Available online at www.sciencedirect.com

ScienceDirect

Journal of Algebra 319 (2008) 3075–3086

JOURNAL OF Algebra

www.elsevier.com/locate/jalgebra

Mixing and generation in simple groups ☆

Aner Shalev

Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel

Received 21 May 2007

Available online 1 November 2007

Communicated by Efim Zelmanov

Abstract

Let G be a finite simple group. We show that a random walk on G with respect to the conjugacy class x^G of a random element $x \in G$ has mixing time 2. In particular it follows that $(x^G)^2$ covers almost all of G, which could be regarded as a probabilistic version of a longstanding conjecture of Thompson. We also show that if w is a non-trivial word, then almost every pair of values of w in G generates G.

Keywords: Random walks; Finite simple groups; Word maps; Characters; Probabilistic methods

1. Introduction

Random walks on finite (almost) simple groups G with respect to a conjugacy class C as a generating set have been studied extensively in the past decades. See Diaconis and Shahshahani [DS] for transpositions in symmetric groups, Lulov [Lu] and Vishne [V] for homogeneous classes in symmetric groups, and [H,Gl,LiSh2,LiSh5] for groups of Lie type. A main problem investigated is determining the mixing time $T(C, G)$ of the random walk, namely the time required till we reach an almost uniform distribution on G. In most cases this mixing time is still not known. For background see also [D1,D2].

Here we obtain a somewhat surprising result, showing that this mixing time $T(G, C)$ is usually the smallest possible, namely 2.

☆ Supported by the Israel Science Foundation and by the Bi-National Science Foundation United States–Israel Grant 2004-052.

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

0021-8693/$ – see front matter
doi:10.1016/j.jalgebra.2007.07.031

Theorem 1.1. *Let G be a finite simple group, let $x \in G$ be randomly chosen, and let $C = x^G$ be its conjugacy class. Then the probability that $T(C, G) = 2$ tends to 1 as $|G| \to \infty$.*

This means that the product of two random elements of a "typical" class C is almost uniformly distributed on G.

By Theorem 1.8 of [LiSh5] and the remark following it, if G is a (large) simple group of Lie type and C is a regular semisimple class in G, then $T(G, C) = 2$. This can be used to prove Theorem 1.1 for Lie type groups $X(q)$ where $q \to \infty$. However, this argument does not work for alternating groups and Lie type groups over bounded fields. Our proof of Theorem 1.1 employs a different strategy and is based on a result of independent interest which we obtain for arbitrary finite groups (see Theorem 2.4 below).

Our method also yields results on random walks when the generating set changes over time. We show that for almost all x, y in a finite simple group G, a product of a random element of x^G with a random element of y^G is almost uniformly distributed on G (see Theorem 2.5 below).

The results above immediately imply the following.

Corollary 1.2. *Let G be a finite simple group and fix $\epsilon > 0$. Let $x, y \in G$ be randomly chosen.*

(i) *The probability that $|(x^G)^2|/|G| \geqslant 1 - \epsilon$ tends to 1 as $|G| \to \infty$.*
(ii) *The probability that $|x^G y^G|/|G| \geqslant 1 - \epsilon$ tends to 1 as $|G| \to \infty$.*

We note that for alternating groups $G = A_n$, the probability that $(x^G)^2 = G$ already tends to 1, as was recently shown in [LaSh].

A longstanding conjecture of Thompson states that every finite simple group G has a conjugacy class C such that $C^2 = G$. This is known in various cases but very much open in general. See [EG] for background and latest results. Corollary 1.2 shows that the square of a class of a random element of G covers almost all of G. This provides positive evidence towards Thompson's conjecture, suggesting that C^2 might be equal to G for many classes C.

A key tool in the proof of the above results is the study of the so-called Witten zeta function ζ^G encoding the character degrees of a finite group G. Let $\operatorname{Irr} G$ denote the set of complex irreducible characters of G. For a real number s define

$$\zeta^G(s) = \sum_{\chi \in \operatorname{Irr} G} \chi(1)^{-s}.$$

This function was studied and applied extensively in [LiSh3,LiSh4,LiSh5,GSh]. Here we show that, *if G_i is any family of finite groups satisfying $\zeta^{G_i}(2/3) \to 1$, and $x \in G_i$ is chosen at random, then $T(x^{G_i}, G_i) = 2$ with probability tending to* 1. See Theorem 2.4 below for this, and more general results of this type. This theorem implies Theorem 1.1 for all families of simple groups except $L_2(q), L_3(q), U_3(q)$, which are dealt with directly.

