Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT

ELSEVIER

Journal of Algebra 268 (2003) 264–285

JOURNAL OF Algebra

www.elsevier.com/locate/jalgebra

Residual properties of the modular group and other free products [☆]

Martin W. Liebeck [a] and Aner Shalev [b,*]

[a] *Department of Mathematics, Imperial College, London SW7 2BZ, England, UK*
[b] *Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel*

Received 14 February 2002

Communicated by Efim Zelmanov

Abstract

Using a probabilistic approach we establish new residual properties of the modular group $PSL_2(\mathbb{Z})$, and of more general free products. We prove that the modular group is residually in any infinite collection of finite simple groups not containing a Suzuki group $Sz(q)$ or a 4-dimensional symplectic group $PSp_4(q)$ with q a power of 2 or 3. This result is best possible, since the groups excluded are not quotients of the modular group. We also show that if S is a collection of classical groups of unbounded rank, then an arbitrary free product $A * B$ of nontrivial finite groups, not both 2-groups, is residually S, and prove results about free products $A * \mathbb{Z}$.

1. Introduction

A group G is said to be residually (in) a set S of groups if the kernels of all epimorphisms from G to members of S intersect trivially. A well-known problem raised by Magnus [10] and Gorchakov and Levchuk [5] asks whether a free group of finite rank at least 2 is residually in any infinite collection of finite simple groups. Following partial solutions in [10,16,29,30], this was answered completely in the affirmative by Weigel in a series of papers [26–28]. Using a probabilistic approach, a short proof of this result is given in [4].

[☆] Research partially supported by a grant from the Israel Science Foundation for A.S.

[*] Corresponding author.
E-mail address: shalev@math.huji.ac.il (A. Shalev).

0021-8693/$ – see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0021-8693(03)00374-0

In this paper we show that the probabilistic approach can be applied in the study of residual properties of other classes of groups, and yields new results which were not obtained by standard tools. Here we focus on free products $\Gamma = A * B$ or $A * \mathbb{Z}$, where A, B are finite groups and $\mathbb{Z}$ denotes the integers. We are concerned with the analogue of the Magnus problem for these groups: namely, for which collections S of finite simple groups is Γ residually S? This is solved in [25] in the case where S consists of alternating groups, but all other cases have remained open until now.

Of particular interest is the free product $C_2 * C_3$ since this is isomorphic to the much studied modular group $PSL_2(\mathbb{Z})$; our first theorem gives a complete answer to the problem in this case.

Theorem 1.1. *Let S be an infinite collection of finite simple groups not containing $PSp_4(q)$ (q a power of 2 or 3) or $Sz(q)$. Then $PSL_2(\mathbb{Z})$ is residually S.*

Of course the Suzuki groups $Sz(q)$ do not contain elements of order 3, so must be excluded; the PSp_4 exceptions are also genuine, as these groups are not quotients of the modular group by [13, Section 6].

For general free products of finite groups, we prove

Theorem 1.2. *Let A, B be nontrivial finite groups, not both 2-groups, and let S be a collection of finite simple classical groups of unbounded ranks. Then the free product $A * B$ is residually S.*

Our last theorem concerns free products of finite groups with $\mathbb{Z}$.

Theorem 1.3. *Let S be an infinite collection of finite simple groups. Then*

(i) *$C_2 * \mathbb{Z}$ is residually S;*
(ii) *$C_3 * \mathbb{Z}$ is residually S, provided S does not contain $Sz(q)$;*
(iii) *if A is any nontrivial finite group, and S consists of classical groups of unbounded ranks, then $A * \mathbb{Z}$ is residually S.*

Note that part (iii) of Theorem 1.3 follows from Theorem 1.2; however, we provide a direct proof which requires fewer tools.

For simple groups of bounded rank such definitive results as Theorems 1.2 and 1.3(iii) are not always possible, as A, B may not be embeddable in such groups. Nevertheless as a by-product of our methods we are able to prove the following partial result (see Proposition 5.7): let p be a fixed prime, let A be a nontrivial subgroup of $PSL_2(p)$, and let S be a collection of simple groups of Lie type of bounded rank in characteristic p, not containing $Sz(q)$. Then the free product $A * \mathbb{Z}$ is residually S.

Let us briefly describe the general strategy of our proofs and the role of probabilistic arguments. In order to show that a group Γ is residually S it suffices to find, for each $1 \neq w \in \Gamma$, an epimorphism ϕ from Γ to some group $G \in S$ satisfying $\phi(w) \neq 1$. Consider first Theorem 1.1. Let $\Gamma = C_2 * C_3$, with canonical generators x, y of orders 2 and 3 respectively, and write $w = w(x, y)$. Then a homomorphism ϕ from Γ to G is determined by the values $a := \phi(x)$ and $b := \phi(y)$, where $a, b \in G$ satisfy $a^2 = b^3 = 1$. The condition that ϕ is an epimorphism can be written as $\langle a, b\rangle = G$, and the condition $\phi(w) \neq 1$ amounts to saying that $w(a, b) \neq 1$. The idea is now to let a be a randomly chosen involution in G, and b a randomly chosen element of order 3, and to show that each of these two conditions holds with probability tending to 1 as $|G| \to \infty$; hence the two conditions can be satisfied simultaneously for some group $G \in S$ as required. For the generation condition $\langle a, b\rangle = G$, we use random $(2, 3)$-generation results from [13] and [8]. The main effort therefore focuses on showing that, for random elements $a, b \in G$ of orders 2 and 3 respectively, the probability that $w(a, b) \neq 1$ tends to 1 as $|G| \to \infty$. For technical reasons we actually fix particular large conjugacy classes of elements of orders 2 and 3 in G and choose a, b at random from these classes; this means that we require slight refinements of the random generation results of [13]. We shall discuss below a little more our methods for proving the last statement about $w(a, b)$.

For Theorem 1.2, in which $\Gamma = A * B$, we shall prove a random (A, B)-generation result of independent interest (see Theorem 2.3 below). This result shows that, if A and B are embedded in a natural fashion in classical groups G of sufficiently large dimension, then randomly chosen G-conjugates of A and B generate G with probability tending to 1 as $|G| \to \infty$. The proof of this theorem relies on the recent paper [15], dealing with random generation of classical groups of large rank by elements of prime orders r, s (not both 2).

The layout of the paper is as follows. In Section 2 we present the results on random generation of simple groups which will be used in subsequent sections. Section 3 contains the proof of Theorems 1.1–1.3 in the case where the collection S consists of classical groups G of unbounded dimensions n. We study words $w(T) \in G * \langle T\rangle \cong G * \mathbb{Z}$, and their behaviour when we substitute a random element $t \in G$ for T. The main result of this section, Theorem 3.7, provides a criterion for $w(t)$ to be nontrivial with probability tending to 1 as $n \to \infty$. This result, when combined with relevant results on random generation in Section 2, yields Theorem 1.2 and the unbounded rank cases of Theorems 1.1 and 1.3.

Sections 4 and 5 are devoted to the proofs of Theorems 1.1 and 1.3 in the case where S consists of simple groups of bounded rank. We discuss Theorem 1.1 here. For the purpose of proving this, we may assume that the groups in S are of the form $S(q) = (G_{\sigma_q})'$, where $G = G(K_p)$ is a simple adjoint algebraic group of fixed type over an algebraically closed field K_p of characteristic p, σ_q is a Frobenius morphism, q is a power of p, and $q \to \infty$ (also p may vary). Our generation results in Section 2 show that if $a, b \in S(q)$ have orders 2, 3 respectively, and are such that $\dim a^G$, $\dim b^G$ are maximal, then $S(q)$ is generated by randomly chosen elements from the classes $a^{S(q)}, b^{S(q)}$ with probability tending to 1 as $q \to \infty$. We prove in Theorem 4.1 that G possesses a subgroup $PSL_2(K_p)$ containing such elements a, b. We further show that, provided K_p is the algebraic closure of a local field, this $PSL_2(K_p)$ contains the free product $C_2 * C_3$ (see Corollary 5.5). Combining this with a little algebraic geometry yields the desired fact that $w(a, b) \neq 1$ with probability tending to 1, and Theorem 1.1 follows quickly from this and the random generation result.

2. Probabilistic generation

In this section we present a number of results concerning probabilistic generation of finite simple groups. All are either taken from, or are easy consequences of, results in [6,7, 13–15].

For a finite group G, and subsets A, B of G, define $P_{A,B}(G)$ to be the probability that $\langle a, b\rangle = G$ for randomly chosen $a \in A$, $b \in B$. In other words,

$$P_{A,B}(G) = \frac{|\{(a,b) \in A \times B\colon\ \langle a,b\rangle = G\}|}{|A \times B|}.$$

Write also $P_{A,*}(G)$ instead of $P_{A,G}(G)$, and for $g \in G$ write $P_{g,B}(G) = P_{\{g\},B}(G)$ and $P_g(G) = P_{\{g\},*}(G)$. Thus $P_g(G)$ is the probability that for randomly chosen $t \in G$ we have $\langle g, t\rangle = G$.

If r is a positive integer, denote by $I_r(G)$ the set of elements of order r in G, and set $i_r(G) = |I_r(G)|$. As in [13], define

$$P_{r,s}(G) = P_{I_r(G),I_s(G)}(G), \qquad P_{r,*}(G) = P_{I_r(G),G}(G).$$

Thus $P_{r,s}(G)$ is the probability that randomly chosen elements of orders r, s in G generate G.

