MOST WORDS ARE GEOMETRICALLY ALMOST UNIFORM

MICHAEL LARSEN

Abstract. If w is a word in $d > 1$ letters and G is a finite group, evaluation of w on a uniformly randomly chosen d-tuple in G gives a random variable with values in G, which may or may not be uniform. It is known [LST] that if G ranges over finite simple groups of given root system and characteristic, a positive proportion of words w give a distribution which approaches uniformity in the limit as $|G| \to \infty$. In this paper, we show that the proportion is in fact 1.

1. Introduction

A *word* for the purposes of this paper is an element of the free group F_d. For any finite group G, the word w defines a word map $w_G \colon G^d \to G$ by substitution; we denote it w when G is understood. If U_G defines the uniform measure on G, we can measure the failure of random values of w to be uniform by comparing the pushforward $w_* U_{G^d}$ to the uniform distribution U_G. We say w is *almost uniform* for an infinite family of finite groups G if

$$\lim_{|G| \to \infty} \|w_* U_{G^d} - U_G\| = 0,$$

where $\|\cdot\|$ denotes the L^1 norm, and G ranges over the groups of the family. We are particularly interested in the family of finite simple groups.

When w is of the form w_0^k for some $k \geq 2$, then w is said to be a *power word*. It is easy to see that power words are not almost uniform for finite simple groups; for instance, in large symmetric groups, most elements are not kth powers at all [P]. There has been speculation as to whether all non-power words are almost uniform for finite simple groups (see, e.g., [Sh, Problem 4.7] and [L, Question 3.1]). Since power words are exponentially thin [LM], one could ask an easier question: is the set of words which are not almost uniform for finite simple groups thin? Or, easier still, does it have density 0? Some words are known to be almost uniform for finite simple groups: primitive words, which are exactly uniform for all groups; the commutator word $x_1 x_2 x_1^{-1} x_2^{-1}$ by [GS], words of the form $x_1^m x_2^n$ by [LS2], and, recently, all words of *Waring type*, i.e., words which can be written as a product of two non-trivial words involving disjoint variables

1991 *Mathematics Subject Classification.* Primary 20P05; Secondary 11G25, 14G15, 20G40.

The author was partially supported by NSF grant DMS-1702152.

[LST, Theorem 1]. The fundamental group of any orientable genus g surface is therefore covered for all $g \geq 1$, and, more generally, various words in which some variables appear exactly twice can also be treated by combining the idea of Parzanchevski-Schul [PS] with the method of Liebeck-Shalev [LiS] All of these words, of course, are in some sense rare and atypical.

From the point of view of algebraic geometry, the easiest families of finite simple groups to consider are those of the form $\underline{G}(\mathbb{F}_{q^n})/Z(\underline{G}(\mathbb{F}_{q^n}))$, where $\underline{G}$ is a simple, simply connected algebraic group over $\mathbb{F}_q$, and n ranges over the positive integers. We say that w is *geometrically almost uniform* for $\underline{G}$ if it is so for this family of groups. In [LST, Theorem 2], it is proved that this property is equivalent to an algebro-geometric condition on w, namely that the morphism of varieties $w_{\underline{G}}\colon \underline{G}^d \to \underline{G}$ (which by a theorem of Borel [B] is dominant) has geometrically irreducible generic fiber. Using this criterion, it is proved in [LST, Theorem 3] that for each d, there exists a set of words of density greater than 1/3 which are almost uniform for $\underline{G}$ for all $\underline{G}/\mathbb{F}_q$. (Note that this does not imply that these words are almost uniform for the family of all finite simple groups of Lie type.)

The main result of this paper is that for each $\underline{G}$ the set of words which are geometrically almost uniform for $\underline{G}$ has density 1. More explicitly:

Theorem 1.1. *Let $\mathbb{F}_q$ and $\underline{G}$ be fixed. Let $(i_1, e_1), (i_2, e_2), \ldots$ be chosen independently and uniformly from $\{1, \ldots, d\} \times \{\pm 1\}$. Let $w = x_{i_1}^{e_1} \cdots x_{i_l}^{e_l}$ be a random word of length l defined in this way. Then the probability that w is geometrically almost uniform for $\underline{G}$ goes to 1 as $l \to \infty$.*

The idea of the proof is as follows. In [LST, Corollary 2.3], it is proved that if the image $\overline{w}$ of w under the abelianization map $F_d \to \mathbb{Z}^d$ is primitive, i.e., if $\gamma(\overline{w}) = 1$, where γ denotes the g.c.d. of its coordinates, then w is almost uniform for every $\underline{G}$, the idea being that $w_{\underline{G}(\mathbb{F}_{q^n})}$ is then surjective for all n, and this implies that $w_{\underline{G}}$ does not factor through a non-birational generically finite morphism $\underline{X}_0 \to \underline{G}$.

Now, the image of a random walk on F_d under the abelianization map is a random walk on $\mathbb{Z}^d$. If $\mathsf{X}_{d,l}$ is the endpoint of a random walk of length l on $\mathbb{Z}^d$, then

$$\limsup_{l\to\infty} \mathbf{P}[\gamma(\mathsf{X}_{d,l}) = 1] < 1$$

for all d, so this is not good enough to get a result which covers almost all words. A new idea is needed.