It is intriguing that, while the value of $\zeta^G(2/3)$ is important in our context, other values of the Witten zeta function play a vital role in other contexts. For instance, in [GSh] we show that if G is a finite group, and $\zeta^G(2)$ is close to 1, then the commutator map on G is almost measure preserving, and in particular almost all elements of G are commutators; this is the case where G is a (large) finite simple group.

Note that $\zeta^G(s) = \sum_{n \geqslant 1} r_n n^{-s}$, where r_n is the number of complex irreducible representations of G of dimension n. We can therefore say that groups with very small representation

growth (in the sense that $\sum r_n n^{-2/3}$ is very close to 1) have mixing properties as in Theorem 1.1 and Corollary 1.2 above.

While the results above deal with products of conjugacy classes, a new method of Gowers [G], which is also based on characters, enables one to study products of two arbitrary subsets of finite groups. Here the relevant invariant is not the full zeta function ζ^G, but the minimal degree of a non-trivial character of G. Using a theorem of Gowers (see Theorem 2.7 below) we easily deduce the following.

Proposition 1.3. *Let G be any finite group, and let k be the minimal degree of a non-trivial character of G. Let A, B be any subsets of G. Then*

(i) $\frac{|AB|}{|G|} \geqslant 1 - (k\frac{|A|}{|G|}\frac{|B|}{|G|})^{-1}$.
(ii) *If (G, A, B) range over a family satisfying $|A||B|/(|G|^2/k) \to \infty$, then $|AB|/|G| \to 1$.*
(iii) *In particular, if $|A|/(|G|/k^{1/2}) \to \infty$, then $|A^2|/|G| \to 1$.*

A related recent result of Nikolov and Pyber [NP] shows that, with the above notation, if $|A| > |G|/k^{1/3}$ then $A^3 = G$. Part (iii) above shows that a weaker condition implies that A^2 covers almost all of G.

We note that parts (ii) and (iii) cannot be applied when G is simple and A, B are conjugacy classes; indeed conjugacy classes in simple groups are too small to satisfy the assumption $|A||B|/(|G|^2/k) \to \infty$. In particular Proposition 1.3 does not imply Corollary 1.2. However, it does imply a result on values of words in simple groups.

Let $w \neq 1$ be a word, namely a non-identity element of some free group, and let $w(G)$ denote the values of w in G. In recent years there is a growing interest in the subsets $w(G)$ for finite simple groups G, in relation to Waring type problems and other problems, see [DPSSh,La,Sh, LaSh].

It is shown in [Sh] that $w(G)^3 = G$ if $|G| \geqslant N_w$, namely every element of G is a product of 3 values of w. This was reproved in [NP] using Gowers method and results from [LaSh]. In [LaSh] it is shown that $w(G)^2 = G$ for alternating groups and groups of Lie type of bounded rank, and it remains open whether this is true in general. Proposition 1.3 can be used to show the following.

Corollary 1.4. *Let w be a non-identity word, and let G be a finite simple group. Then $|w(G)^2|/|G| \to 1$ as $|G| \to \infty$. Moreover, if w_1, w_2 are two non-identity words, then $|w_1(G)w_2(G)|/|G| \to 1$ as $|G| \to \infty$, namely almost all elements of G are a product of a value of w_1 and a value of w_2.*

This result also follows from [Sh] and [LaSh] but Proposition 1.3 above provides a shorter proof. Corollary 1.4 has various interesting instances. For example, it shows that, given an integer $m \geqslant 2$, almost all elements of a (large) finite simple group G can be written as a product of two mth powers in G.

Our final result deals with generation by two random elements of $w(G)$. In recent years many results on random generation of finite simple groups have been obtained, see for instance [KL, LiSh1,LiSh23,GLSSh,GK,LiShrs] and the references therein. Many results focus on random generation of simple groups by elements of given orders; here we focus on elements which are values of a given word.