Now let G be simple. It is proved in [13, 1.1], [14, 1.1] that $P_{2,*}(G) \to 1$ as $|G| \to \infty$; in [13, 7.1(iii)], [14, 1.2] that, provided G is not a Suzuki group, $P_{3,*}(G) \to 1$ as $|G| \to \infty$; and in [13, 1.4], [8] that, provided G is not a Suzuki group or $PSp_4(q)$, we have $P_{2,3}(G) \to 1$ as $|G| \to \infty$. The next result is a slight refinement of these to the case where our simple group G is of fixed Lie type, and our random elements are chosen from specified large conjugacy classes of G.

Proposition 2.1. *Fix a Lie type X, and for each prime power q, let $X(q)$ be the finite simple group of type X over the field $\mathbb{F}_q$ (where q is an odd power of 2 or 3 for $X = {}^2F_4, {}^2B_2, {}^2G_2$). For each q, choose a simple adjoint algebraic group H over an algebraically closed field of characteristic $p = \mathrm{char}(\mathbb{F}_q)$, and a Frobenius morphism σ_q of H, such that $X(q) = (H_{\sigma_q})'$. Let $a, b \in H$ be elements of order 2, 3 respectively, such that the dimensions of the classes a^H, b^H are maximal.*

(A) *The class a^H intersects $X(q)$ nontrivially; so does the class b^H, provided $X(q) \neq {}^2B_2(q)$.*

(B) *Excluding the exception in (A), take $a, b \in X(q)$ and define*

$$A = a^{X(q)}, \qquad B = b^{X(q)}.$$

Then the following hold.

(i) $P_{A,*}(X(q)) \to 1$ *as* $q \to \infty$.

(ii) $P_{B,*}(X(q)) \to 1$ *as* $q \to \infty$, *provided* $X(q) \neq {}^2B_2(q)$.

(iii) $P_{A,B}(X(q)) \to 1$ *as* $q \to \infty$, *provided* $X(q) \neq PSp_4(q)$ *or* ${}^2B_2(q)$.
(iv) *Suppose* $X(q) = PSp_4(q)$ *with* $\operatorname{char}(\mathbb{F}_q) \neq 2, 3$ *for all* q; *then* $P_{A,B}(X(q)) \to 1$ *as* $q \to \infty$, *provided b has two eigenvalues* 1 *on the natural* 4*-dimensional module.*

Proof. (A) The classes of elements of orders 2, 3 of largest dimension in H are calculated in the proofs of [13, 4.1, 4.3], from which it is evident that each such class has a representative in $X(q)$, apart from elements of order 3 when $X(q)$ is a Suzuki group.

(B) For $X(q)$ classical, parts (i), (ii), (iii) are proved using [13, Proposition 2.6], as shown at the end of [13, §2]; and part (iv) is proved in the last half of the proof of [13, 6.3]. For $X(q)$ exceptional we see from the proof of [13, 4.3] that $|A| > (1 - o(1))i_2(X(q))$ and $|B| > (1 - o(1))i_3(X(q))$, whence (i), (ii) and (iii) follow from [14, 1.1], [14, 1.2] and [8] respectively. □

The next result is taken from [6, Theorem 1] and [7, Theorem 2].

Proposition 2.2. *Let G be a finite simple group of Lie type over* $\mathbb{F}_q$.

(i) *Then G contains a conjugacy class C such that* $P_{g,C}(G) > 1/10$ *for all* $1 \neq g \in G$.
(ii) *If we define* $P^-(G) = \min\{P_g(G)\colon 1 \neq g \in G\}$, *then* $P^-(G) \to 1$ *as* $q \to \infty$.

Note that the conclusion of (i) follows for quasisimple groups G as well, taking $g \notin Z(G)$.

If A is a finite group, k is a field, and V is a kA-module, we say that V is a *virtually free* kA-module if $V \downarrow A = F \oplus U$, where $F \neq 0$ is free and $\dim U < 2|A| + 4$. And if W is a vector space over k and $A \leqslant GL(W)$, we say A is embedded *virtually freely* in $GL(W)$ if W is virtually free as a kA-module. In such a situation, if $Z = Z(GL(W))$, then the image of A in $PGL(W)$ is $AZ/Z \cong A$, and we say also that A is embedded virtually freely in $PGL(W)$.

Note that any finite group A can be embedded virtually freely in any classical simple group G with natural module V of dimension with $n \geqslant 2|A| + 2$ over $\mathbb{F}_q$. One way of seeing this is as follows. If $G = PSL(V)$, write $n = m|A| + r$ with $r < |A|$, and regard V as an A-module of the form $(\mathbb{F}_q A)^m \oplus I_r$, where I_r denotes the trivial r-dimensional A-module. Otherwise, observe that $(\mathbb{F}_q A)^2$ admits A-invariant non-degenerate symplectic, orthogonal and unitary forms. By taking V to be a suitable direct sum of such modules with an appropriate trivial module we obtain a virtually free embedding of A in G.

We can now state the (A, B)-generation result referred to in the Introduction.

Theorem 2.3. *Let* A, B *be nontrivial finite groups, not both* 2*-groups. Then there exists a positive integer* $f(|A|, |B|)$ *such that the following holds. If G is a finite classical simple group of rank at least* $f(|A|, |B|)$, *and* A, B *are embedded virtually freely in G, then for randomly chosen* $t \in G$, *the probability that* $\langle A, B^t \rangle = G$ *tends to* 1 *as* $|G| \to \infty$.

Proof. Since A, B are nontrivial and not both 2-groups there are primes r, s not both 2 and elements $a \in A$ and $b \in B$ of orders r, s respectively. Let V be the natural module for G. Since A is embedded virtually freely in G, as an $\langle a \rangle$-module V has a free submodule of bounded codimension. Similarly, as a $\langle b\rangle$-module V has a free submodule of bounded codimension. The proof of the main result of [15] now shows that, assuming the dimension of G is large enough, random conjugates of a, b in G generate G with probability tending to 1 as $|G| \to \infty$. This means that, if $t \in G$ is chosen at random, then the probability that a, b^t generate G tends to 1, and so the probability that $\langle A, B^t\rangle = G$ also tends to 1 as $|G| \to \infty$. □

3. Groups of unbounded rank

Let $G = Cl_n(q) \leqslant SL_n(q)$ be a classical quasisimple group with natural module $V = V_n = (\mathbb{F}_q)^n$. Let $(\ ,\ )$ be the form on V fixed by G (bilinear if G is symplectic or orthogonal, sesquilinear if G is unitary, and identically zero if $G = SL_n(q)$); and if G is orthogonal, let Q be the quadratic form on V fixed by G.

In this section we consider the cases of Theorems 1.1–1.3 where the collection S of simple groups consists of classical groups of unbounded dimension. Hence in this section we assume that n is large.

We begin with an elementary result on linear algebra.

Let I denote the identity matrix in G. For a matrix $a \in G$ let $\mathrm{rk}(a)$ denote its rank, and set

$$\nu(a) = \min\bigl\{\mathrm{rk}(a - \lambda I)\colon\ \lambda \in F_q\bigr\}.$$

Lemma 3.1. *Let $n \geqslant 2d$. Let $a_1, \ldots, a_d \in G$, set $\nu_i = \nu(a_i)$ and let $\nu = \min\{\nu_1, \ldots, \nu_d\}$. Let $v_1, \ldots, v_d \in V$ be randomly chosen linearly independent vectors. Then the probability that the vectors $v_1, \ldots, v_d, v_1a_1, \ldots, v_da_d$ are linearly independent is at least $1 - q^{d-\nu}$.*

Proof. Suppose $v_1, \ldots, v_d, v_1a_1, \ldots, v_da_d$ are linearly dependent. Then there are scalars $\lambda_1, \ldots, \lambda_d, \mu_1, \ldots, \mu_d \in \mathbb{F}_q$, not all zero, such that $\sum_{i=1}^d \lambda_i v_i + \sum_{i=1}^d \mu_i v_i a_i = 0$. Since $v_1, \ldots, v_d$ are linearly independent there is i with $\mu_i \neq 0$.

The number of choices for $\mu_1, \ldots, \mu_d$ not all zero up to multiplication by a common scalar $\lambda \neq 0$ is $(q^d - 1)/(q - 1)$. Now, given $\mu_1, \ldots, \mu_d$, the scalars $\lambda_1, \ldots, \lambda_d$ are uniquely determined (since $v_1, \ldots, v_d$ are linearly independent), and we have

$$\sum_{i=1}^{d} v_i(\mu_i a_i + \lambda_i I) = 0.$$

This equation can be viewed as a system of n linear equations in dn variables (the coordinates of the vectors v_i) over $\mathbb{F}_q$. Let r denote the rank of the $n \times dn$ matrix of this system. Then $r \geqslant \min_{i=1}^d \mathrm{rk}(\mu_i a_i + \lambda_i I)$. There exists i with $\mu_i \neq 0$, so for this value of i we have

$$r \geqslant \mathrm{rk}(\mu_i a_i + \lambda_i I) \geqslant \nu(a_i) \geqslant \nu.$$

By standard linear algebra, the system above has q^{dn-r} solutions, so the probability that d randomly chosen vectors in V form a solution is $q^{-r} \leqslant q^{-\nu}$. Now, d randomly chosen vectors are linearly independent with probability at least $1 - q^{d-n}$. Therefore the probability that $v_1, \dots, v_d$ form a solution is at most $(1 - q^{d-n})^{-1} q^{-\nu}$. Summing over the choices for $\mu_1, \dots, \mu_d$ we see that the probability that $v_1, \dots, v_d, v_1 a_1, \dots, v_d a_d$ are linearly dependent is at most $(q^d - 1)/(q - 1) \cdot (1 - q^{d-n})^{-1} q^{-\nu}$. It is easy to see, using the inequality $2d \leqslant n$, that this expression is bounded above by $q^{d-\nu}$. The result follows. □

Let V_d denote the set of all d-tuples of linearly independent vectors $v_1, \dots, v_d$ in G. Clearly $(1 - o_n(1))|V|^d \leqslant |V_d| \leqslant |V|^d$.

