Products of squares in finite simple groups

Martin W. Liebeck
Imperial College
London SW7 2BZ
UK

E.A. O'Brien
University of Auckland
Auckland
New Zealand

Aner Shalev
Hebrew University
Jerusalem 91904
Israel

Pham Huu Tiep
University of Arizona
Tucson, AZ 85721
USA

Abstract

The Ore Conjecture, proved in [18], states that every element of every finite (non-abelian) simple group is a commutator. In this paper we use similar methods to prove that every element of every finite simple group is a product of two squares. This can be viewed as a non-commutative analogue of Lagrange's four squares theorem.

1 Introduction

In recent years there has been considerable interest in word maps on finite (non-abelian) simple groups G: namely, maps of the form $(g_1, \ldots, g_d) \mapsto w(g_1, \ldots, g_d)$, where w is a non-identity element of the free group F_d of rank d and $g_i \in G$. Let $w(G)$ denote the image of this map and let $w(G)^k$ be the set of all products of k elements of $w(G)$. In [23] it is shown that $w(G)^3 = G$ provided $|G| > N_w$, where N_w depends only on w, and this has recently been improved to $w(G)^2 = G$ in [15, 16, 17]. Clearly there are words w for which $w(G) \neq G$; for example $w = x^2$ is not surjective on any finite non-abelian simple group. More generally, any word which is a proper power is non-surjective on infinitely many simple groups: indeed, if $w = v^k$ and $|G|$ is not coprime to k, then the map $g \mapsto g^k$ is not injective on G, so $w(G) \neq G$. However, some word maps are surjective, and it is a major challenge to determine which.

Trivially, any primitive word – that is, any word that is part of a free generating set of F_d – is surjective on all groups. The same is true for any word of the form $\prod_{i=1}^{d} x_i^{e_i} f$ with $f \in F_d'$, where $e_1, \ldots, e_d$ are integers with greatest common divisor 1 (see [22, 3.1.1]). The first nontrivial example of a word map which is surjective on all finite non-abelian simple groups is the commutator map $[x, y]$; indeed, this is the content of the well known

Liebeck acknowledges the support of a Maclaurin Fellowship from the New Zealand Institute of Mathematics and its Applications. O'Brien acknowledges the support the Marsden Fund of New Zealand (grant UOA 0721). Shalev acknowledges the support of an ERC Advanced Grant 247034, an EPSRC Visiting Fellowship, an Israel Science Foundation Grant, and a Bi-National Science Foundation grant United States-Israel 2008194. Tiep acknowledges the support of the NSF (grant DMS-0901241).

Ore conjecture that every element of a finite non-abelian simple group is a commutator, recently proved in [18]. Its long proof combines character theory and computational methods. In this paper we use these and other ideas to prove another surjectivity result, this time for the word x^2y^2.

Theorem 1 *Every element of every finite non-abelian simple group G is a product of two squares. In other words, if $w(x,y)$ is the word x^2y^2, then $w(G) = G$.*

The word x^2y^2 is of interest for a number of reasons. Firstly, one can think of the theorem as a non-commutative analogue of Lagrange's four squares theorem. Secondly, it was shown in [8] that the word x^2y^2 is almost measure-preserving on finite simple groups – namely, the inverse image of a subset S of G of proportion $p = |S|/|G|$ has proportion $p + o(1)$ in $G \times G$ as $|G| \to \infty$; but its surjectivity remained open. Note also that by the general result of [17], every element of a sufficiently large finite simple group is a product of two squares; however, it is intriguing that no single exception exists. Thirdly, as for commutators, there is a character theoretic connection essentially going back to Hurwitz (see Lemma 2.2). This paves the way to character theoretic methods which are used in our proof of the theorem.

Our proof for alternating groups and groups of Lie type in odd characteristic is short, using results in [1, 5, 12], and sporadic groups are handled computationally. This leaves the groups $G(q)$ of Lie type in characteristic 2. Using [9] and other tools, we reduce to consideration of classical groups with $q = 2$ or 4. The proof for these groups occupies most of the paper, and uses a similar approach to that in [18], involving character theory, induction on the dimension, and computation to establish base cases.

2 Preliminaries

We begin with a couple of trivial observations.

Lemma 2.1 *Let G be a finite group.*

(i) *If $x \in G$ is an element of odd order, then x is a square.*

(ii) *Suppose the number of squares in G is greater than $\frac{1}{2}|G|$. Then every element of G is a product of two squares.*

Lemma 2.2 *Let G be a finite group. For $g \in G$, the number of pairs $(x,y) \in G \times G$ such that $g = x^2y^2$ is*

$$|G| \cdot \sum_{\chi \in \mathrm{Irr}(G),\ \chi \text{ real}} \frac{\chi(g)}{\chi(1)}.$$

Proof This is a special case of [13, Satz 1]. ∎

Proposition 2.3 (i) *Theorem 1 holds for alternating groups.*

(ii) *Theorem 1 holds for simple groups of Lie type over fields of odd characteristic.*

(iii) *Theorem 1 holds for sporadic simple groups.*

Proof (i) Let $G = \mathsf{A}_n$. If n is odd (resp. even), then every element of G is a product of two n-cycles (resp. $(n-1)$-cycles), by [12] and [1]. Hence every element is a product of two elements of odd order, giving the conclusion by Lemma 2.1(i).

(ii) Let $G = G(q)$ be a simple group of Lie type over a field $\mathbb{F}_q$ with q odd. By a result of Ellers and Gordeev [5, Theorem 3, Corollary], every element of G is a product of two unipotent elements. Since these have odd order, the conclusion follows.

(iii) This follows by a routine check of the character tables of the sporadic groups, using Lemma 2.2. ∎

It follows from the proposition that the only remaining groups to handle are simple groups of Lie type over fields $\mathbb{F}_q$ with $q = 2^k$.