Theorem 1.5. *Let w be a non-identity word, and G a finite simple group. Suppose $x, y \in w(G)$ are chosen at random. Then x, y generate G with probability tending to 1 as $|G| \to \infty$.*

The case of groups of Lie type was recently established by Nikolov and Pyber [NP]. Therefore it remains to prove the theorem for alternating groups, which is what we do here. The proof for alternating groups is harder, since random generation does not follow just from results on the size of the subset $w(G)$, and more delicate arguments should be invoked.

We note that similar arguments show that, if w_1, w_2 are non-identity words, then random elements of $w_1(G)$ and of $w_2(G)$ generate G with probability tending to 1. This leads to various particular cases of interest. For example we have

Corollary 1.6. *Let k, m be positive integers. Then a finite simple group G is almost surely generated by a random kth power and a random mth power in G.*

This paper is organized as follows. In Section 2 we show how the behavior of the function ζ^G for a finite group G can be used to bound the mixing time of the random walks under consideration. We then prove results 1.1–1.3. Section 3 deals with word values and random generation in simple groups. This is where results 1.4–1.5 are proved.

2. Characters and random walks

We start with properties of the Witten zeta function

$$\zeta^G(s) = \sum_{\chi \in \operatorname{Irr} G} \chi(1)^{-s}.$$

In what follows we let L_n (U_n) denote the projective special linear (unitary) group in dimension n.

Lemma 2.1. *Let G be a finite simple group.*

(i) $\zeta^G(2) \to 1$ *as* $|G| \to \infty$.
(ii) *If* $G \neq L_2(q), L_3(q), U_3(q)$, *then* $\zeta^G(2/3) \to 1$ *as* $|G| \to \infty$.

Proof. Part (i) follows from Theorem 1.1 of [LiSh4].

Part (ii) follows from Theorems 1.1 and 1.2 of [LiSh5] (see also the remark following Theorem 1.1 of [LiSh4]). □

Given a finite group G let $P = P_G$ be the uniform distribution on G. If Q is another probability distribution on G we set

$$\|Q - P\| = \sum_{g \in G} \big|Q(g) - P(g)\big|,$$

the L_1-distance between P and Q (which is equivalent to the so-called variation distance, see [D1]).

Lemma 2.2. *Let $s > 0$ be a real number, G a finite group, and let $x \in G$ be randomly chosen. Then the probability that*

$$\big|\chi(x)\big| \leqslant \chi(1)^s \quad \text{for all } \chi \in \operatorname{Irr} G$$

is greater than $2 - \zeta^G(2s)$.

Proof. Let $P = P_G$ as above.

Given $\chi \in \operatorname{Irr} G$ let $E(\chi) = \{x \in G\colon\ |\chi(x)| > \chi(1)^s\}$. Then

$$|G| = \sum_{x\in G} \big|\chi(x)\big|^2 > \big|E(\chi)\big|\chi(1)^{2s}.$$

Hence $P(E(\chi)) < \chi(1)^{-2s}$. Clearly $P(E(1)) = 0$. It follows that

$$P\Bigg(\bigcup_{\chi\in\operatorname{Irr} G} E(\chi)\Bigg) \leqslant \sum_{\chi\neq 1} P\big(E(\chi)\big) < \zeta^G(2s) - 1.$$

Thus the probability of the complementary event is greater than $1 - (\zeta^G(2s) - 1) = 2 - \zeta^G(2s)$. The lemma is proved. □

Given $x \in G$ and a positive integer t define

$$d_t(x) = \sum_{1\neq\chi\in\operatorname{Irr} G} \big|\chi(x)\big|^{2t}/\chi(1)^{2t-2}.$$

Let $C = x^G$, and let P_C^t denote the probability distribution on G after t steps of the random walk with C as a generating set. We are interested in the L_1-distance $\|P_C^t - P\|$ between P_C^t and the uniform distribution P on G. When this distance is smaller than $1/e$ we say that the mixing time is $\leqslant t$. By the upper bound lemma of [DS] we have

$$\big\|P_C^t - P\big\|^2 \leqslant d_t(x).$$

Therefore, to show that $T(x^G, G) = 2$ it suffices to show that $d_2(x)$ is sufficiently small.

Our next result shows that the value of ζ^G at the point $2/3$ has special importance in our context.