Corollary 3.2. *Suppose d is fixed, $a_1, \dots, a_d \in G$, and $\nu(a_1), \dots, \nu(a_d) \to \infty$ as $n \to \infty$. Choose linearly independent vectors $v_1, \dots, v_d \in V$ at random. Then the probability that $v_1, \dots, v_d, v_1 a_1, \dots, v_d a_d$ are linearly independent tends to 1 as $n \to \infty$.*

Consider the free product $G * \mathbb{Z} = G * \langle T \rangle$, and let $w = w(T) \in G * \mathbb{Z}$ be a non-identity element. Then we can write

$$w = a_1 T^{k_1} a_2 T^{k_2} \cdots a_l T^{k_l},$$

where $a_i \in G$, $k_i \in \mathbb{Z}$, and $a_2, \dots, a_l \neq 1$, and $k_1, \dots, k_{l-1} \neq 0$. We call a_i the coefficients of w, and define $\nu(w) = \min(\nu(a_i)\colon a_i \neq 1)$.

For each $t \in G$, we define the specialisation $w(t) \in G$ to be the image of w under the homomorphism from $G * \langle T \rangle \to G$ induced by the identity map on G and the map sending T to t. Our aim is to show that if $\nu(w) \to \infty$, and $t \in G$ is chosen at random, then the probability that $w(t)$ is non-scalar tends to 1. Replacing w by $T^{-m} w T^m$ for $m \neq 0$, k_1 or $-k_l$, we may assume that $a_1 = 1$ and $k_l \neq 0$.

Let $d = |k_1| + \cdots + |k_l|$ (the degree of w with respect to T), and write $w = w_1 w_2 \cdots w_d$, where each $w_i = w_i(T) = g_i T^{\varepsilon_i}$, where $\varepsilon_i \in \{1, -1\}$ and $g_i \in G$ (possibly 1).

Let $i_1, \dots, i_e$ be the set of indices i, $1 \leqslant i \leqslant d$, for which $g_i \neq 1$. We will say that a sequence of vectors $v_1, \dots, v_{d+1}$ in V is *w-good* if the vectors $v_1, \dots, v_{d+1}, v_{i_1} g_{i_1}, \dots, v_{i_e} g_{i_e}$ are linearly independent. Further, we call such a sequence *w-feasible* if it is w-good, and there is $t \in G$ such that $v_i w_i(t) = v_{i+1}$ for all $i = 1, \dots, d$. Note that these equations in t take the form

$$(v_i g_i)t = v_{i+1} \text{ if } \varepsilon_i = 1, \quad \text{and} \quad v_{i+1} t = v_i g_i \text{ if } \varepsilon_i = -1 \quad (1 \leqslant i \leqslant d). \tag{†}$$

Clearly if the sequence $v_1, \dots, v_{d+1}$ is w-feasible, then (v_i, v_i) (and also $Q(v_i)$ if G is orthogonal) does not depend on i; in particular, if one of the v_i is singular, then so are all of them.

By Corollary 3.2, almost all sequences of vectors $v_1, \dots, v_{d+1}$ are w-good, provided $k_1, \dots, k_l$ are fixed, and each $\nu(g_{i_j})$ tends to infinity. By definition the elements g_{i_j} lie in the set $\{a_1, \dots, a_l\}$.

Corollary 3.3. *Let w be as above, and suppose $\nu(w) \to \infty$. Then the probability that a sequence $v_1, \dots, v_{d+1}$ of linearly independent vectors is w-good tends to 1. The same holds if v_1 is some fixed non-zero vector and only $v_2, \dots, v_{d+1}$ are chosen at random.*

Proof. The first statement is immediate from Corollary 3.2, and the second from the proof of Lemma 3.1. □

For $d \geqslant 1$, define $\alpha(d)$ to be 0, $\binom{d}{2}$, d^2, $\binom{d+1}{2}$ if $G = SL_n(q), Sp_n(q), SU_n(q^{1/2}), O_n(q)$, respectively.

Lemma 3.4. *Suppose the sequence $v_1, \ldots, v_{d+1} \in V$ is w-feasible and consists of singular vectors. Then the number of elements $t \in G$ satisfying $v_i w_i(t) = v_{i+1}$ for all $i = 1, \ldots, d$ is at least $(1 - o_n(1))|V|^{-d} q^{\alpha(d)}|G|$.*

Proof. The conditions on t can be written in the form $u_i t = u_i'$ $(1 \leqslant i \leqslant d)$ for suitable vectors u_i, u_i'. Let S be the set of elements $t \in G$ satisfying these conditions. Then S is non-empty by w-feasibility.

Now let $U = \langle u_1, \ldots, u_d \rangle$. Observe that $u_1, \ldots, u_d$ are linearly independent since $v_1, \ldots, v_{d+1}$ is w-good. Hence $\dim U = d$. Let H be the pointwise stabilizer of U in G. Then S is a coset of H, hence $|S| = |H|$. Proposition 14 of [4] shows that $|G : H| = (1 + o_n(1))|V|^d q^{-\alpha(d)}$, and the conclusion follows. □

We will make use of the following easy observation.

Lemma 3.5. *Let $g \in G$, and let $U \leqslant V$ be an f-dimensional subspace. Then the number of vectors $v \in V$ for which $vg \in \langle U, v \rangle$ is at most $q^{n+f+1-\nu(g)}$.*

Proof. If v is such a vector, then there is $\lambda \in \mathbb{F}_q$ such that $v(g - \lambda I) \in U$. There are q choices for λ, and given λ we have $\mathrm{rk}(g - \lambda I) \geqslant \nu(g)$. Hence the kernel of the map $g - \lambda I : V \to V$ has size at most $q^{n-\nu(g)}$, so the inverse image of U has size at most $q^f \cdot q^{n-\nu(g)}$. The conclusion follows. □

Lemma 3.6. *Let w be as above, and let $\nu = \nu(w)$. Fix a non-zero singular vector $v_1 \in V$. Then the number of sequences $v_2, \ldots, v_{d+1} \in V$ such that $v_1, \ldots, v_{d+1}$ is w-feasible exceeds $(1 - o_\nu(1))|V|^d q^{-\alpha(d)}$.*

Proof. In the case $G = SL_n(q)$ every w-good sequence is w-feasible, so the conclusion follows from Corollary 3.3. So suppose G is symplectic, unitary or orthogonal.

A sequence $v_1, \ldots, v_{d+1}$ is w-feasible if and only if there exists $t \in G$ such that Eqs. (†) hold, and also the set $Y = \{v_i, v_j g_j \colon 1 \leqslant i, j \leqslant d+1,\ g_j \neq 1\}$ is linearly independent. Observe that if the sequence is w-feasible, then Y consists of singular vectors, since v_1 is singular. Equations (†) are of the form $\alpha_i t = \beta_i$ $(1 \leqslant i \leqslant d)$, where $\alpha_i, \beta_i \in Y$; by Witt's lemma, for singular α_i, β_i, the existence of a solution $t \in G$ is equivalent to the system of equations

$$(\alpha_i, \alpha_j) = (\beta_i, \beta_j) \quad \text{for all } i, j. \tag{$\diamond$}$$

We now estimate the number of sequences $v_1, \ldots, v_d$ of singular vectors such that ($\diamond$) holds and Y is independent. This is done recursively as follows. The vector v_1 is already given. Suppose $1 \leqslant k \leqslant d$ and singular vectors $v_1, \ldots, v_k$ are given, such that $Y_k = \{v_i, v_j g_j\colon 1 \leqslant i, j \leqslant k,\ g_j \neq 1\}$ is linearly independent, and $(\alpha_i, \alpha_j) = (\beta_i, \beta_j)$ for all $1 \leqslant i, j < k$. To form the next vector in a w-feasible sequence, v_{k+1} has to satisfy the following:

(1) $(\alpha_i, \alpha_k) = (\beta_i, \beta_k)$ for $1 \leqslant i < k$,
(2) v_{k+1} is singular,
(3) Y_{k+1} is linearly independent.

The restrictions in (1) yield $k-1$ linear equations on v_{k+1}. Since the set Y_k is linearly independent, these equations are linearly independent, hence their solution space is some coset $u + U$ of a subspace $U \leqslant V$ of codimension $k-1$.

