev & P.H. Tiep, Surjective word maps and Burnside's p^aq^b theorem, preprint.
[21] R. Guralnick & P.H. Tiep, Cross characteristic representations of even characteristic symplectic groups, *Trans. Amer. Math. Soc.* **356** (2004), 4969–5023.
[22] J.I. Hall, M.W. Liebeck & G.M. Seitz, Generators for finite simple groups, with applications to linear groups, *Quart. J. Math. Oxford* **43** (1992), 441–458.
[23] H.A. Helfgott, Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$, *Ann. of Math.* **167** (2008), 601–623.
[24] H.A. Helfgott, Growth in $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$, *J. Eur. Math. Soc.* **13** (2011), 761–851.
[25] H.A. Helfgott & A. Seress, On the diameter of permutation groups, *Ann. of Math.* **179** (2014), 611–658.
[26] D.G. Higman, Intersection matrices for finite permutation groups, *J. Algebra* **6** (1967), 22–42.
[27] E. Hrushovski & A. Pillay, Definable subgroups of algebraic groups over finite fields, *J. Reine Angew. Math.* **462** (1995), 69–91.
[28] S. Jambor, An $L3-U3$-quotient algorithm for finitely presented groups, PhD Thesis, RWTH Aachen University (2012).
[29] S. Jambor, M.W. Liebeck & E.A. O'Brien, Some word maps that are non-surjective on infinitely many finite simple groups, *Bull. Lond. Math. Soc.* (2013), **45**, 907–910.
[30] G.A. Jones, Varieties and simple groups, *J. Austral. Math. Soc.* **17** (1974), 163–173.
[31] M. Kassabov & N. Nikolov, Words with few values in finite simple groups, *Q. J. Math.* **64** (2013), 1161–1166.
[32] M. Kassabov & T.R. Riley, Diameters of Cayley graphs of Chevalley groups, *European J. Combin.* **28** (2007), 791–800.
[33] V. Landazuri & G.M. Seitz, On the minimal degrees of projective representations of the finite Chevalley groups, *J. Algebra* **32** (1974), 418–443.
[34] M.J. Larsen & R. Pink, Finite subgroups of algebraic groups, *J. Amer. Math. Soc.* **24** (2011), 1105–1158.
[35] M. Larsen & A. Shalev, Characters of symmetric groups: sharp bounds and applications, *Invent. Math.* **174** (2008), 645–687.
[36] M. Larsen & A. Shalev, Word maps and Waring type problems, *J. Amer. Math. Soc.* **22** (2009), 437–466.
[37] M. Larsen & A. Shalev, On the distribution of values of certain word maps, preprint, arXiv:1308.1286.
[38] M. Larsen, A. Shalev & P.H. Tiep, Waring problem for finite simple groups, *Ann. of Math.* **174** (2011), 1885–1950.
[39] R. Lawther & M.W. Liebeck, On the diameter of a Cayley graph of a simple group of Lie type based on a conjugacy class, *J. Comb. Theory Ser. A* **83** (1998), 118–137.
[40] A. Lev, The covering number of the group $PSL_n(F)$, *J. Algebra* **182** (1996), 60–84.
[41] M. Levy, Images of word maps in almost simple groups and quasisimple groups, *Internat.J. Algebra Comput.* **24** (2014), 47–58.
[42] M.W. Liebeck, H.D. Macpherson & K. Tent, Primitive permutation groups of bounded orbital diameter, *Proc. Lond. Math. Soc.* **100** (2010), 216–248.
[43] M.W. Liebeck, N. Nikolov & A. Shalev Product decompositions in finite simple groups, *Bull. London Math. Soc.* **44** (2012), 469–472.
[44] M.W. Liebeck, E.A. O'Brien, A. Shalev & P.H. Tiep, The Ore conjecture, *J. Eur. Math. Soc.* **12** (2010), 939–1008.
[45] M.W. Liebeck, E.A. O'Brien, A. Shalev & P.H. Tiep, Products of squares in finite simple groups, *Proc. Amer. Math. Soc.* **140** (2012), 21–33.

[46] M.W. Liebeck & L. Pyber, Finite linear groups and bounded generation, *Duke Math. J.* **107** (2001), 159–171.
[47] M.W. Liebeck & G.M. Seitz, The maximal subgroups of positive dimension in exceptional algebraic groups, *Mem. Amer. Math. Soc.* **169** (2004), No. 802, 1–227.
[48] M.W. Liebeck & G.M. Seitz, *Unipotent and nilpotent classes in simple algebraic groups and Lie algebras*, Amer. Math. Soc. Surveys and Monographs **180** (2012).
[49] M.W. Liebeck & A. Shalev, Classical groups, probabilistic methods, and the $(2,3)$-generation problem, *Ann. of Math.* **144** (1996), 77–125.
[50] M.W. Liebeck & A. Shalev, Diameters of simple groups: sharp bounds and applications, *Ann. of Math.* **154** (2001), 383–406.
[51] A. Lubotzky, *Discrete Groups, Expanding Graphs and Invariant Measures*, Progress in Mathematics, vol. 125, Birkhäuser Verlag, Basel, 1994.
[52] A. Lubotzky, Images of word maps in finite simple groups, *Glasg. Math. J.* **56** (2014), 465–469.
[53] A. Lubotzky, Expander Graphs in Pure and Applied Mathematics, arXiv:1105.2389.
[54] H.D. Macpherson & K. Tent, Pseudofinite groups with NIP theory and definability in finite simple groups, Groups and model theory, 255–267, *Contemp. Math.* **576**, Amer. Math. Soc., Providence, RI, 2012.
[55] N. Nikolov & L. Pyber, Product decompositions of quasirandom groups and a Jordan type theorem, *J. Eur. Math. Soc.* **13** (2011), 1063–1077.
[56] O. Ore, Some remarks on commutators, *Proc. Amer. Math. Soc.* **2** (1951), 307–314.
[57] W. Plesken & A. Fabiańska, An L_2-quotient algorithm for finitely presented groups, *J. Algebra* **322** (2009), 914–935.
[58] L. Pyber & E. Szabó, Growth in finite simple groups of Lie type of bounded rank, preprint, arXiv:1005.1858.
[59] J. Ramo, Strongly real elements of orthogonal groups in even characteristic, *J. Group Theory* **14** (2011), 9–30.
[60] D. Segal, *Words: notes on verbal width in groups*, London Math. Soc. Lecture Note Series **361**, Cambridge University Press, Cambridge, 2009.
[61] A. Selberg, On the estimation of Fourier coefficients of modular forms, *Proc. Symp. Pure Math.* **8** (1965), 1–15.
[62] A. Shalev, Word maps, conjugacy classes, and a noncommutative Waring-type theorem, *Ann. of Math.* **170** (2009), 1383–1416.
[63] R.C. Thompson, Commutators in the special and general linear groups, *Trans. Amer. Math. Soc.* **101** (1961), 16–33.
[64] P.H. Tiep & A. Zalesskii, Real conjugacy classes in algebraic groups and finite groups of Lie type, *J. Group Theory* **8** (2005), 291–315.
[65] I.M. Vinogradov, *The method of trigonometrical sums in the theory of numbers*, (translated, revised and annotated by K. F. Roth and Anne Davenport), Interscience Publishers, London and New York (1954).