Proposition 2.3. *Let t be a positive integer, G a finite group, and let $\alpha = \zeta^G(2 - 4/(t+1)) - 1$. Let $x \in G$ be randomly chosen and set $C = x^G$. Then the probability that $\|P_C^t - P\|^2 \leqslant \alpha$ is at least $1 - \alpha$.*

In particular, if $\alpha = \zeta^G(2/3) - 1$, then the probability that $\|P_C^2 - P\|^2 \leqslant \alpha$ is at least $1 - \alpha$.

Proof. Let $s = (t-1)/(t+1)$. Then $2s = 2 - 4/(t+1)$. Let E be the set of elements $x \in G$ satisfying

$$\big|\chi(x)\big| \leqslant \chi(1)^s \quad \text{for all } \chi \in \operatorname{Irr} G.$$

By Lemma 2.2 we have $P(E) \geqslant 2 - \zeta^G(2s) = 1 - \alpha$. Suppose $x \in E$. Then we have

$$d_t(x) \leqslant \sum_{\chi \neq 1} \big(\chi(1)^s\big)^{2t} / \chi(1)^{2t-2}.$$

Since $2t - 2 - 2ts = 2s$ we obtain

$$\big\| P_C^t - P \big\|^2 \leqslant d_t(x) \leqslant \sum_{\chi \neq 1} \chi(1)^{-2s} = \zeta^G(2s) - 1 = \alpha.$$

The result follows. □

Proposition 2.3 immediately yields the following.

Theorem 2.4. *Fix a positive integer t. Let G range over a family of finite groups satisfying $\zeta^G(2 - 4/(t+1)) \to 1$. Let $x \in G$ be randomly chosen. Then the probability that $T(x^G, G) \leqslant t$ tends to* 1.

In particular if $\zeta^G(2/3) \to 1$ then the probability that $T(x^G, G) = 2$ tends to 1.

This theorem shows that if, for some $\epsilon > 0$, we have $\zeta^G(2 - \epsilon) \to 0$, then the mixing time $T(x^G, G)$ is bounded almost surely.

We can now prove Theorem 1.1. Let G be a finite simple group. Suppose first that $G \neq L_2(q), L_3(q), U_3(q)$. Then by Lemma 2.1(ii) we have

$$\zeta^G(2/3) \to 1 \quad \text{as } |G| \to \infty.$$

Therefore the required conclusion follows from Theorem 2.4

Now suppose $G = L_2(q), L_3(q)$ or $U_3(q)$ (or indeed, any Lie type group of bounded rank). Let E be the set of regular semisimple elements of G. Then as $|G| \to \infty$ we have $q \to \infty$, and $P(E) \geqslant 1 - O(q^{-1})$ by [GL]. Thus $P(E) \to 1$ as $|G| \to \infty$. Now, by Lemma 4.4 of [Sh] there is an absolute constant c such that for $x \in E$ we have

$$\big|\chi(x)\big| \leqslant c \quad \text{for all } x \in E.$$

Thus for such x we have

$$d_2(x) \leqslant \sum_{\chi \neq 1} c^4 \chi(1)^{-2} = c^4\big(\zeta^G(2) - 1\big).$$

Combining this with Lemma 2.1(i) we see that for $x \in E$ we have $d_2(x) \to 0$ as $|G| \to \infty$. This completes the proof of Theorem 1.1.

There is a variant of Theorem 1.1 dealing with two distinct conjugacy classes. Let $x, y \in G$ be randomly chosen, and consider their conjugacy classes x^G, y^G. Let Q be the distribution on G of the random variable obtained by multiplying random elements of x^G and y^G. Then standard Fourier-type arguments show that

$$\|Q - P\|^2 \leqslant d(x, y),$$

where

$$d(x, y) = \sum_{1 \neq \chi \in \operatorname{Irr} G} |\chi(x)|^2 |\chi(y)|^2 / \chi(1)^2.$$

If G is simple then the arguments above show that the probability that $d(x, y) \leqslant \epsilon$ tends to 1 as $|G| \to \infty$. This gives rise to

Theorem 2.5. *Let G be a finite simple group, and let $x, y \in G$ be randomly chosen. Then the product of random elements of x^G and y^G is almost uniformly distributed.*

The above results have immediate applications to the size of $(x^G)^2$ and $x^G y^G$. Indeed, $d_2(x) \to 0$ implies $|(x^G)^2|/|G| \to 1$, and $d(x, y) \to 0$ implies $|x^G y^G|/|G| \to 1$. This gives rise to Corollary 1.2.