Now let N_k be the number of singular vectors in the coset $u + U$. Using [4, Proposition 11] we see that $N_k = q^{n-k+1}$ if G is symplectic, $N_k = (1 + o_n(1))q^{n-k+1/2}$ if G is unitary, and $N_k = (1 + o_n(1))q^{n-k}$ if G is orthogonal.

Obviously, the number of vectors v_{k+1} satisfying conditions (1) and (2) is precisely N_k. Condition (3) amounts to requiring that $v_{k+1} \notin \langle Y_k \rangle$, and also that $v_{k+1} g_{k+1} \notin \langle Y_k, v_{k+1} \rangle$ if $g_{k+1} \neq 1$.

Since $|Y_k| \leqslant 2k$ the first restriction above leaves us with at least $N_k - q^{2k}$ choices for v_{k+1}. If $g_{k+1} \neq 1$ then $\nu(g_{k+1}) \geqslant \nu$.

Define $M_k = N_k - q^{2k} - q^{n+2k+2-\nu}$. Using Lemma 3.5 with $g = g_{k+1}$, we see that there are at least M_k vectors $v_{k+1} \in V$ satisfying (1)–(3) above. Altogether it follows that there are at least $\prod_{k=1}^{d} M_k$ w-feasible sequences $v_1, \ldots, v_{d+1}$.

Note that $k \leqslant d$ is bounded, hence $M_k \geqslant (1 - o_\nu(1))N_k$. This implies that $\prod_{k=1}^{d} M_k \geqslant (1 - o_\nu(1)) \prod_{k=1}^{d} N_k$. Finally, we have $\prod_{k=1}^{d} N_k = (1 + o_n(1))|V|^d q^{-\alpha(d)}$. The result follows. □

We can now prove the main result of this section.

Theorem 3.7. *Let $w \in G * \mathbb{Z}$ be as above, and suppose $\nu(w) \to \infty$ as $n \to \infty$. Choose $t \in G$ at random. Then the probability that $w(t)$ is non-scalar tends to* 1.

Proof. Fix a non-zero singular vector $v_1 \in V$, and write $\nu = \nu(w)$.

Then there are at least $(1 - o_\nu(1))|V|^d q^{-\alpha(d)}$ w-feasible sequences $v_1, \ldots, v_{d+1}$. Now, for each w-feasible sequence $v_1, \ldots, v_{d+1}$, the number of $t \in G$ satisfying $v_i w_i(t) = v_{i+1}$ $(i = 1, \ldots, d)$ is at least $(1 - o_n(1))|V|^{-d} q^{\alpha(d)}|G|$ by Lemma 3.4. Summing up over the w-feasible sequences we see that at least

$$\bigl(1 - o_\nu(1)\bigr)|V|^d q^{-\alpha(d)} \cdot \bigl(1 - o_n(1)\bigr)|V|^{-d} q^{\alpha(d)}|G| = \bigl(1 - o_\nu(1)\bigr)|G|$$

elements $t \in G$ satisfy $v_1 w(t) = v_{d+1}$ for some $v_{d+1} \notin \langle v_1 \rangle$. Hence the probability that $w(t)$ is non-scalar tends to 1 as $\nu \to \infty$. □

Recall from Section 2 the definition of a virtually free embedding of a finite group in G.

Corollary 3.8. *Let A be a fixed nontrivial finite group. Let $1 \neq w \in A * \mathbb{Z} = A * \langle T \rangle$. For $t \in G$ let $\phi_t : A * \mathbb{Z} \to G$ be a homomorphism induced by embedding A in G virtually freely, and by sending T to t. Then, as $t \in G$ is chosen at random, the probability that $\phi_t(w)$ is non-scalar tends to* 1 *as $n \to \infty$.*

Proof. Let $A < G$ as above. As the embedding is virtually free, we have $\nu(a) \geqslant (n - 2|A| - 4)/|A|$ for all $1 \neq a \in A$. Therefore, if $n \to \infty$ so does $\nu(w)$. The required conclusion now follows from Theorem 3.7. □

For the applications we shall also need the following, slightly more technical, result.

Corollary 3.9. *Let A be a fixed nontrivial finite group. Let $1 \neq w \in A * \mathbb{Z} = A * \langle T \rangle$, and for $t \in G$ let $\phi_t : A * \mathbb{Z} \to G$ be as above. Fix a conjugacy class $C = x^G$ in G. Then, as $t \in C$ is chosen at random, the probability that $\phi_t(w)$ is non-scalar tends to* 1 *as $\nu(x) \to \infty$.*

Proof. We rewrite w by replacing T with $T^{-1}xT$. In this way we obtain a non-identity word in $G * \langle T \rangle$ such that its coefficients a_i are either non-identity elements of A, or x. We see that $\nu(a_i) \to \infty$ for all the coefficients a_i, yielding the result by Theorem 3.7. □

The final corollary concerns free products of arbitrary finite groups. Let A, B be fixed nontrivial finite groups. Fix embeddings $f : A \to G$, $g : B \to G$, such that $f(A)$ and $g(B)$ are virtually free in G.

Corollary 3.10. *For $t \in G$ let $\psi_t : A * B \to G$ be the homomorphism induced by sending $a \in A$ to $f(a) \in G$, and $b \in B$ to $g(b)^t \in G$. If $1 \neq w \in A * B$, and $t \in G$ is chosen at random, then the probability that $\psi_t(w)$ is non-scalar tends to 1 as $n \to \infty$.*

Proof. This follows in a similar manner: we rewrite $w = a_1b_1a_2b_2\ldots$ by replacing each b_i by $T^{-1}b_iT$. This gives a non-identity word in $w' \in G * \langle T \rangle$ such that $\nu(w') \geqslant \min((n - 2|A| - 4)/|A|, (n - 2|B| - 4)/|B|)$, hence $\nu(w') \to \infty$. The conclusion now follows from Theorem 3.7. □

At this point the proofs of Theorems 1.2 and 1.3(iii) can be quickly deduced. Observe that the cases of Theorems 1.1 and 1.3(i), (ii) where the collection S consists of simple groups of unbounded rank follow immediately from Theorems 1.2 and 1.3(iii).

Proof of Theorem 1.2. Assume the hypothesis of Theorem 1.2. Given $1 \neq w \in A * B$, it suffices to find a group $H \in S$ and an epimorphism from $A * B$ to H sending w to a non-identity element. For $H \in S$ let $G = Cl_n(q) \leqslant SL_n(q)$ be a quasisimple group with $G/Z(G) = H$. For $t \in G$, define $\psi_t : A * B \to G$ as in Corollary 3.10, and let $\bar{\psi_t} : A * B \to H$ be the composition of ψ_t with the canonical map $\pi : G \to H$. By Corollary 3.10, the probability that $\bar{\psi_t}(w) \neq 1$ tends to 1 as $n \to \infty$. The image of $\bar{\psi_t}$ is $\langle \pi(f(A)), \pi(g(B)^t) \rangle$, and by Theorem 2.3 this image is equal to H with probability tending to 1 as $n \to \infty$. Hence if n is large enough, there exists $t \in G$ such that $\bar{\psi_t}$ is an epimorphism and $\bar{\psi_t}(w) \neq 1$, as required. □

Proof of Theorem 1.3(iii). Assume the hypothesis, and let $1 \neq w \in A * \mathbb{Z}$. As above, for $H \in S$ let $G = Cl_n(q)$ with $G/Z(G) = H$. By Proposition 2.2(i) and the remark following it, G has conjugacy class $C = x^G$ such that $P_{g,C}(G) > 1/10$ for all $g \in G\backslash Z(G)$; moreover, we see from [6] that $\nu(x) \to \infty$ as $n \to \infty$. For $t \in C$ define $\phi_t : A * \mathbb{Z} \to G$ as in Corollary 3.9. By Corollary 3.9 the probability that $\phi_t(w)$ is non-scalar tends to 1; and by Proposition 2.2(i), for at least $1/10$ of the elements $t \in C$ we have $\langle A, t\rangle = G$. Hence, if n is sufficiently large, both conditions hold for some t. Therefore ϕ_t is an epimorphism from $A * \mathbb{Z}$ to G sending w to a non-scalar. The result follows by composing ϕ_t with the canonical map $G \to H$. □

4. Groups of bounded rank, I: a preliminary result

In our proof of the bounded rank cases of Theorems 1.1 and 1.3, we shall require the following result, showing that in any classical algebraic group G, a subgroup of type A_1 can be found which contains involutions and elements of order 3 lying in G-classes of maximal dimension.

Theorem 4.1. *Let G be a simple adjoint algebraic group over an algebraically closed field K of characteristic p, and let σ be a Frobenius morphism of G. Assume that the fixed point group G_σ is not a Suzuki group ${}^2B_2(q)$. Then there exist elements $a, b \in G_\sigma$ with a of order* 2 *and b of order* 3, *such that the following hold*:

(i) $\dim a^G$ *and* $\dim b^G$ *are maximal among the dimensions of G-conjugacy classes of elements of order* 2 *and* 3, *respectively, and*
(ii) *there is an embedding $\phi : PSL_2(K) \to G$ such that* $\mathrm{Im}(\phi)$ *intersects both a^G and b^G nontrivially.*

Proof. *Case* 1: *G classical.* Assume that G is classical. The classes of largest dimension of elements of order 2 or 3 in G are given by [13, 4.1] and its proof. We record here the dimensions $k(G)$, $l(G)$ of the largest classes of elements of order 2, 3 respectively:

G	$k(G)$	$l(G)$
$PSL_n(K)$	$[n^2/2]$	$[2n^2/3]$
$PSp_{2m}(K),\ PSO_{2m+1}(K)$	$m^2 + m$	$[2(2m^2 + m)/3]$
$PSO_{2m}(K)$	$m^2 + 1 - (2, m)$	$[2(2m^2 - m)/3]$

In all cases there are either one or two classes of involutions of dimension $k(G)$, and one or two classes of elements of order 3 of dimension $l(G)$; suitable elements in each of these classes can be read off from the proof of [13, 4.1]. By Proposition 2.1(A), each class has a representative in G_σ.