Lemma 2.4 *Let $G = G(q)$, $q = 2^k$. The proportion of elements of G that are of odd order is greater than $\frac{1}{2}$ unless one of the following holds:*

(i) $q = 2$;

(ii) $G = SL_4(4)$, $Sp_{2n}(4)$ *or* $\Omega^{\pm}_{2n}(4)$.

Proof We use the results of [9], which give estimates for the proportion $s(G)$ of elements of even order in G. Theorem 1.1 of [9] says that

$$s(G) < \frac{3}{q-1} + \frac{2}{(q-1)^2}.$$

For $q \geq 8$ this is less than $\frac{1}{2}$, giving the result.

Assume now that $q = 4$. If $G = L_n(4)$ $(n \neq 4)$, the proof of [9, 2.3] (see p. 5) gives $s(G) < \frac{1}{q-1} + \frac{1}{(q-1)^2} < \frac{1}{2}$. If $G = U_n(4)$ then counting a little more carefully in the proof of [9, 2.3] also gives $s(G) < \frac{1}{2}$. The same holds for exceptional groups over $\mathbb{F}_4$ apart from $G_2(4)$ and $F_4(4)$, by [9, 3.1]. For the latter groups, the precise numbers of semisimple elements are listed in [20], and the proportion of them is more than $\frac{1}{2}$. ∎

Lemma 2.5 *The conclusion of Theorem* 1 *holds for the following groups:*

(i) $SL_n(2)$ $(3 \leq n \leq 6)$;

(ii) $SU_n(2)$ $(4 \leq n \leq 9)$;

(iii) $Sp_{2n}(2)$ $(3 \leq n \leq 6)$;

(iv) $Sp_{2n}(4)$ $(n = 2, 3)$;

(v) $\Omega^{\pm}_{2n}(2)$ $(4 \leq n \leq 6)$;

(vi) $\Omega^{\pm}_8(4)$;

(vii) ${}^3D_4(2)$, ${}^2F_4(2)'$, $F_4(2)$, $E^{\pm}_6(2)$, $E_7(2)$.

Proof For all of these groups except $E_7(2)$, we applied 2.2 to the character table. Some of the tables are available in the Character Table Library of GAP [7]; the remainder were constructed directly using the MAGMA [2] implementation of the algorithm of Unger [26]. For $E_7(2)$ one finds using [20] that the proportion of semisimple elements in the group is just over $\frac{1}{2}$. ∎

3 Completion of the proof

It remains to prove Theorem 1 for the following groups:

$SL_n(2)$ $(n \geq 7)$,
$SU_n(2)$ $(n \geq 10)$,
$Sp_{2n}(2)$ $(n \geq 7)$, $Sp_{2n}(4)$ $(n \geq 4)$,
$\Omega^{\pm}_{2n}(2)$ $(n \geq 7)$, $\Omega^{\pm}_{2n}(4)$ $(n \geq 5)$,
$E_8(2)$.

For these groups, the proof follows closely that given in [18]: there we proved that certain key elements $g \in G$ were commutators by using the character theoretic criterion that g is a commutator if $\sum_{\chi \in \mathrm{Irr}(G)} \frac{\chi(g)}{\chi(1)} \neq 0$. In many cases we established this by simply showing that

$$\sum_{1_G \neq \chi \in \mathrm{Irr}(G)} \frac{|\chi(g)|}{\chi(1)} < 1. \tag{1}$$

Of course (1) is sufficient to prove that the sum in Lemma 2.2 is nonzero, and hence that g is equal to x^2y^2 for some $x, y \in G$.

Before proceeding, we eliminate the one exceptional group in the above list.

Lemma 3.1 *Every element of $E_8(2)$ is a product of two squares.*

Proof The proof of the Ore conjecture for $G = E_8(2)$ in [18, §7] was achieved by establishing that for every $g \in G$, either (1) holds or g lies in a subsystem subgroup of G which is a central product of perfect groups of Lie type of rank at most 4 over $\mathbb{F}_2$. By Lemma 2.5, we know that Theorem 1 holds for such subgroups. Hence it holds for G. ∎

3.1 Some generic cases

It remains to prove the theorem for the classical groups over $\mathbb{F}_2$ and $\mathbb{F}_4$ listed above. As in [18] we study a key subset of elements of these groups, defined as follows.

Definition Let $G = Cl_n(q) = SL_n(q)$, $SU_n(q)$, $Sp_n(q)$ or $\Omega_n(q)$. An element $x \in G$ is *breakable* if it lies in a natural proper subgroup $Cl_r(q) \times Cl_{n-r}(q)$ (stabilizing a non-degenerate r-space if $G \neq SL_n(q)$), and one of the following holds:

(1) both factors $Cl_r(q)$ and $Cl_{n-r}(q)$ are perfect groups;

(2) $Cl_r(q)$ is perfect, and the projection of x to $Cl_{n-r}(q)$ is a product of two squares in $Cl_{n-r}(q)$.

Otherwise, x is *unbreakable*.

A simple induction argument shows that Theorem 1 for $G = Cl_n(q)$ follows immediately if we prove that every unbreakable element in G is a product of two squares. Indeed, let $x \in G$ and suppose that x is breakable, so $x = (x_1, x_2) \in Cl_r(q) \times Cl_{n-r}(q)$ satisfies (1) or (2) in the above definition. In either case, by induction x_1, x_2 are products of two squares in $Cl_r(q)$, $Cl_{n-r}(q)$ respectively, say $x_i = y_i^2 z_i^2$ for $i = 1, 2$; then $x = (y_1, y_2)^2 (z_1, z_2)^2$.

Lemma 3.2 *Theorem* 1 *holds for* $G = Sp_{2n}(2)\,(n \geq 7)$, $\Omega^{\pm}_{2n}(2)\,(n \geq 7)$, *or* $\Omega^{\pm}_{2n}(4)\,(n \geq 5)$.