Our method yields a related quantitative result for arbitrary finite groups.

Corollary 2.6. *Let G be a finite group, and let $\epsilon = (\zeta^G(2/3) - 1)^{1/2}$. Let $x, y \in G$ be chosen at random. Then*

(i) $P(|(x^G)^2| \geqslant (1 - \epsilon)|G|) \geqslant 1 - \epsilon^2$.
(ii) $P(|x^G y^G| \geqslant (1 - \epsilon)|G|) \geqslant (1 - \epsilon^2)^2$.

Proof. We use the second part of Proposition 2.3 and its notation. Thus $\epsilon = \alpha^{1/2}$ and the probability that $\|P_C^2 - P\| \leqslant \epsilon$ is at least $1 - \epsilon^2$.

Note that if Q is a probability distribution on G satisfying $\|Q - P\| \leqslant \epsilon$, then the support of Q has size at least $(1 - \epsilon)|G|$. Hence $\|P_C^2 - P\| \leqslant \epsilon$ implies $|C^2| \geqslant (1 - \epsilon)|G|$. This completes the proof of part (i).

To prove part (ii), we deduce from Lemma 2.2 that the probability that $|\chi(x)| \leqslant \chi(1)^{1/3}$ and $|\chi(y)| \leqslant \chi(1)^{1/3}$ for all $\chi \in \operatorname{Irr} G$ is at least $(2 - \zeta^G(2/3))^2 = (1 - \epsilon^2)^2$. Now, assuming x, y satisfy the above inequalities we have $d(x, y) \leqslant \zeta^G(2/3) - 1$. Combining this with the discussion and notation preceding Theorem 2.5 we see that $\|Q - P\| \leqslant \epsilon$ with probability at least $(1 - \epsilon^2)^2$. Finally, $\|Q - P\| \leqslant \epsilon$ implies $|x^G y^G| \geqslant (1 - \epsilon)|G|$.

This completes the proof. □

We now prove Proposition 1.3. Our main tool is the following result of Gowers (see p. 22 of [G]).

Theorem 2.7. *Let G be a finite group and k the minimal degree of a non-trivial character of G. If A, B, C are subsets of G such that $|A\|B\|C| > |G|^3/k$, then there are $a \in A$, $b \in B$, $c \in C$ such that $ab = c$.*

Now, given $A, B \subseteq G$, define $C = G \setminus AB$. Then there are no $a \in A, b \in B, c \in C$ with $ab = c$. Hence, by Gowers' Theorem, $|A\|B\|C| \leqslant |G|^3/k$, and so

$$|G| - |AB| = |C| \leqslant |G|^3 \big/ \big(k|A\|B|\big).$$

This yields $|AB|/|G| \geqslant 1 - |G|^2/(k|A\|B|)$, proving part (i) of Proposition 1.3. Parts (ii) and (iii) follow from part (i).

3. Word values and generation

We first provide a short proof of Corollary 1.4 using Gowers' method. Let $w_1, w_2 \neq 1$ be words. If $G = A_n$ and n is large enough then it is shown in [LaSh] that

$$w_1(G)w_2(G) = G.$$

So it remains to deal with simple groups of Lie type. Suppose G is such a group of rank r over the field with q elements. Let k be the minimal degree of a non-trivial character of G.

We now apply Proposition 1.3 with $A = w_1(G)$ and $B = w_2(G)$. To show that $|w_1(G)w_2(G)|/|G| \to 1$ it suffices to show that, for $i = 1, 2$,

$$\big|w_i(G)\big|\big/\big(|G|/k^{1/2}\big) \to \infty \quad \text{as } |G| \to \infty.$$

Now, by [LS] we have

$$k \geqslant cq^r,$$

where $c > 0$ is some absolute constant. It therefore suffices to show that, for any word $w \neq 1$,

$$\big|w(G)\big|\big/\big(|G|/q^{r/2}\big) \to \infty \quad \text{as } |G| \to \infty.$$

For r bounded, or G symplectic or orthogonal, this follows from stronger lower bounds on $|w(G)|$ given in [LaSh]. For r unbounded and G special linear or unitary, this follows from results 1.7 and 1.8 in the paper [NP] of Nikolov and Pyber.