Let V be the natural module associated with G, so that $G = PSL(V)$, $PSp(V)$ or $PSO(V)$.

Write $A = SL_2(K)$. We adopt the usual notation for irreducible KA-modules: for $0 \leqslant r \leqslant p - 1$, denote by $V_A(r)$ the irreducible KA-module with high weight r. This module has dimension $r + 1$, and can be realised as the space $H_A(r)$ of all homogeneous polynomials of degree r in two variables, with the natural SL_2-action. If T is a maximal torus of A, the weights of T on $V_A(r)$ are $r, r-2, r-4, \ldots, -r$. If r is odd then $V_A(r)$ is a faithful symplectic module for $A = SL_2(K)$; and if r is even then $V_A(r)$ is an orthogonal module for $A/Z(A) = PSL_2(K)$.

(A) Assume first that $p \geqslant 3$, $\dim V = 2m$ is even, and $G = PSL(V)$ or $PSp(V)$. In this case, define a 6-dimensional KA-module V_6 as follows:

$$\begin{aligned} V_6 &= V_A(3) \oplus V_A(1) \quad \text{if } p \geqslant 5, \quad \text{and} \\ V_6 &= V_A(1) \otimes V_A(2) \quad \text{if } p = 3. \end{aligned}$$

Observe that V_6 is a symplectic module for A (with the direct sum being a perpendicular sum). Write

$$a = \operatorname{diag}(i, -i), \ b = \operatorname{diag}\bigl(\omega, \omega^{-1}\bigr) \in A,$$

where i is a fourth root of unity and ω a cube root of unity in K. Then

$$\begin{aligned} a^{V_6} &= \operatorname{diag}(i, i, i, -i, -i, -i), \quad \text{and} \\ b^{V_6} &= \begin{cases} \operatorname{diag}(\omega, \omega, \omega^2, \omega^2, 1, 1) & (p > 3), \\ \operatorname{diag}(J_3, J_3) & (p = 3) \end{cases} \end{aligned}$$

(where J_3 denotes a Jordan block of size 3).

Now let $2m = 6k + r$ with k an integer and $r = 0, 2$ or 4. Define an r-dimensional A-module V_r as follows:

r	V_r
0	0
2	$V_A(1)$
$r = 4,\ p > 3$	$V_A(3)$
$r = 4,\ p = 3,\ G = PSL(V)$	$H_A(3)$
$r = 4,\ p = 3,\ G = PSp(V)$	$V_A(1)^2$

(where as above, $H_A(3)$ denotes the space of homogeneous polynomials of degree 3 in two variables with the natural A-action). Finally, let

$$V = (V_6)^k \oplus V_r.$$

Then V is a KA-module of dimension $2m$ admitting a non-degenerate A-invariant symplectic form, and such that $Z(A)$ acts as $\langle -1 \rangle$ on V. The representation of A on V therefore embeds $PSL_2(K)$ in G.

It remains to show that a^V and b^V belong to the G-classes of maximal dimension among elements of orders 2, 3. If $r = 0$ then

$$a^V = \left(i^{(3k)}, -i^{(3k)}\right),$$
$$b^V = \begin{cases} (1^{(2k)}, \omega^{(2k)}, \omega^{-1\,(2k)}) & (p > 3), \\ (J_3^{(2k)}) & (p = 3), \end{cases}$$

where bracketed superscripts indicate multiplicities. If $G = PSL(V)$ then $\dim a^G = \dim GL_{6k} - 2\dim GL_{3k} = \frac{1}{2}(\dim V)^2$; and if $p > 3$ then $\dim b^G = \dim GL_{6k} - 3\dim GL_{2k}$, while if $p = 3$ then $\dim b^G = \dim GL_{6k} - 2(2k)^2 - \dim GL_{2k}$ (see the proof of [13, 4.1] for formulae for the dimensions of centralizers of unipotent elements of order 3), whence $\dim b^G = \frac{2}{3}(\dim V)^2$. Hence from the table above we see that $\dim a^G$ and $\dim b^G$ are maximal, as required. Similar calculations give the conclusion when $G = PSp(V)$: here $C_G(a)^0 = GL_{3k}$, while if $p > 3$ then $C_G(b) = GL_{2k}Sp_{2k}$ and if $p = 3$ then $\dim C_G(b) = (2k)^2 + \dim Sp_{2k}$ (see [13, 4.1] again).

Likewise, if $r = 2$ then

$$a^V = \left(i^{(3k+1)}, -i^{(3k+1)}\right),$$
$$b^V = \begin{cases} (1^{(2k)}, \omega^{(2k+1)}, \omega^{-1\,(2k+1)}) & (p > 3), \\ (J_3^{(2k)}, J_2) & (p = 3), \end{cases}$$

and again we check that $\dim a^G, \dim b^G$ are maximal.

Finally, let $r = 4$. Then $a^V = (i^{(3k+2)}, -i^{(3k+2)})$; if $p > 3$ then

$$b^V = \left(1^{(2k+2)}, \omega^{(2k+1)}, \omega^{-1\,(2k+1)}\right);$$

and if $p = 3$ then

$$b^V = \left(J_3^{(2k+1)}, J_1\right) \quad \text{or} \quad \left(J_3^{(2k)}, J_2^{(2)}\right),$$

according as $G = PSL(V)$ or $PSp(V)$ respectively. Once again we calculate that $\dim a^G$ and $\dim b^G$ are as in the above table, hence are maximal.

The completes the proof in case (A).

(B) Assume in this case that $p \geqslant 3$ and either $G = PSO(V)$, or $G = PSL(V)$ with $\dim V$ odd (this covers all cases with $p \geqslant 3$ remaining after (A)).

Define a 12-dimensional KA-module V_{12} as follows:

$$V_{12} = V_A(4) \oplus V_A(2)^2 \oplus V_A(0) \quad \text{if } p \geqslant 5,$$
$$V_{12} = \left(V_A(2) \otimes V_A(2)\right) \oplus V_A(2) \quad \text{if } p = 3.$$

Then

$$a^{V_{12}} = \left(1^{(6)}, -1^{(6)}\right) \quad \text{and}$$
$$b^{V_{12}} = \begin{cases} (1^{(4)}, \omega^{(4)}, \omega^{-1\,(4)}) & (p > 3), \\ (J_3^{(4)}) & (p = 3). \end{cases}$$

Observe that V_{12} is an orthogonal module for $A/Z(A) = PSL_2(K)$.

Now for $0 \leqslant r \leqslant 11$, define an r-dimensional A-module V_r as follows:

r	V_r
0	0
1	$V_A(0)$
2	$V_A(0)^2$
3	$V_A(2)$
4	$V_A(2) \oplus V_A(0)$
5	$V_A(4),\ p \geqslant 5$
	$H_A(4),\ p = 3,\ G = PSL(V)$
	$V_A(2) \oplus V_A(0)^2,\ p = 3,\ G = PSO(V)$
6	$V_A(2)^2$
7	$V_A(2)^2 \oplus V_A(0)$
8	$V_A(2)^2 \oplus V_A(0)^2$
9	$V_A(2) \otimes V_A(2)$
10	$V_A(2)^3 \oplus V_A(0)$
11	$V_A(4) \oplus V_A(2)^2,\ p \geqslant 5$
	$H_A(4) \oplus V_A(2)^2,\ p = 3,\ G = PSL(V)$
	$V_A(2)^3 \oplus V_A(0)^2,\ p = 3,\ G = PSO(V)$

Let $\dim V = 12k + r$ with $0 \leqslant r \leqslant 11$, and define an A-action on V by setting

$$V = \left(V_{12}^k\right) \oplus V_r.$$

Then the representation of A on V embeds $A/Z(A) = PSL_2(K)$ in G. Calculating as above with a^V and b^V (and recalling that r is odd if $G = PSL(V)$ in this case (B)), we find that $\dim a^G$ and $\dim b^G$ are maximal.