Proof This follows almost immediately from the proofs of the Ore conjecture for these groups in [18], as follows. Consider first $G = Sp_{2n}(2)$. By [10, 6.2], there is a set $\mathcal{W}$ of five Weil characters such that every nontrivial irreducible character of G not in $\mathcal{W}$ has degree at least $(2^{2n}-1)(2^{n-1}-1)(2^{n-1}-4)/30$. Moreover the characters in $\mathcal{W}$ have distinct degrees (see [10, §3]), so in particular they are all real. For $x \in G$ let

$$F_1(x) = \sum_{\chi \in \mathcal{W}} \frac{\chi(x)}{\chi(1)}, \quad F_2(x) = \sum_{1_G \neq \chi \in \mathrm{Irr}(G) \setminus \mathcal{W}} \frac{|\chi(x)|}{\chi(1)}.$$

By Lemma 2.2 it is sufficient to show that $|F_1(x)| + F_2(x) < 1$ for all unbreakable $x \in G$. This is established in Lemmas 4.4–4.7 of [18].

An entirely similar discussion applies when G is an orthogonal group. ■

The remaining groups $SL_n(2)$, $SU_n(2)$, $Sp_{2n}(4)$ were not handled in the above way in [18], so require detailed arguments here.

3.2 Special linear and symplectic groups

Lemma 3.3 *Theorem* 1 *holds for* $G = SL_n(2)\,(n \geq 7)$.

Proof Let $x \in G$ be unbreakable. We claim first that

$$|C_G(x)| \leq 2^n \cdot 3^2. \tag{2}$$

The proof is similar to analogous proofs in [18], for example [18, 4.7]. If x is unipotent and unbreakable, then it has only Jordan blocks of size 2 or at least $n-2$ (size 2 is allowed as $SL_2(2)$ is non-perfect); it follows easily that $x = J_n$ or (J_{n-2}, J_2) where J_i denotes a Jordan block of size i. These have centralizers of order 2^n or 2^{n+2} respectively, so (2) holds. If x is not unipotent, write $x = su$ with s, u commuting semisimple and unipotent elements. As x is unbreakable, one of the following holds:

(i) $C_G(s) = GL_{n/k}(2^k)$ for some k dividing n, and $u = J_{n/k} \in GL_{n/k}(2^k)$;

(ii) $n = 2m$, $C_G(s) = GL_m(4)$ and $u = (J_{m-1}, J_1) \in GL_m(4)$;

(iii) $n = 2m+2$, $C_G(s) = GL_m(4) \times GL_1(4)$ or $GL_m(4) \times GL_2(2)$ and $u = (J_m, J_1)$ or $J_m, J_2)$ respectively.

The centralizer orders $|C_G(x)|$ are given by [19, §2], and the largest is that in case (ii), where $|C_G(x)| = 4^m |GL_1(4)|^2$. Hence (2) is proved.

Next, define $D = (2^n-1)(2^{n-1}-4)/3$. By [24, Theorem 3.1], G has exactly one nontrivial irreducible character χ_0 of degree less than D, and χ_0 is the nontrivial constituent of the permutation character of G on 1-spaces, of degree $2^n - 2$. We saw above that $\dim C_V(x) \leq 2$ (where $V = V_n(2)$), so

$$\frac{|\chi_0(x)|}{\chi_0(1)} \leq \frac{3}{2^n - 2}.$$

The number $k(G)$ of conjugacy classes of G satisfies $k(G) \leq (2.5) \cdot 2^{n-1}$ by [6, 3.6]. Hence by [18, 2.6],

$$\sum_{\chi \in \mathrm{Irr}(G),\ \chi(1) \geq D} \frac{|\chi(x)|}{\chi(1)} \leq \frac{k(G)^{1/2} |C_G(x)|^{1/2}}{D} \leq \frac{\sqrt{2.5} \cdot 2^{(n-1)/2} \cdot 2^{n/2} \cdot 3}{D} < 0.2.$$

It follows that

$$\sum_{1_G \neq \chi \in \mathrm{Irr}(G)} \frac{|\chi(x)|}{\chi(1)} < \frac{3}{2^n - 2} + 0.2 < 1,$$

and so by Lemma 2.2 the conclusion follows. ∎

Lemma 3.4 *Theorem* 1 *holds for* $G = Sp_{2n}(4)$ $(n \geq 4)$.

Proof Let $x \in G$ be unbreakable. As $Sp_2(4)$ is perfect, x cannot stabilize a proper non-degenerate subspace. It follows that if x is unipotent then $x = J_{2n}$ or J_n^2 ($V(2n)$ or $W(n)$ in the notation of [19, §3]); and if $x = su$ is not unipotent, then $C_G(s) = GL^{\pm}_{n/k}(4^k)$ and $u = J_{n/k} \in GL^{\pm}_{n/k}(4^k)$, for some k dividing n. The centralizer orders are given in [19], and the largest is for $x = J_n^2$, which has centralizer order $4^{2n-1} \cdot |Sp_2(4)|$ if n is even, $4^{2n} \cdot |O_2^{\pm}(4)|$ if n is odd. Hence

$$|C_G(x)| < 4^{2n+2}.$$

Now we consider characters of G. By [10, 6.2], there is a collection $\mathcal{W}$ of 7 irreducible characters such that every nontrivial irreducible character not in $\mathcal{W}$ has degree at least

$$D = \frac{(4^{2n} - 1)(4^{n-1} - 1)(4^{n-1} - 4^2)}{2(4^4 - 1)}.$$

The 7 characters in $\mathcal{W}$ are labelled $\alpha_n, \beta_n, \rho_n^1, \rho_n^2, \tau_n^1, \zeta_n^1, \zeta_n^2$, and all are real. The values of $\alpha_n + \beta_n$, $\rho_n^1 + \rho_n^2$, τ_n^1, ζ_n^1, and ζ_n^2 are described explicitly in [10, §3], and we see easily as in the proof of [18, Lemma 4.5] that

$$|F_1(x)| = \left| \sum_{\chi \in \mathcal{W}} \frac{\chi(x)}{\chi(1)} \right| < 0.2.$$

Also $k(G) \leq (15.2) \cdot 4^n$ by [6, 3.13], so as in the previous proof

$$F_2(x) = \sum_{\chi \in \mathrm{Irr}(G),\ \chi(1) \geq D} \frac{|\chi(x)|}{\chi(1)} \leq \frac{\sqrt{15.2} \cdot 2^n \cdot 4^{n+1}}{D} < 0.2.$$

Hence $|F_1(x)| + F_2(x) < 1$ and the conclusion follows. ∎

3.3 Unitary groups

The proof for $SU_n(2)$ with $n \geq 10$ is similar to the previous ones, but we give more detail as unitary groups were handled by a different method in [18].