By a probabilistic analysis, we prove that for each d,

$$\lim_{M\to\infty} \liminf_{l\to\infty} \mathbf{P}[1 \leq \gamma(\mathsf{X}_{d,l}) \leq M] = 1.$$

Thus, it suffices to prove that for each $d \geq 2$ and $k > 0$, in the limit as l goes to infinity, the fraction of w of length l with $\gamma(\overline{w}) = k$ for which w is almost uniform in rank $\leq r$ goes to 1. For any such w and any group G, the image of w_G contains all kth powers in G. For $k > 1$, this no longer implies geometric irreducibility of the generic fiber of $w_{\underline{G}}$, but it puts very strong constraints on which quasi-finite morphisms $\underline{X}_0 \to \underline{G}$ it can factor through.

To see how to exploit such constraints, consider the following toy problem. Suppose c polynomial map $f\colon \mathbb{A}^1 \to \mathbb{A}^1$ is defined over $\mathbb{F}_q$; for all n, $f(\mathbb{F}_{q^n})$ contains all squares in $\mathbb{F}_{q^n}$; and for some n_0, $f(\mathbb{F}_{q^{n_0}})$ contains a non-square. We claim this implies f is purely inseparable.

Indeed, consider the curve $\underline{C}\colon y^2 = f(x)$. For $\underline{C}$ to fail to be geometrically irreducible would mean that $f(x) = g(x)^2$ for some $g(x) \in \overline{\mathbb{F}}_q[x]$. Either $g(x) \in \mathbb{F}_q[x]$ or $f(x) = ah(x)^2$ for some non-square $a \in \mathbb{F}_q$ and some $h(x) \in \mathbb{F}_q[x]$. In the first case, $f(\mathbb{F}_{q^{n_0}})$ contains only squares in $\mathbb{F}_{q^{n_0}}$, contrary to assumption. In the second case, for all $n \geq 1$, $f(\mathbb{F}_{q^n})$ contains no non-zero square in $\mathbb{F}_{q^n}$.

Thus, the conditions on the image of f imply that $\underline{C}$ is geometrically irreducible, so it has $(1+o(1))q^n$ points over $\mathbb{F}_{q^n}$ by the Lang-Weil estimate. Consider the y-map, that is, the morphism of degree $\deg f$ from $\underline{C}$ to the affine line given by the function y. By the Chebotarev density theorem for finite extensions of $\mathbb{F}_q(t)$, in the limit as $n \to \infty$, a fixed positive proportion of points in $\mathbb{A}^1(\mathbb{F}_{q^n})$ have preimage in $\underline{C}(\mathbb{F}_{q^n})$ consisting of $\deg f$ points. Since the y-map is surjective on $\mathbb{F}_{q^n}$-points, this implies that f is purely inseparable.

To apply this idea in the word map setting, one needs to find elements in $w(\underline{G}(\mathbb{F}_{q^n})^d)$ which play the role of non-square elements in $f(\mathbb{F}_{q^n})$. We do not need to find them for all w, just for almost all in an asymptotic sense. An approach to achieving this is to fix a d-tuple $\mathbf{g} \in \underline{G}(\mathbb{F}_{q^n})^d$ and estimate the probability that $w(\mathbf{g})$ is a "non-square" element. For large enough n, one can view $w(\mathbf{g})$ as uniformly distributed in $\underline{G}(\mathbb{F}_{q^n})$. In order to get the probability of success to approach 1, it is necessary to use not a single $\mathbf{g}$ but a sufficiently large number of independent choices $\mathbf{g}_1, \ldots, \mathbf{g}_N$. The existence of N elements of $\underline{G}(\mathbb{F}_{q^n})^d$ which are independent in this sense (in the limit $n \to \infty$) depends on $\underline{G}(\mathbb{F}_{q^n})^N$ being d-generated. There is a substantial literature, going back to work of Philip Hall [H], concerning the size of minimal generating sets of G^N, where G is a finite simple group. We use a recent result of Maróti and Tamburini [MT].

I would like to thank Aner Shalev for his useful comments on an earlier version of this paper. I also want to express my gratitude to the referee for pointing out a number of inaccuracies in an earlier draft of this paper and suggesting several improvements in the exposition.

2. Varieties over Finite Fields

Throughout this section, a *variety* will always mean a geometrically integral affine scheme of finite type over a finite field. Let $A \subset B$ be an inclusion of finitely generated $\mathbb{F}_q$-algebras such that $\underline{X} :=$ Spec A and $\underline{Y} :=$ Spec B are normal varieties. Let $\phi\colon \underline{Y} \to \underline{X} =$ Spec A correspond to the inclusion $A \subset B$. Let K and L denote the fraction fields of A and B respectively. Let

K_0 denote the separable closure of K in L, which is a finite extension of K since L is finitely generated. Let A_0 denote the integral closure of A in K_0, $\underline{X}_0$ the spectrum of A_0, and $\psi\colon \underline{X}_0 \to \underline{X}$ the morphism corresponding to the inclusion $A \subset A_0$. As $B \supset A$ is integrally closed in $L \supset K_0$ it follows that B contains A_0, so ϕ factors through ψ.