(C) In this case suppose that $p = 2$ and $G = PSL(V)$ or $PSp(V)$. Again let $A = SL_2(K)$, with

$$a = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},\quad b = \operatorname{diag}\left(\omega, \omega^{-1}\right) \in A,$$

elements of orders 2 and 3. Define a 6-dimensional A-module V_6 by

$$V_6 = \left(V_A(1) \otimes V_A(1)\right) \oplus V_A(1).$$

This is a symplectic (but not orthogonal) A-module, and

$$a^{V_6} = \left(J_2^{(3)}\right), \qquad b^{V_6} = \left(1^{(2)}, \omega^{(2)}, \omega^{-1(2)}\right).$$

Consider the embedding of A in $Sp(V_6) = Sp_6$. In the notation of [1, §7], a^{V_6} lies in the class represented by the element b_3. By [1, 7.7], it follows that $V_6(a) = \{v \in V_6\colon (v, va) = 0\} \neq V_6$. Hence if we embed V_6 in a symplectic A-module of the form $V = V_6 \perp W$, then the involution $t = a^{V_6} \perp a^W$ satisfies $V(t) \neq V$, hence by [1, 7.7] lies in class b_k or c_k of $Sp(V)$, for some k. Note that the largest involution classes in $Sp_{2m}(K)$

are represented by b_m if m is odd and by c_m if m is even (see the proof of [13, 4.1]; these elements act as matrices $(J_2^{(m)})$.

For $0 \leqslant r \leqslant 5$ define an A-module V_r as follows:

r	V_r
0	0
1	$V_A(0)$
2	$V_A(1)$
3	$V_A(1) \oplus V_A(0)$
4	$V_A(1) \otimes V_A(1)$
5	$V_A(1)^2 \oplus V_A(0)$

For r even, V_r is a symplectic A-module. Let $\dim V = 6k + r$ with $0 \leqslant r \leqslant 5$, and define an A-action on V by setting $V = V_6^k \oplus V_r$. This action is symplectic when r is even, and using the above remarks on the classes b_m, c_m in the case where G is symplectic, we calculate that $\dim a^G$ and $\dim b^G$ are maximal.

(D) It remains to handle the case where $p = 2$ and $G = SO(V)$. Define a 12-dimensional A-module V_{12} by

$$V_{12} = \big(V_A(1) \otimes V_A(1)\big)^2 \oplus V_A(1)^2.$$

This is an orthogonal module (the two $V_A(1)$s being totally singular subspaces), with $a^{V_{12}} = (J_2^{(6)})$ and $b^{V_{12}} = (1^{(4)}, \omega^{(4)}, \omega^{-1(4)})$. Calculation shows that in the action of a on the 4-dimensional orthogonal module $V_4 = V_A(1) \otimes V_A(1)$ we have $V_4(a) \neq V_4$. Hence in the notation of [1, §8] we have a^{V_4} conjugate to c_2, and so by previous observations, $a^{V_{12}}$ is conjugate to $c_6 \in SO_{12}$.

For r even with $0 \leqslant r \leqslant 10$, define an A-module V_r of dimension r as follows:

r	V_r
0	0
2	$V_A(0)^2$
4	$V_A(1) \otimes V_A(1)$
6	$V_A(1)^2 \oplus V_A(0)^2$
8	$(V_A(1) \otimes V_A(1))^2$
10	$(V_A(1) \otimes V_A(1)) \oplus V_A(1)^2 \oplus V_A(0)^2$

Let $\dim V = 12k + r$ with r even and $0 \leqslant r \leqslant 10$, and define an A-action on V by setting $V = V_{12}^k \oplus V_r$. This makes V an orthogonal A-module, and the usual calculations show that $\dim a^G$ and $\dim b^G$ are maximal.

This completes the proof of the theorem in the case where G is classical.

Case 2: *G exceptional*. Assume now that G is of exceptional type G_2, F_4, E_6, E_7 or E_8. Then G has a maximal rank subgroup $D = A_2$, A_1C_3, A_1A_5, A_2A_5 or A_4A_4 respectively. Write $A = SL_2(K)$, with elements a, b as above. We define an embedding $\phi : A \to D$ as follows. For $G \neq G_2$, let R_1, R_2 be the simple factors of D, so that $D \cong (R_1 \times R_2)/Z$ for some central subgroup Z. In the table below we specify representations $\phi_i : A \to R_i$; then for $a \in A$, $\phi(a)$ is defined to be the image modulo Z of $(\phi_1(a), \phi_2(a))$. For $i = 1, 2$ the representation ϕ_i is specified by giving the restriction to A of the natural module for R_i. (For $G = G_2$, $D = A_2$ we give the restriction of the natural A_2-module.)

G	D	Representations ϕ_1, ϕ_2
G_2	A_2	$V_A(2)\ (p \geqslant 3)$ $V_A(1) \oplus V_A(0)\ (p = 2)$
$F_4,\ E_6$	$A_1C_3,\ A_1A_5$	$V_A(1),\ V_A(5)\ (p > 5)$ $V_A(1),\ V_A(1) \otimes V_A(2)\ (p = 3 \text{ or } 5)$ $V_A(1),\ (V_A(1) \otimes V_A(1)) \oplus V_A(1)\ (p = 2)$
E_7	A_2A_5	$V_A(2),\ V_A(5)\ (p > 5)$ $V_A(2),\ V_A(1) \otimes V_A(2)\ (p = 3 \text{ or } 5)$ $V_A(1) \oplus V_A(0),\ (V_A(1) \otimes V_A(1)) \oplus V_A(1)\ (p = 2)$
E_8	A_4A_4	$V_A(4),\ V_A(4)\ (p \geqslant 5)$ $H_A(4),\ H_A(4)\ (p = 3)$ $V_A(1) \oplus V_A(1) \oplus V_A(0),\ V_A(1) \oplus V_A(1) \oplus V_A(0)\ (p = 2)$

The restriction of the adjoint module $L(G)$ to D is given by [20, 1.8], as follows, where we just give the high weights of the relevant irreducibles:

G	D	$(L(G)/L(D)) \downarrow D$
G_2	A_2	$\lambda_1 \oplus \lambda_2$
F_4	A_1C_3	$1 \otimes \lambda_3$
E_6	A_1A_5	$1 \otimes \lambda_3$
E_7	A_2A_5	$(\lambda_1 \otimes \lambda_2) \oplus (\lambda_2 \otimes \lambda_4)$
E_8	A_4A_4	$(\lambda_1 \otimes \lambda_2) \oplus (\lambda_2 \otimes \lambda_1) \oplus (\lambda_4 \otimes \lambda_3) \oplus (\lambda_3 \otimes \lambda_4)$

(Note that the prime restrictions in the hypothesis of [20, 1.8] are present just to ensure the complete reducibility of $L(D)$, and do not affect the proof otherwise.) To calculate with the above modules, note that the irreducible for A_n with high weight λ_i is just the ith alternating power of the natural module; and for C_3 the irreducible with high weight λ_3 has dimension 14, and is the alternating cube of the natural module factored out by the natural module.

From these restrictions we see immediately that the image $\phi(A)$ of A in G is $PSL_2(K)$, and we can calculate the actions of the elements a and b on $L(G)$. When these elements are semisimple (i.e., $p \neq 2$, $p \neq 3$ respectively), we find that $\dim C_{L(G)}(a) = 6$, 24, 38, 63, 120 and $\dim C_{L(G)}(b) = 4$, 16, 24, 43, 80, according as $G = G_2$, F_4, E_6, E_7, E_8 respectively. It follows that $C_G(a)^0 = A_1A_1$, A_1C_3, A_1A_5, A_7, D_8 and $C_G(b)^0 = A_1T_1$, A_2^2, A_2^3, A_2A_5, A_8 respectively. These are well known to be the smallest dimensional centralizers of semisimple elements of orders 2 and 3 (see, for example, [12, 1.2]). The corresponding classes a^G, b^G are fixed by any Frobenius morphism σ, and hence by [24, I, 3.4] have a representative in G_σ.

To conclude, suppose a or b is unipotent (so $p = 2$ or 3). Calculating from the embedding of $\phi(A)$ in D and the restriction $L(G) \downarrow D$, we find the Jordan block sizes of a and b on $L(G)$. From this information, the tables in [11] specify the unipotent classes of G containing a, b, which are as follows:

	$G = G_2$	F_4	E_6	E_7	E_8
$p = 2$, a in class	$\tilde{A}_1$	$A_1 + \tilde{A}_1$	$3A_1$	$4A_1$	$4A_1$
$p = 3$, b in class	$G_2(a_1)$	$\tilde{A}_2 + A_1$	$2A_2 + A_1$	$2A_2 + A_1$	$2A_2 + 2A_1$

(For $G = F_4$, $p = 3$ we also need to calculate with the action on the 25-dimensional module $V_{F_4}(\lambda_4)$ to show that b is in class $\tilde{A}_2 + A_1$ rather than $\tilde{A}_1 + A_2$.) It now follows from the classification of unipotent classes and centralizers in [3,18,19,22,23] that the classes a^G, b^G have maximal dimension. Finally, these classes are all fixed by σ, hence have a representative in G_σ. This completes the proof of the theorem. □

5. Groups of bounded rank, II: proof of Theorems 1.1 and 1.3

In this section we prove Theorems 1.1 and 1.3(i), (ii) in the case where the collection S consists of simple groups of Lie type of bounded rank.

The first few lemmas show that the relevant free products $C_2 * C_3$, $C_2 * \mathbb{Z}$ and $C_3 * \mathbb{Z}$ lie in $PSL_2(F)$ for any local field F. The proofs then proceed by combining this with Theorem 4.1 and the results from Section 2 on probabilistic generation.