Lemma 3.5 *If* $x \in G = SU_n(2)$ *is unbreakable, then* $|C_G(x)| \leq 2^{3n-6} \cdot 3^4$. *Moreover, if* $n = 10$ *then* $|C_G(x)| \leq 2^{15} \cdot 3^4$.

Proof The non-perfect special unitary groups are $SU_r(2)$ with $r \leq 3$, so x does not lie in a subgroup $SU_r(2) \times SU_{n-r}(2)$ with $r \in \{1, 2, 3\}$. If x is unipotent then it is J_n, (J_{n-2}, J_2) or (J_{n-3}, J_3), all of which have centralizer order smaller than the bound in the conclusion (see [19, §2]). Otherwise $x = su$ with s a nontrivial semisimple element, and the largest possible

centralizer is achieved when n is a multiple of 3, $C_G(s) = G \cap GU_{3m}(2^k)$ ($n = 3mk$, k odd), and $u = (J_m^3) \in GU_{3m}(2^k)$. Writing $q = 2^k$ we have

$$\begin{aligned} |C_G(x)| \leq C_{GU_{3m}(q)}(u) &= q^{9m-9}|GU_3(q)| \\ &= 2^{3n}.\tfrac{(q^3+1)(q^2-1)(q+1)}{q^6}. \end{aligned}$$

The right hand side attains a maximum when $k = 1$ and $q = 2$; it achieves the bound in the conclusion. The stronger bound for $n = 10$ was established computationally – it occurs for the element $x = (\omega J_2, \omega J_2, \omega J_3, x_0)$ where ω is a cube root of 1 in $\mathbb{F}_4$ and x_0 is an element of order 9 in $GU_3(2)$. ∎

Consider the natural module $V = \mathbb{F}_4^n$ for $G := GU_n(2)$. Any $g \in G$ is *indecomposable* if V cannot be decomposed into a direct sum of pairwise orthogonal g-invariant nonzero non-degenerate subspaces. A nonzero non-degenerate subspace U of V is a *λ-block* for g if U is g-invariant, $g|_U$ is indecomposable, and $\det(g|_U) = \lambda$. A *block* for g is a λ-block for some $\lambda \in \mathbb{F}_4^\times = \{1, \omega, \omega^2\}$.

Lemma 3.6 *Assume $g \in SU_n(2)$ is unbreakable and $n \geq 10$. Then V is a direct sum of at most 6 blocks.*

Proof Write

$$V = A_1 \oplus A_2 \oplus \ldots \oplus A_r \oplus B_1 \oplus B_2 \oplus \ldots \oplus B_s \oplus C_1 \oplus C_2 \oplus \ldots \oplus C_t$$

as a direct sum of $r + s + t$ blocks for g, where A_i is a 1-block of dimension a_i, B_j is an ω-block of dimension b_j, C_k is an ω^2-block of dimension c_k, $r, s, t \geq 0$, and

$$1 \leq a_1 \leq \ldots \leq a_r, \quad 1 \leq b_1 \leq \ldots \leq b_s, \quad 1 \leq c_1 \leq \ldots \leq c_t.$$

Arguing by contradiction, we assume that $r + s + t \geq 7$.

1) First we claim that V cannot be written as an orthogonal sum of three g-invariant non-degenerate subspaces $A \oplus B \oplus C$ with $\dim A, \dim B, \dim C \geq 2$ and $\det(g|_A) = \det(g|_B) = \det(g|_C) = 1$. Indeed, assume the contrary and (without loss) that $a := \dim A \leq b := \dim B \leq c := \dim C$. Then $c \geq 4$ as $n \geq 10$, and $a + b \geq 4$. Now observe that $g \in SU_c(2) \times SU_{a+b}(2)$, and so g is breakable.

2) Here we show that $\min(s, t) \leq 1$. For, assuming that $s, t \geq 2$ we can define

$$A := B_1 \oplus C_1, \quad B := B_2 \oplus C_2, \quad C := (A \oplus B)^\perp.$$

Then $\dim A, \dim B \geq 2$ and $\dim C \geq r + s + t - 4 \geq 3$. Furthermore, the choices of A, B, C ensure that $g|_A$, $g|_B$, and $g|_C$ all have determinant 1. Hence g is breakable by 1).

3) Next we consider the case $r \geq 2$. If $s, t \geq 1$, then we can define

$$A := A_1 \oplus A_2, \quad B := B_1 \oplus C_1, \quad C := (A \oplus B)^\perp$$

and conclude that g is breakable by 1). Furthermore, if $s \geq 3$ for instance, then we define

$$A := A_1 \oplus A_2, \quad B := B_1 \oplus B_2 \oplus B_3, \quad C := (A \oplus B)^\perp$$

and again conclude that g is breakable by 1). So we must have that $\min(s,t) = 0$ and $\max(s,t) \le 2$. Replacing g by g^{-1} if necessary, we may assume that $t = 0$ and $s \le 2$. In this case, $1 = \det(g) = \omega^s$, hence $s = 0$ and $r \ge 7$. Now if $d := a_1 + a_2 + a_3 \ge 4$ then

$$g \in SU(A_1 \oplus A_2 \oplus A_3) \times SU((A_1 \oplus A_2 \oplus A_3)^\perp) = SU_d(2) \times SU_{n-d}(2)$$

with $n - d \ge 4$, and so g is breakable. So $a_1 = a_2 = a_3 = 1$. If in addition $3 + a_4 \le n - 4$, then

$$g \in SU(A_1 \oplus \ldots \oplus A_4) \times SU((A_1 \oplus \ldots \oplus A_4)^\perp) = SU_{3+a_4}(2) \times SU_{n-3-a_4}(2)$$

and again g is breakable. Thus $3 + a_4 \ge n - 3$ and so $a_4 \ge n - 6 \ge 4$. This is impossible, since $n - 3 - a_4 > a_5 + a_6 \ge 2a_4 \ge 8$.