This completes the proof of Corollary 1.4.

For the proof of Theorem 1.5 we need some preparations. Let $G = A_n$, an alternating group. We let $\mathcal{M}$ denote a set of representatives of conjugacy classes of maximal subgroups of G. For an integer $1 \leqslant k \leqslant n/2$ let

$$M_k = (S_k \times S_{n-k}) \cap A_n,$$

the setwise stabilizer of $\{1, \ldots, k\}$ in A_n.

Lemma 3.1. *Fix a positive integer $k \leqslant n/2$ and let $\mathcal{T} \subset \mathcal{M}$ be all subgroups in $\mathcal{M}$ which are not conjugate to $M_1, \ldots, M_{k-1}$. Then*

$$\sum_{M \in \mathcal{T}} |G : M|^{-1} = O\left(\binom{n}{k}^{-1}\right),$$

where the implied constant depends on k.

Proof. It follows from [LMSh] that A_n has at most $n^{o(1)}$ conjugacy classes of primitive maximal subgroup M. It is well known that each such subgroup M has size at most $n^{cn^{1/2}}$ (see [C]). Thus the contribution of the primitive subgroups to the sum on the left-hand side is at most $n^{o(1)}(n!/2n^{cn^{1/2}})^{-1}$ which is marginal.

The number of classes of transitive imprimitive maximal subgroups M is $d(n) - 2$, where $d(n) = n^{o(1)}$ is the number of divisors of n. Each such subgroup has index at least $2^{n/2}$ and so again their contribution is marginal.

It therefore remains to deal with intransitive maximal subgroups M_l, where $k \leqslant l \leqslant n/2$. This easily yields the result. □

Lemma 3.2. *For any constant $c_1 > 1$ there is a constant $c_2 > 1$ (depending on c_1) such that, as $x \in A_n$ is randomly chosen, we have*

$$P\left(\left|C_{A_n}(x)\right| < c_2^{\sqrt{n}}\right) \geqslant 1 - c_1^{-\sqrt{n}}.$$

Proof. For a finite group G let $k(G)$ denote the number of its conjugacy classes. Then for a real number $\alpha > 0$ we have

$$P\left(\left|C_G(x)\right| < \alpha k(G)\right) \geqslant 1 - \alpha^{-1}.$$

See Lemma 5.3 of [Sh]. Letting $\alpha = c_1^{\sqrt{n}}$ and using the well-known bound $k(A_n) \leqslant a^{\sqrt{n}}$ we obtain the conclusion with $c_2 = c_1 a$. □

A main tool in our proof of Theorem 1.5 is a theorem from [LaSh], quoted below.

Theorem 3.3. *Let w be a non-identity word, and let $\epsilon > 0$. Then for all sufficiently large n we have*

$$\left|w(A_n)\right| \geqslant n^{-4-\epsilon}|A_n|.$$

We now prove Theorem 1.5. As mentioned in the introduction, it suffices to deal with alternating groups $G = A_n$. Let $W = w(A_n)$, and let $x, y \in W$ be randomly chosen. Fix $0 < \epsilon < 1/2$ and suppose n is large enough, so that

$$|W|/|A_n| \geqslant n^{-4-\epsilon}.$$

Step 1. We first show that the probability that $x, y \in W$ belong to some maximal subgroup M of G which is not conjugate to $M_1, \ldots, M_8$ (stabilizers of k-sets, $k \leqslant 8$) tends to 0 as $n \to \infty$.

Indeed, let T denote the union $\bigcup_M M \times M$ over all maximal subgroups of G not conjugate to some M_k, $k \leqslant 8$. Let $\mathcal{T}$ denote a set of representatives of conjugacy classes of these subgroups.

Then

$$|T|/|G|^2 \leqslant |G|^{-2} \sum_{M \in \mathcal{T}} |G : M||M|^2 = \sum_{M \in \mathcal{T}} |G : M|^{-1} \leqslant c\binom{n}{9}^{-1},$$

by Lemma 3.1.