Lemma 5.1. *Let A, B be groups and let $C \leqslant A \cap B$ be a common subgroup. Consider the free product with amalgamation $D = A *_C B$. Let $A_1 \leqslant A$ and $B_1 \leqslant B$ be subgroups satisfying $A_1 \cap C = B_1 \cap C = 1$. Then the subgroup of D generated by A_1 and B_1 is the free product $A_1 * B_1$.*

Proof. This follows immediately from the normal form theorem for free products with amalgamation; see, for instance, [17, Chapter IV, Theorem 2.6]. □

Now let F be a local field, O its ring of integers, and π the maximal ideal in O. Define

$$C = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in SL_2(O):\ a_{21} \in \pi \right\} \leqslant SL_2(O),$$

and let $\overline{C} < PSL_2(O)$ be the image of C modulo $Z(SL_2(O))$.

Lemma 5.2. *With the above notation we have*

$$PSL_2(F) \cong PSL_2(O) *_{\overline{C}} PSL_2(O).$$

Proof. By a result of Ihara (see Corollary 1 on p. 79 of [21]) we have $SL_2(F) \cong SL_2(O) *_C SL_2(O)$. Factoring out the centers yields the result. □

Corollary 5.3. *$PSL_2(\mathbb{Z})$ can be embedded in $PSL_2(F)$ for any local field F.*

Proof. Define

$$a_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \; b_0 = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \in SL_2(O).$$

Then the images a, b of a_0, b_0 in $PSL_2(O)$ do not lie in the subgroup C. Applying Lemma 5.1 we see that in $PSL_2(O) *_{\overline{C}} PSL_2(O)$, taking a and b on different sides, we have $\langle a, b\rangle \cong \langle a\rangle * \langle b\rangle \cong C_2 * C_3 \cong PSL_2(\mathbb{Z})$. The result follows using Lemma 5.2. □

Lemma 5.4.

(i) *Let $C_2 * C_3 = \langle x\rangle * \langle y\rangle$ (with x of order 2 and y of order 3). Then*

$$\langle x, yxy\rangle = \langle x\rangle * \langle yxy\rangle \cong C_2 * \mathbb{Z} \quad \textit{and}$$
$$\langle xyx, yxy\rangle = \langle xyx\rangle * \langle yxy\rangle \cong C_3 * \mathbb{Z}.$$

In particular, $C_2 * \mathbb{Z}$ and $C_3 * \mathbb{Z}$ can be embedded in $PSL_2(F)$ for any local field F.

(ii) *If F is a local field of characteristic p such that $|O/\pi| > p^2$, then $PSL_2(F)$ has a subgroup isomorphic to $PSL_2(p) * \mathbb{Z}$.*

Proof. Part (i) follows from the normal form for elements of $C_2 * C_3$.

Now consider part (ii). We claim that there is a subgroup $L < PSL_2(O)$ with $L \cong PSL_2(p)$ and $L \cap \overline{C} = 1$ (where O and $\overline{C}$ are as above). To see this, observe that by hypothesis, $PSL_2(O)$ has a subgroup $M \cong PSL_2(q)$ with $q > p^2$, such that $M \cap \overline{C}$ is the image modulo scalars of $\{(\begin{smallmatrix} a & b \\ 0 & a^{-1}\end{smallmatrix})\colon a \in \mathbb{F}_q^*, \; b \in \mathbb{F}_q\}$. In other words, $M \cap \overline{C} = M_{\langle v\rangle}$, where $v = (0, 1)$. Now let L_0 be the copy of $PSL_2(p)$ in M consisting of matrices with entries in $\mathbb{F}_p$. The condition that a matrix in L_0 fix a 1-space $\langle(1, \alpha)\rangle$ with $\alpha \in \mathbb{F}_q$ implies that α satisfies a quadratic equation over $\mathbb{F}_p$. As $q > p^2$, we can therefore find $\alpha \in \mathbb{F}_q$ such that the stabilizer in L_0 of $\langle(1, \alpha)\rangle$ is 1. Consequently there is a conjugate L of L_0 such that $L_{\langle v\rangle} = 1$. Then $L \cap \overline{C} = 1$, proving the claim.

We can also find an involution $b \in PSL_2(O)$ with $b \notin \overline{C}$. Hence by Lemma 5.2 we have $L * \langle b\rangle \leqslant PSL_2(F)$. Pick $l \in L$ with l of order at least 3.

We claim finally that the subgroup $\langle L^b, lbl\rangle$ is equal to the free product $L^b * \langle lbl\rangle \cong PSL_2(p) * \mathbb{Z}$. This again follows from the normal form for elements of free products. □

Since $PSL_2(F)$ can be embedded in any simple (untwisted) Chevalley group $G(F)$, we obtain the following, which may be of some independent interest.

Corollary 5.5. *Let $G(F)$ be a simple Chevalley group over a local field F. Then $G(F)$ contains the free products $C_2 * C_3$, $C_2 * \mathbb{Z}$, $C_3 * \mathbb{Z}$. If F has characteristic p and the residue field has order at least p^3, then $G(F)$ contains $PSL_2(p) * \mathbb{Z}$.*

We shall apply the above results in the case where $F = \mathbb{F}_p((t))$ with p a prime. Let K_p be the algebraic closure of this field.

In our proofs of Theorems 1.1 and 1.3 we shall make use of the following elementary result from algebraic geometry (see, for example, [9, 2.18]). For this result we need to set up notation for Frobenius morphisms of simple algebraic groups rather precisely, as follows. Let G be a simple algebraic group over an algebraically closed field K of characteristic p, generated by root elements $x_\alpha(t)$ (α a root, $t \in K$) in the standard way (see, for example, [2]). For q a power of p let ϕ_q be the field morphism of G which sends $x_\alpha(t) \to x_\alpha(t^q)$ for all α, t. Finally, define $\tau : G \to G$ to be either 1, or a nontrivial graph endomorphism of G commuting with ϕ_q, as in [2]. Thus for $G = A_n, D_n$ or E_6 we have $\tau^2 = 1$ (or $\tau^3 = 1$ for $G = D_4$), while for $(G, p) = (F_4, 2), (B_2, 2)$ or $(G_2, 3)$ we have $\tau^2 = \phi_p$. A general Frobenius endomorphism of G is a conjugate of $\sigma_q = \tau\phi_q$ for some τ, ϕ_q as above.

Lemma 5.6. *Let G, σ_q be as above, and suppose that $V \subset G$ is a σ_q-stable subvariety which is defined over K by at most e polynomial equations of degree at most f. Then there is a constant $c = c(e, f, \dim G)$ such that*

$$|V_{\sigma_q}| < cq^{\dim V}.$$

We are now ready to prove Theorems 1.1 and 1.3(i), (ii) in the case where the collection S consists of simple groups of Lie type of bounded rank.

Proof of Theorem 1.1 (*for bounded rank case*). It is sufficient to establish that $PSL_2(\mathbb{Z})$ is residually in any infinite set of simple groups of the form $X(q) \neq PSp_4(2^f)$, $PSp_4(3^f)$, $Sz(2^f)$, where X is a fixed Lie type (possibly twisted) and $q \to \infty$. Such groups $X(q)$ are of the form $(G_{\sigma_q})'$, where $G = G(K_p)$ is an adjoint simple algebraic group of fixed type over K_p, σ_q is a Frobenius morphism (as defined in the preamble to Lemma 5.6), and q is a power of p. (Recall that K_p denotes the algebraic closure of the local field $\mathbb{F}_p((t))$.) Note that as $q \to \infty$ the prime p may vary.

Consider such a group $X(q) = (G_{\sigma_q})'$. By Theorem 4.1 there is a subgroup Y of G with $Y \cong PSL_2(K_p)$ such that $Y \cap X(q)$ contains elements a, b of orders 2, 3 and $\dim a^G, \dim b^G$ are maximal among the dimensions of G-classes of elements of these orders. Moreover, by the proof of Corollary 5.3 there exists $y \in Y$ such that $\langle a, b^y\rangle \cong \langle a\rangle * \langle b^y\rangle \cong C_2 * C_3$.

Fix a nontrivial word $1 \neq w = w(Y, Z) \in C_2 * C_3$. Then $w(a, b^y) \neq 1$. Define

$$V = \bigl\{t \in G\colon\ w\bigl(a, b^t\bigr) = 1\bigr\}.$$

Then V is a subvariety of G, and is proper since $y \notin V$. Hence $\dim V < \dim G$. By definition of V, the number of polynomial equations over K_p defining V, and their degrees, depend only on the word w and the type of G (and not on p). Hence from Lemma 5.6 we see that there is a constant $c = c(w, \dim G)$ such that

$$|V_{\sigma_q}| < cq^{\dim V}.$$

We have $\dim V < \dim G$, and from the order formulae for simple groups, $|X(q)| < c'q^{\dim G}$ for some absolute constant c'. It follows that

$$\frac{|\{t \in X(q)\colon\ w(a,b^t) \neq 1\}|}{|X(q)|} \geqslant 1 - c_1 q^{\dim V - \dim G} \geqslant 1 - c_2 q^{-1} \to 1 \quad \text{as } q \to \infty. \tag{1}$$

Now Proposition 2.1(B)(iii) shows that, provided $X(q) \neq PSp_4(q)$, we have

$$\frac{|\{t \in X(q)\colon\ \langle a,b^t\rangle = X(q)\}|}{|X(q)|} \to 1 \quad \text{as } q \to \infty. \tag{2}$$

When $X(q) = PSp_4(q)$, we have $p > 3$ by the hypothesis of Theorem 1.1; the proof of Theorem 4.1 shows that the element b is conjugate to $\mathrm{diag}(\omega, \omega^{-1}, 1, 1)$ (where ω is a cube root of 1), and hence Proposition 2.1(B)(iv) shows (2) holds in this case as well.