4) We have shown that $r \le 1$. By 2) we may assume $t \le 1$, whence $s = n - r - t \ge 5$. Since $1 = \det(g) = \omega^{s+2t}$, we must have $3|(s-t)$, and so $s \ge 6$. Defining

$$A := B_1 \oplus B_2 \oplus B_3, \quad B := B_4 \oplus B_5 \oplus B_6, \quad C := (A \oplus B)^\perp,$$

by 1) we see that the unbreakability of g implies that $\dim C \le 1$. But $\dim C \ge (r + s + t) - 6$, so we must have that $r + s + t = 7$, and either $r = t = 0$ and $s = 7$, or $r + t = 1$ and $s = 6$. The former is impossible as $3|(s-t)$. Similarly, in the latter case we have $t = 0$ and $r = 1$, and moreover $C = A_1$ has dimension 1. Recalling that A_1 is a 1-block, we see $g|_{A_1} = 1$ and so $g \in SU_{n-1}(2)$, again a contradiction. ∎

For $\lambda \in \mathbb{F}_{q^2}^\times$ and $g \in GU_n(q)$, define $e(g,\lambda)$ to be the dimension (over $\mathbb{F}_{q^2}$) of $\mathrm{Ker}(g - \lambda \cdot 1_V)$. If $f \in \mathbb{F}_{q^2}[t]$ is a monic irreducible polynomial with a root $\alpha \in \overline{\mathbb{F}}_{q^2}$, define $\check{f}$ to be the unique monic irreducible polynomial over $\mathbb{F}_{q^2}$ such that $\check{f}(\alpha^{-q}) = 0$.

Corollary 3.7 *If $n \ge 10$ and $g \in SU_n(2)$ is unbreakable, then*

$$\sum_{\lambda \in \mathbb{F}_4^\times} e(g,\lambda) \le 6.$$

Proof Consider an α-block U of g for $\alpha \in \mathbb{F}_4^\times$. Since $g|_U$ is indecomposable, the semisimple part s of $g|_U$ must have characteristic polynomial $f(t)^k$, where $f \in \mathbb{F}_4[t]$ is irreducible and $f = \check{f}$, or $(f(t)\check{f}(t))^k$, where $f \in \mathbb{F}_4[t]$ is irreducible and $f \ne \check{f}$. In both cases, if $\deg(f) \ge 2$, then $f(\lambda) \ne 0$ for $\lambda \in \mathbb{F}_4^\times$ and so $\dim \mathrm{Ker}(g|_U - \lambda \cdot 1_U) = 0$. On the other hand, if $\deg(f) = 1$, then $f(t) = t - \beta$ for some $\beta \in \mathbb{F}_4^\times$, whence $f = \check{f}$ (i.e. we are in the former case) and $s = \beta \cdot 1_U$. Then the indecomposability of $g|_U$ implies that the unipotent part of g acts on U as a single Jordan block of size k. Thus $\dim \mathrm{Ker}(g|_U - \lambda \cdot 1_U)$ equals 1 if $\lambda = \beta$ and 0 otherwise. Now the statement follows immediately from Lemma 3.6. ∎

Lemma 3.8 *Let p be a prime, q a power of p, and let $s \in G := GU_n(q)$ be a semisimple element such that s and αs are conjugate, where $1 = \alpha^{q+1} \ne \alpha \in \mathbb{F}_{q^2}^\times$. Then*

$$N(s) := (G : C_G(s))_{p'} > q^{n^2/4-2}.$$

Proof 1) As usual, we may decompose V into an orthogonal sum $V_1 \oplus \ldots \oplus V_m$ of s-invariant non-degenerate subspaces, with the characteristic polynomial of $s|_{V_i}$ being either $f_i(t)^k$, where $f_i \in \mathbb{F}_{q^2}[t]$ is irreducible and $f_i = \check{f}_i$, or $(f_i(t)\check{f}_i(t))^k$, where $f_i \in \mathbb{F}_{q^2}[t]$ is irreducible and $f_i \neq \check{f}_i$, and moreover $f_i \neq f_j, \check{f}_j$ whenever $i \neq j$. Since s and αs are conjugate, the the map $J : x \mapsto \alpha x$ on $\bar{F}_q$ preserves the set of all eigenvalues of s. As $\alpha = \alpha^{q^2}$, J induces an action on the set of irreducible factors f_i of the characteristic polynomial of s. Moreover, since $\alpha = \alpha^{-q}$, this action of J commutes with the map $f_i \mapsto \check{f}_i$. Thus J induces an action (which we also denote by J) on the set of summands $V_1, \ldots, V_m$, with say l orbits. We consider the decomposition

$$V = W_1 \oplus W_2 \oplus \ldots \oplus W_l,$$

where each W_j is the sum of all V_i belonging to one orbit of J.

2) Observe that, by our construction, $s_j := s|_{W_j}$ and αs_j are conjugate in $GU(W_j)$. (Indeed, if $xsx^{-1} = \alpha s$ for some $x \in GU(V)$, then $s|_{x(W_j)}$ has characteristic polynomial equal to the image under J of the characteristic polynomial of $s|_{W_j}$. It follows that $x(W_j) = W_j$ and $x|_{W_j}$ conjugates s_j to αs_j.) In particular, $\dim W_j > 1$ (as $\alpha \neq 1$). Furthermore, $C_G(s)$ preserves each W_j (since it fixes each V_i).