Since $|W \times W| \geqslant |G|^2 n^{-8-2\epsilon}$ it follows that

$$|T|/|W \times W| \leqslant n^{8+2\epsilon} \Big/ \left(c\binom{n}{9}\right) \to 0 \quad \text{as } n \to \infty.$$

This concludes Step 1 of the proof.

Step 2. We show that the probability that $x, y \in W$ belong to some maximal subgroup M of G conjugate to some M_k with $1 \leqslant k \leqslant 8$ tends to 0.

Fix a constant $c_1 > 1$, and let c_2 be as in Lemma 3.2. Let

$$W_1 = \left\{ g \in W \colon \left| C_{A_n}(g) \right| < c_2^{\sqrt{n}} \right\}.$$

Then Lemma 3.2 yields

$$|W \setminus W_1| \leqslant c_1^{-\sqrt{n}} |A_n|.$$

Since $|W| \geqslant n^{-4-\epsilon}|A_n|$ it follows that

$$|W_1|/|W| \geqslant 1 - n^{4+\epsilon} c_1^{-\sqrt{n}} \to 1.$$

Hence in proving the claim in Step 2 we may assume $x, y \in W_1$.

Let f be the number of fixed points of x (on $\{1, \ldots, n\}$) and e the number of cycles in x of length at least 2. Then

$$f! \cdot 2^e/2 \leqslant \left| C_{A_n}(x) \right| < c_2^{\sqrt{n}}.$$

This implies

$$f \leqslant c_3 \sqrt{n}/\log n \quad \text{and} \quad e \leqslant c_4 \sqrt{n}.$$

Given k with $1 \leqslant k \leqslant 8$ and $g \in G$ let $f_k(g)$ denote the number of fixed points of g in its action on k-subsets of $\{1, \ldots, n\}$ (namely the number of invariant subsets of g of size k).

To bound the number $f_k(x)$ note that each k-subset invariant under x is a union of x-orbits. If i is the number of orbits of size 1 in this union then $i \leqslant k$ and there are at most $(k-i)/2$ orbits of size at least 2. Hence

$$f_k(x) \leqslant \sum_{0 \leqslant i \leqslant k} \binom{f}{i} \left\{ \sum_{0 \leqslant j \leqslant (k-i)/2} \binom{e}{j} \right\} \leqslant \sum_{0 \leqslant i \leqslant k} f^i (e+1)^{(k-i)/2}.$$

Set $e_1 = (e+1)^{1/2}$. Then we have

$$f_k(x) \leqslant \sum_{0 \leqslant i \leqslant k} f^i e_1^{k-i} \leqslant (f + e_1)^k.$$

Our bounds on f, e imply $f + e_1 \leqslant c_5 n^{1/2}/\log n$. Hence

$$f_k(x) \leqslant (c_5/\log n)^k n^{k/2}.$$

In a similar manner we have

$$f_k(y) \leqslant (c_5/\log n)^k n^{k/2}.$$

Let $C = x^G$, $D = y^G$. Then $C, D \subseteq W_1$ since W_1 is a normal subset of G. The probability that a random element of C lies in M_k is $|C \cap M_k|/|C| = f_k(x)/\binom{n}{k}$.

Let $Q(C, D)$ be the probability that random elements of C and D both lie in some conjugate of M_k for some $1 \leqslant k \leqslant 8$. Then

$$Q(C, D) \leqslant \sum_{k=1}^{8} |G : M_k| \frac{f_k(x)}{\binom{n}{k}} \cdot \frac{f_k(y)}{\binom{n}{k}} = \sum_{k=1}^{8} \frac{f_k(x) f_k(y)}{\binom{n}{k}}.$$

Our bounds on $f_k(x)$, $f_k(y)$ now yield

$$Q(C, D) \leqslant \sum_{k=1}^{8} \frac{(c_5/\log n)^{2k} n^k}{\binom{n}{k}} \leqslant c_6 (\log n)^{-2}.$$

Since this holds for any choice of classes $C, D \subseteq W_1$ this concludes the proof of Step 2.

Combining Steps 1 and 2 we see that the probability that randomly chosen elements $x, y \in W$ both lie in some maximal subgroup M of G tends to 0. Thus x, y generate G with probability tending to 1.