From (1) and (2), it follows that if q is large enough, there exists $t \in X(q)$ such that $\langle a, b^t\rangle = X(q)$ and $w(a,b^t) \neq 1$.

Thus we have shown that $PSL_2(\mathbb{Z}) \cong C_2 * C_3$ is residually $X(q)$, completing the proof of Theorem 1.1. □

Proof of Theorem 1.3(i), (ii) (*for bounded rank case*). Again it is enough to show that $C_2 * \mathbb{Z}$ and $C_3 * \mathbb{Z}$ are residually $X(q)$, where X is a fixed Lie type (not ${}^2B_2$ in the $C_3 * \mathbb{Z}$ case) and $q \to \infty$. Choose $G = G(K_p)$ and σ_q such that $X(q) = (G_{\sigma_q})'$, as above.

Again by Theorem 4.1 there is a subgroup $Y \cong PSL_2(K)$ of G such that $Y \cap X(q)$ contains elements a, b of orders 2, 3 lying in G-classes of maximal dimension, and $\langle a, b^y\rangle = \langle a\rangle * \langle b^y\rangle \cong C_2 * C_3$.

From Lemma 5.4 we know that $\langle a, b^y a b^y\rangle = \langle a\rangle * \langle b^y a b^y\rangle \cong C_2 * \mathbb{Z}$. Fix a nontrivial word $1 \neq w \in C_2 * \mathbb{Z}$. Then $w(a, b^y a b^y) \neq 1$. Therefore the subvariety $V' = \{t \in G\colon w(a,t) = 1\}$ is proper in G. Hence using Lemma 5.6 as before we see that

$$\frac{|\{t \in X(q)\colon\ w(a,t) \neq 1\}|}{|X(q)|} \to 1 \quad \text{as } q \to \infty. \tag{3}$$

Also, by Proposition 2.1(B)(i),

$$\frac{|\{t \in X(q)\colon\ \langle a,t\rangle = X(q)\}|}{|X(q)|} \to 1 \quad \text{as } q \to \infty. \tag{4}$$

By (3) and (4), if q is large enough there exists $t \in X(q)$ such that $\langle a,t\rangle = X(q)$ and $w(a,t) \neq 1$. This completes the proof of Theorem 1.3(i).

Finally, for Theorem 1.3(ii), observe that using Lemma 5.4 it follows that there exists an element $z \in G$ such that $\langle b, z\rangle = \langle b\rangle * \langle z\rangle \cong C_3 * \mathbb{Z}$. Let $1 \neq w \in C_3 * \mathbb{Z}$, and define the subvariety $V'' = \{t \in G\colon w(b,t) = 1\}$. This is then proper in G, and the proof goes through exactly as above. □

To complete the paper we prove the partial result on free products $A * \mathbb{Z}$ with $A \leqslant PSL_2(p)$ referred to in Section 1.

Proposition 5.7. *Fix a prime p, let A be a nontrivial subgroup of $PSL_2(p)$, and let S be a collection of finite simple groups of Lie type of bounded rank in characteristic p not containing $Sz(q)$. Then the free product $A * \mathbb{Z}$ is residually S.*

Proof. As before let X be a fixed Lie type and $X(q) = (G_{\sigma_q})'$, where $G = G(K_p)$. We need to show that $A * \mathbb{Z}$ is residually in any infinite set of groups $X(q)$.

Now $X(q)$ has a subgroup $L \cong PSL_2(p)$ which lies in a subgroup $Y \cong PSL_2(K_p)$ of G. Moreover L is unique up to conjugacy in Y. Since K_p contains $\mathbb{F}_{p^e}((t))$ for any e, Lemma 5.4(ii) implies that there is an injection $\phi : PSL_2(p) * \mathbb{Z} \to Y$ sending $PSL_2(p)$ to L. Let T be a generator of the $\mathbb{Z}$ factor, and let $z = \phi(T) \in Y$.

Now let $1 \neq w \in A * \mathbb{Z}$. For $t \in G$, define $\phi_t : A * \langle T \rangle \to G$ to be the homomorphism acting on A as ϕ and sending T to t. Then $\phi_z(w) \neq 1$, so we see in the usual way using Lemma 5.6 that

$$\frac{|\{t \in X(q) \colon \phi_t(w) \neq 1\}|}{|X(q)|} \to 1 \quad \text{as } q \to \infty.$$

Using Proposition 2.2(ii), we see that, as $q \to \infty$ and $t \in X(q)$ is randomly chosen, the probability that $\langle \phi(A), t \rangle = X(q)$ tends to 1. Therefore for sufficiently large q there exists $t \in X(q)$ such that $\phi_t(w) \neq 1$ and $\mathrm{Im}(\phi_t) = X(q)$. The conclusion follows. □

References

[1] M. Aschbacher, G.M. Seitz, Involutions in Chevalley groups over fields of even order, Nagoya Math. J. 63 (1976) 1–91.

[2] R.W. Carter, Simple Groups of Lie Type, Wiley–Interscience, New York, 1972.

[3] B. Chang, The conjugate classes of Chevalley groups of type (G_2), J. Algebra 9 (1968) 190–211.

[4] J.D. Dixon, L. Pyber, Á. Seress, A. Shalev, Residual properties of free groups and probabilistic methods, J. Reine Angew. Math. 556 (2003) 159–172.

[5] Yu.M. Gorchakov, V.M. Levchuk, On approximation of free groups, Algebra i Logika 9 (1970) 415–421.

[6] R. Guralnick, W.M. Kantor, Probabilistic generation of finite simple groups, J. Algebra 234 (2000) 743–792.

[7] R. Guralnick, M.W. Liebeck, J. Saxl, A. Shalev, Random generation of finite simple groups, J. Algebra 219 (1999) 345–355.

[8] R. Guralnick, F. Lübeck, A. Shalev, Zero-one laws for finite Chevalley groups, in preparation.

[9] E. Hrushovski, The first order theory of the Frobenius automorphisms, in preparation.

[10] R. Katz, W. Magnus, Residual properties of free groups, Comm. Pure Appl. Math. 22 (1969) 1–13.

[11] R. Lawther, Jordan block sizes of unipotent elements in exceptional algebraic groups, Comm. Algebra 23 (1995) 4125–4156.

[12] M.W. Liebeck, G.M. Seitz, On finite subgroups of exceptional algebraic groups, J. Reine Angew. Math. 515 (1999) 25–72.

[13] M.W. Liebeck, A. Shalev, Classical groups, probabilistic methods, and the $(2,3)$-generation problem, Ann. of Math. 144 (1996) 77–125.

[14] M.W. Liebeck, A. Shalev, Simple groups, probabilistic methods, and a conjecture of Kantor and Lubotzky, J. Algebra 184 (1996) 31–57.

[15] M.W. Liebeck, A. Shalev, Random (r,s)-generation of finite classical groups, Bull. London Math. Soc. 34 (2002) 185–188.

[16] A. Lubotzky, On a problem of Magnus, Proc. Amer. Math. Soc. 98 (1986) 583–585.

[17] R.C. Lyndon, P.E. Schupp, Combinatorial Group Theory, Springer, Berlin, 1977.

[18] K. Mizuno, The conjugate classes of Chevalley groups of type E_6, J. Fac. Sci. Univ. Tokyo 24 (1977) 525–563.

[19] K. Mizuno, The conjugate classes of Chevalley groups of type E_7 and E_8, Tokyo J. Math. 3 (1980) 291–461.

[20] G.M. Seitz, Maximal subgroups of exceptional algebraic groups, Mem. Amer. Math. Soc. 90 (1991), No. 441.

[21] J.-P. Serre, Trees, Springer, Berlin, 1980.

[22] K. Shinoda, The conjugacy classes of Chevalley groups of type (F_4) over finite fields of characteristic 2, J. Fac. Sci. Univ. Tokyo 21 (1974) 133–159.

[23] T. Shoji, The conjugacy classes of Chevalley groups of type (F_4) over finite fields of characteristic $p \neq 2$, J. Fac. Sci. Univ. Tokyo 21 (1974) 1–17.

[24] T.A. Springer, R. Steinberg, Conjugacy classes, in: A. Borel et al. (Eds.), Seminar on Algebraic Groups and Related Finite Groups, Lecture Notes in Math., Vol. 131, Springer, 1986, pp. 168–266.

[25] M.C. Tamburini, J.S. Wilson, A residual property of free products, Math. Z. 186 (1984) 525–530.

[26] T.S. Weigel, Residual properties of free groups, J. Algebra 160 (1993) 16–41.

[27] T.S. Weigel, Residual properties of free groups II, Comm. Algebra 20 (1992) 1395–1425.

[28] T.S. Weigel, Residual properties of free groups III, Israel J. Math. 77 (1992) 65–81.

[29] J. Wiegold, Free groups are residually alternating of even degree, Arch. Math. (Basel) 28 (1977) 337–339.

[30] J.S. Wilson, A residual property of free groups, J. Algebra 138 (1991) 36–47.