3) The desired bound is obvious when $n \leq 5$. We proceed by induction on $n \geq 6$. First we consider the case $l \geq 2$. The observations in 2) allow us to apply the induction hypothesis to $s_1 = s|_{W_1}$ and $s_1' := s|_{W_1'}$ with $W_1' := (W_1)^\perp$. Hence

$$N(s_1) = \big(GU(W_1) : C_{GU(W_1)}(s_1)\big)_{p'} > q^{a^2/4-2},$$
$$N(s_1') = \Big(GU(W_1') : C_{GU(W_1')}(s_1')\Big)_{p'} > q^{b^2/4-2},$$

where $a := \dim W_1 \geq 2$ and $b := \dim W_1' \geq 2$. Observe that

$$q^{k(k+1)/2} < f(q,k) := \prod_{i=1}^{k}(q^i - (-1)^i) \leq \frac{q+1}{q} \cdot q^{k(k+1)/2} \tag{3}$$

for any $k \geq 1$. It follows that

$$\frac{|GU(V)|_{p'}}{|(GU(W_1) \times GU(W_1'))|_{p'}} = \frac{f(q,n)}{f(q,a)f(q,b)} > \frac{q^{\frac{n(n+1)}{2}}}{\left(\frac{q+1}{q}\right)^2 \cdot q^{\frac{a(a+1)+b(b+1)}{2}}} > q^{ab-2}.$$

Since $n = a + b \geq 6$ and $a, b \geq 2$, we have $ab \geq 8$. Consequently,

$$N(s) > N(s_1)N(s_1') \cdot \frac{f(q,n)}{f(q,a)f(q,b)} > q^{\frac{a^2+b^2}{4}-4} \cdot q^{ab-2} \geq q^{\frac{n^2}{4}-2},$$

as stated.

4) Now we may assume that $l = 1$. First we consider the case where $f_i = \check{f}_i$. The characteristic polynomial of s is $(f_1 f_2 \ldots f_m)^k$, and $\deg f_i = r$ is odd. Then $C_G(s) \cong GU_k(q^r)^m$ and $n = krm$. Applying (3) we get

$$N(s) = \frac{f(q,n)}{f(q^r,k)^m} > \frac{q^{\frac{n(n+1)}{2}}}{\left(\frac{q^r+1}{q^r}\right)^m \cdot q^{mr\frac{k(k+1)}{2}}} > q^{\frac{n^2}{4}+\frac{n}{4}(mkr-2k)-\frac{3m}{5}}$$

(since $(q^r+1)/q^r \le 3/2 < q^{3/5}$). If $mr \ge 3$ and $n \ge 8$, then $mkr - 2k \ge n/3$ and $n^2/12 > 3m/5$, yielding $N(s) > q^{n^2/4}$. The same holds if $mr \ge 3$ and $n = 6, 7$. Assume $mr \le 2$. Since r is odd, we have $r = 1$. If in addition $m = 1$ then s is scalar and so s and αs cannot be conjugate. Thus $m \ge 2$ and so $N(s) > q^{n^2/4-2}$.

Finally we consider the case where $f_i \ne \check{f}_i$. The characteristic polynomial of s is $(f_1 f_2 \dots f_m \check{f}_1 \check{f}_2 \dots \check{f}_m)^k$, and $\deg f_i = r \ge 1$. Then $C_G(s) \cong GL_k(q^{2r})^m$ and $n = 2krm$. Applying (3), we get

$$N(s) > \frac{f(q,n)}{q^{rmk(k+1)}} > \frac{q^{\frac{n(n+1)}{2}}}{q^{mrk(k+1)}} > q^{\frac{n^2}{2} - \frac{n}{2}k} \ge q^{\frac{n^2}{4}},$$

and so we are done with the inductive step. ■

Now we prove the following theorem which is of independent interest:

Theorem 3.9 *Let q be a power of a prime p and let $\ell = 0$ or a prime coprime to $\gcd(n, q+1)$. Assume V is an ℓ-modular absolutely irreducible representation of $GU_n(q)$ which is reducible on restriction to $SU_n(q)$. Then*

$$\dim V > q^{\frac{n^2}{4}-2}.$$

Proof 1) Let $G := GU_n(q)$, $S := SU_n(q)$, $Z := Z(G)$. Consider the subgroups A, B of G which contain S and such that $A/S = O_\ell(G/S)$ and $B/S = O_{\ell'}(G/S)$. Since $(G : ZS) = \gcd(n, q+1)$ is coprime to ℓ, we have $A \le ZS$. For $X \lhd G$, let $\kappa_X^G(V)$ denote the total number of irreducible constituents of the X-module $V|_X$. Similarly, we choose U to be an irreducible constituent of the A-module $V|_A$, and let $\kappa_Y^A(U)$ denote the total number of irreducible constituents of the Y-module $U|_Y$ for $Y \lhd A$. Since $A \le ZS$, every S-irreducible constituent W of $U|_S$ is A-invariant. But A/S is cyclic, hence W extends to A. In other words, $\kappa_S^A(U) = 1$. By our assumptions, $\kappa_S^G(V) > 1$, and by [14, Lemma 3.3],

$$\kappa_S^G(V) = \kappa_A^G(V) \cdot \kappa_S^A(U).$$

Hence $\kappa_A^G(V) > 1$.

Recall that G/A is a cyclic ℓ'-group. The latter inequality implies by [14, Lemma 3.2] that there is some nontrivial irreducible ℓ-modular representation L of G which is trivial on A such that $V \simeq V \otimes L$.

2) Observe that the dual group G^* of $G = GU_n(q)$ can be identified with G. Hence, $\mathrm{Irr}(G)$ is the disjoint union of the rational series $\mathcal{E}(G,(x))$, where (x) runs over the set of conjugacy classes (x) of semisimple elements $x \in G$, cf. [4, 21]. Furthermore, according to the main result of [3], $\mathrm{Irr}(G)$ can be partitioned into the disjoint union of $\mathcal{E}_\ell(G,(y))$, where each $\mathcal{E}_\ell(G,(y))$, labelled by the conjugacy class (y) of semisimple ℓ'-elements $y \in G$, and defined by

$$\mathcal{E}_\ell(G,(y)) = \bigcup_{t \in C_G(y),\ t \text{ is an } \ell-\text{element}} \mathcal{E}(G,(yt)),$$

is a union of ℓ-blocks.