Theorem 1.5 is proved.

We note that the same argument shows that if W is any normal subset of A_n satisfying $|W| \geqslant n^{-c \log n} |A_n|$ (with c a suitable constant) then A_n is almost surely generated by two random elements of W. A more refined argument can be used to deduce the same conclusion under weaker assumptions on $|W|$ (e.g. $|W| \geqslant c^{-\sqrt{n}} |A_n|$).

References

[C] P.J. Cameron, Finite permutation groups and finite simple groups, Bull. London Math. Soc. 13 (1981) 1–22.

[D1] P. Diaconis, Group Representations in Probability and Statistics, IMS Lecture Notes Monogr. Ser., vol. 11, 1988.

[D2] P. Diaconis, Random walks on groups: Characters and geometry, in: Groups St Andrews 2001 in Oxford, vol. I, in: London Math. Soc. Lecture Note Ser., vol. 304, Cambridge Univ. Press, Cambridge, 2003, pp. 120–142.

[DS] P. Diaconis, M. Shahshahani, Generating a random permutation with random transpositions, Z. Wahrsch. Verw. Geb. 57 (1981) 159–179.

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

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

[GSh] S. Garion, A. Shalev, Commutator maps, measure preservation, and T-systems, Trans. Amer. Math. Soc., in press.

[Gl] D. Gluck, Characters and random walks on finite classical groups, Adv. Math. 129 (1997) 46–72.

[G] W.T. Gowers, Quasirandom groups, preprint, 2007.

[GK] R.M. Guralnick, W.M. Kantor, The probability of generating a simple group, J. Algebra 234 (2000) 743–792.

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

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

[H] M. Hildebrand, Generating random elements in $SL_n(F_q)$ by random transvections, J. Algebraic Combin. 1 (1992) 133–150.

[KL] W.M. Kantor, A. Lubotzky, The probability of generating a finite classical group, Geom. Dedicata 36 (1990) 67–87.
[LS] V. Landazuri, G.M. Seitz, On the minimal degrees of projective representations of the finite Chevalley groups, J. Algebra 32 (1974) 418–443.
[La] M. Larsen, Word maps have large image, Israel J. Math. 139 (2004) 149–156.
[LaSh] M. Larsen, A. Shalev, Word maps and Waring type problems, preprint, 2007.
[LMSh] M.W. Liebeck, B.M.S. Martin, A. Shalev, On conjugacy classes of maximal subgroups of finite simple groups, and a related zeta function, Duke Math. J. 128 (2005) 541–557.
[LiSh1] M.W. Liebeck, A. Shalev, The probability of generating a finite simple group, Geom. Dedicata 56 (1995) 103–113.
[LiSh23] M.W. Liebeck, A. Shalev, Classical groups, probabilistic methods and the $(2, 3)$-generation problem, Ann. of Math. 144 (1996) 77–125.
[LiSh2] M.W. Liebeck, A. Shalev, Diameters of finite simple groups: Sharp bounds and applications, Ann. of Math. 154 (2001) 383–406.
[LiShrs] M.W. Liebeck, A. Shalev, Random (r, s)-generation of finite classical groups, Bull. London Math. Soc. 34 (2002) 185–188.
[LiSh3] M.W. Liebeck, A. Shalev, Fuchsian groups, coverings of Riemann surfaces, subgroup growth, random quotients and random walks, J. Algebra 276 (2004) 552–601.
[LiSh4] M.W. Liebeck, A. Shalev, Fuchsian groups, finite simple groups, and representation varieties, Invent. Math. 159 (2005) 317–367.
[LiSh5] M.W. Liebeck, A. Shalev, Character degrees and random walks in finite groups of Lie type, Proc. London Math. Soc. 90 (2005) 61–86.
[Lu] N. Lulov, Random walks on symmetric groups generated by conjugacy classes, PhD thesis, Harvard University, 1996.
[NP] N. Nikolov, L. Pyber, Product decompositions of quasirandom groups and a Jordan type theorem, preprint, 2007.
[Sh] A. Shalev, Word maps, conjugacy classes, and a non-commutative Waring-type theorem, Ann. of Math., in press.
[V] U. Vishne, Mixing and covering in symmetric groups, J. Algebra 205 (1998) 119–140.