Assume that V belongs to the union $\mathcal{E}_\ell(G,(s))$ of blocks labelled by the conjugacy class of a semisimple ℓ'-element $s \in G$. Since S acts trivially on

L but L is nontrivial, we can also find a nontrivial ℓ'-element $z \in Z$ such that the Brauer character of L is just the restriction to ℓ'-elements in G of the semisimple character χ_z labelled by z. According to [4, Proposition 13.30] and its proof, the tensor product with χ_z defines a bijection between the series $\mathcal{E}(G,(x))$ and $\mathcal{E}(G,(xz))$, hence also between the unions of blocks $\mathcal{E}_\ell(G,(s))$ and $\mathcal{E}_\ell(G,(sz))$. Since $V \simeq V \otimes L$, we conclude that $(s) = (sz)$, i.e. s and sz are conjugate in G. By Lemma 3.8, $N(s) = (G : C_G(s))_{p'} > q^{n^2/4-2}$. Finally, by [11, Proposition 1], $\dim V$ is divisible by $N(s)$, whence the statement follows. ■

Lemma 3.10 *Let $S = SU_n(2)$ with $n \geq 10$ and let*

$$D = \begin{cases} \frac{(2^n-1)(2^{n-1}+1)(2^{n-2}-1)}{3^4}, & n \text{ even.} \\ \frac{(2^n+1)(2^{n-1}-1)(2^{n-2}-2^3)}{3^4}, & n \text{ odd} \end{cases}$$

If $1_S \neq \chi \in \mathrm{Irr}(S)$ and $\chi(1) < D$, then χ is either one of three Weil characters, or one of the characters D°_α defined in [18, Proposition 6.3].

Proof Let $G = GU_n(2)$ and let $\theta \in \mathrm{Irr}(G)$ be such that χ is an irreducible constituent of $\theta|_S$. Clearly, $\theta(1) \leq 3\chi(1) < 3D < 2^{n^2/4-2}$. It then follows by Theorem 3.9 that $\theta|_S$ is irreducible, whence $\chi = \theta|_S$. In particular, $\theta(1) = \chi(1) < D$. Now the statement follows from [18, Proposition 6.6]. ■

Lemma 3.11 *If $g \in S = SU_n(2)$ is unbreakable, $n \geq 10$, and D is as in the previous lemma, then*

$$F_1(g) = \sum_{1<\chi(1)<D,\ \chi \text{ real}} \frac{|\chi(g)|}{\chi(1)} < \begin{cases} 0.09, & \text{if } n \geq 11, \\ 0.28, & \text{if } n = 10. \end{cases}$$

Proof 1) It is well known that among the three Weil characters $\zeta^i_{n,2}$ of S, only the unipotent character $\zeta := \zeta^0_{n,2}$ is real. Next we show that the character D°_α is real if and only if $\alpha \in \mathrm{Irr}(GU_2(2))$ is real. The characters D°_α are constructed in [18, §6.1] by embedding a central product $H * S$ in $GU_{2n}(2)$ for $H := GU_2(2)$ and restricting the reducible Weil character

$$\varphi(g) = (-2)^{\dim_{\mathbb{F}_4} \mathrm{Ker}(g-1)}$$

of $GU_{2n}(2)$ to $H * S$. In particular,

$$\varphi|_{H*G} = \sum_{\alpha \in \mathrm{Irr}(H)} \alpha \otimes D_\alpha.$$

By [18, Proposition 6.3], $D^\circ_\alpha = D_\alpha - \kappa_\alpha \cdot 1_S$ with $\kappa_\alpha \in \{0,1\}$; in particular, D°_α is real if and only if D_α is real. Since φ is real,

$$\sum_{\alpha \in \mathrm{Irr}(H)} \bar\alpha \otimes \bar D_\alpha = \bar\varphi|_{H*G} = \varphi|_{H*G} = \sum_{\alpha \in \mathrm{Irr}(H)} \bar\alpha \otimes D_{\bar\alpha},$$

whence $\bar D_\alpha = D_{\bar\alpha}$. Also by [18, Proposition 6.3], for $\alpha, \beta \in \mathrm{Irr}(H)$, $D_\alpha = D_\beta$ precisely when $\alpha = \beta$. We conclude that D°_α is real precisely when α is real.

2) Observe that $H = GU_2(2)$ has exactly three real irreducible characters: $\alpha_1 = 1_H$, α_2 the Steinberg character (of degree 2), and one more, α_3,

of degree 1 (which is $\chi_{q-1}^{(1,2)}$ in the notation of Table III of [18]). Thus the summation in $F_1(g)$ involves 4 characters: ζ, and $\chi_i = D_{\alpha_i}^\circ$ for $1 \le i \le 3$. First by [25, Lemma 4.1] we have

$$\zeta(g) = \frac{(-1)^n}{3} \sum_{\lambda \in \mathbb{F}_4^\times} (-2)^{e(g,\lambda)}.$$

By Corollary 3.7, $\sum_{\lambda \in \mathbb{F}_4^\times} e(g,\lambda) \le 6$. It follows that $|\zeta(g)| \le (2^6+2)/3 = 22$. Next,

$$D_\alpha(g) = \frac{1}{|H|} \sum_{x \in H} \overline{\alpha(x)} \varphi(xg)$$

for $\alpha \in \mathrm{Irr}(H)$. The computations in the proof of [18, Proposition 6.9] and Corollary 3.7 show that, for every $x \in H$, $|\varphi(xg)| \le 2^{12}$. Since $|\alpha(x)| \le \alpha(1)$, it follows that $|D_\alpha(g)| \le 2^{12}\alpha(1)$ and so $|D_\alpha^\circ(g)| \le 2^{12}\alpha(1) + 1$. Now for $i = 1, 3$ we have $\alpha_i(1) = 1$ and $\chi_i(1) \ge (2^n - 2)(2^{n-1} - 4)/9$, whereas for $i = 2$ we have $\alpha_i(1) = 2$ and $\chi_i(1) > 2(2^n - 2)(2^{n-1} - 4)/9$, cf. [18, Table III]. Also, $\zeta(1) \ge (2^n - 2)/3$. It follows that

$$F_1(g) < \frac{66}{2^n - 2} + 2 \cdot \frac{2^{12} + 1}{(2^n - 2)(2^{n-1} - 4)/9} + \frac{2^{13} + 1}{2(2^n - 2)(2^{n-1} - 4)/9}$$

which is less than 0.09 if $n \ge 11$ and 0.28 if $n = 10$. ■

Now we prove the main result of this subsection, which also completes the proof of Theorem 1:

Proposition 3.12 *Theorem* 1 *holds for* $G = SU_n(2)$ *with* $n \ge 10$.

Proof It suffices to show that every unbreakable $g \in G$ is a product of two squares. We have $k(G) < (8.26) \cdot 2^{n-1}$ by [6, 3.10], so in the usual fashion we see that

$$F_2(g) = \sum_{\chi(1) \ge D} \frac{|\chi(g)|}{\chi(1)} \le \frac{\sqrt{8.26} \cdot 2^{(n-1)/2} \cdot 2^{(3n-6)/2} \cdot 3^2}{D},$$

where D is as defined in Lemma 3.10. For $n \ge 11$ this yields $F_2(g) < 0.74$. For $n = 10$ we use the stronger bound for $|C_G(x)|$ in Lemma 3.5 to obtain $F_2(x) \le \sqrt{8.26} \cdot 2^{9/2} \cdot 2^{15/2} \cdot 3^2/D$, which is less than 0.07. On the other hand, $F_1(g) < 0.09$ for $n \ge 11$ and $F_1(g) < 0.28$ for $n = 10$ by Lemma 3.11. Thus $F_1(g) + F_2(g) < 0.83$ for all unbreakable $g \in G$, and so the statement follows. ■

References

[1] E. Bertram, Even permutations as a product of two conjugate cycles, *J. Comb. Theory Ser. A*, **12** (1972), 368–380.

[2] W. Bosma, J. Cannon and C. Playoust, The Magma algebra system I: The user language, *J. Symbolic Comput.* **24** (1997), 235–265.

[3] M. Broué and J. Michel, Blocs et séries de Lusztig dans un groupe réductif fini, *J. Reine Angew. Math.* **395** (1989), 56–67.

[4] F. Digne and J. Michel, *Representations of Finite Groups of Lie Type*, London Mathematical Society Student Texts 21, Cambridge University Press, 1991.

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

[6] J. Fulman and R. M. Guralnick, Bounds on the number and sizes of conjugacy classes in finite Chevalley groups, preprint.

[7] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.4.9*, 2006, `http://www.gap-system.org`.

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

[9] R.M. Guralnick and F. Lübeck, On p-singular elements in Chevalley groups in characteristic p, in: '*Groups and computation*, III (Columbus, OH, 1999), pp. 169–182, Ohio State Univ. Math. Res. Inst. Publ., 8, de Gruyter, Berlin, 2001.

[10] R. Guralnick and P.H. Tiep, Cross characteristic representations of even characteristic symplectic groups, *Trans. Amer. Math. Soc.* **356** (2004), 4969–5023.

[11] G. Hiss and G. Malle, Low dimensional representations of special unitary groups, *J. Algebra* **236** (2001), 745–767.

[12] D. Husemoller, Ramified coverings of Riemann surfaces, *Duke Math. J.* **29** (1962), 167–174.

[13] A. Kerber and B. Wagner, Gleichungen in endlichen Gruppen, *Arch. Math.* **35** (1980), 252–262.

[14] A. S. Kleshchev and P.H. Tiep, Representations of finite special linear groups in non-defining characteristic, *Adv. Math.* **220** (2009), 478–504.

[15] M. Larsen and A. Shalev, Characters of symmetric groups: sharp bounds and applications, *Invent. Math.* **174** (2008), 645–687.

[16] M. Larsen and A. Shalev, Word maps and Waring type problems, *J. Amer. Math. Soc.* **22** (2009), 437–466.

[17] M. Larsen, A. Shalev and P.H. Tiep, The Waring problem for finite simple groups, preprint.

[18] M.W. Liebeck, E.A. O'Brien, A. Shalev and P.H. Tiep, The Ore Conjecture, *J. Europ. Math. Soc.*, to appear.

[19] M.W. Liebeck and G.M. Seitz, Unipotent and nilpotent classes in simple algebraic groups and Lie algebras, preprint.

[20] Frank Lübeck, Data for Finite Groups of Lie Type and Related Algebraic Groups. `www.math.rwth-aachen.de/~Frank.Luebeck/chev`.

[21] G. Lusztig, *Characters of Reductive Groups over a Finite Field*, Annals of Math. Studies **107**, Princeton Univ. Press, Princeton, 1984.

[22] D. Segal, *Words: notes on verbal width in groups*, London Math. Soc. Lecture Note Series **361**, Cambridge University Press, Cambridge, 2009.

[23] A. Shalev, Word maps, conjugacy classes, and a non-commutative Waring-type theorem, *Annals of Math.* **170** (2009), 1383–1416.

[24] P.H. Tiep and A. Zalesskii, Minimal characters of the finite classical groups, *Comm. Algebra* **24** (1996), 2093–2167.

[25] P.H. Tiep and A. E. Zalesskii, Some characterizations of the Weil representations of the symplectic and unitary groups, *J. Algebra* **192** (1997), 130–165.

[26] W.R. Unger, Computing the character table of a finite group, *J. Symbolic Comput.* **41** (2006), 847–862.