Proposition 2.1. *For all positive integers n,*

$$\phi(\underline{Y}(\mathbb{F}_{q^n})) \subset \psi(\underline{X}_0(\mathbb{F}_{q^n})), \tag{2.1}$$

and

$$|\psi(\underline{X}_0(\mathbb{F}_{q^n}))| - |\phi(\underline{Y}(\mathbb{F}_{q^n}))| = o(q^{n \dim \underline{X}}). \tag{2.2}$$

Moreover ψ is an isomorphism if and only if ϕ has geometrically irreducible generic fiber; if not, there exists $\epsilon > 0$ and a positive integer m such that

$$|\psi(\underline{X}_0(\mathbb{F}_{q^n}))| < (1-\epsilon) q^{n \dim \underline{X}} \tag{2.3}$$

if m divides n.

Proof. As $A \subset A_0 \subset B$, the morphism ϕ factors through ψ, implying (2.1).

By [EGA IV$_2$, Proposition 4.5.9], $K = K_0$ if and only if the generic fiber of ϕ is geometrically irreducible. By the same proposition, the generic fiber of $\underline{Y} \to \underline{X}_0$ is always geometrically irreducible. By [EGA IV$_3$, Théorème 9.7.7], there is a dense open subset of $\underline{X}_0$ over which the fibers of $\underline{Y} \to \underline{X}_0$ are all geometrically irreducible. Let $\underline{C}$ denote the complement of this subset, endowed with its structure of reduced closed subscheme of $\underline{X}_0$.

It is well known that the Lang-Weil estimate is uniform in families. There does not seem to be a canonical reference for this fact, but a proof is sketched, for instance in [LS1, Proposition 3.4] and in [T, Theorem 5]. From this, it follows that if n is sufficiently large, for every point of $\underline{X}_0(\mathbb{F}_{q^n})$ over which the morphism $\underline{Y} \to \underline{X}_0$ has geometrically irreducible fiber, there exists an $\mathbb{F}_{q^n}$-point in this fiber. In particular, every point in $\underline{X}_0(\mathbb{F}_{q^n}) \setminus \underline{C}(\mathbb{F}_{q^n})$ lies in the image of $\underline{Y}(\mathbb{F}_{q^n}) \to \underline{X}_0(\mathbb{F}_{q^n})$. By the easy part of the Lang-Weil bound,

$$|\underline{C}(\mathbb{F}_{q^n})| = O(q^{n \dim \underline{C}}) \leq O(q^{n(\dim \underline{X}_0 - 1)}).$$

Thus, the complement of the image of $\underline{Y}(\mathbb{F}_{q^n}) \to \underline{X}_0(\mathbb{F}_{q^n})$ has cardinality $o(q^{n \dim \underline{X}})$, which implies (2.2).

If ϕ is not geometrically irreducible, then $[K_0 : K] > 1$. Let K_1 denote the Galois closure of K_0/K in a fixed separable closure $\overline{K}$. We choose m so that $\mathbb{F}_{q^m}$ contains the algebraic closure of $\mathbb{F}_q$ in K_1. If we are content to limit consideration to $\mathbb{F}_{q^n}$-points of $\underline{X}$ and $\underline{X}_0$, where m divides n, we may replace $\underline{X}$ and $\underline{X}_0$ by the varieties $\underline{X}_{\mathbb{F}_{q^m}}$ and $(\underline{X}_0)_{\mathbb{F}_{q^m}}$ respectively, obtained by base change. This has the effect of replacing K, K_0, and K_1 by $K\mathbb{F}_{q^m}$, $K_0\mathbb{F}_{q^m}$, and $K_1\mathbb{F}_{q^m} = K_1$ respectively. Replacing q by q^m, we may now assume that $\mathbb{F}_q$ is algebraically closed in K_1.

Now, $\mathrm{Gal}(K_1/K)$ acts faithfully on A_1 as $\mathbb{F}_q$-algebra. As A is integrally closed in K and A_1 is the integral closure of A in K_1, it follows that

$$A \subset A_1^{\mathrm{Gal}(K_1/K)} \subset A_1 \cap K = A,$$

so $A = A_1^{\mathrm{Gal}(K_1/K)}$; likewise, $A_0 = A_1^{\mathrm{Gal}(K_1/K_0)}$. Geometrically, this means that $\underline{X}$ and $\underline{X}_0$ are the quotients of $\underline{X}_1 := \mathrm{Spec}\, A_1$ by $\mathrm{Gal}(K_1/K)$ and $\mathrm{Gal}(K_1/K_0)$ respectively. We denote these quotient maps π and π_0 respectively. Thus we have the diagram

$$\begin{tikzcd}
\underline{X}_1 \arrow[d, "\pi_0"'] \arrow[dd, bend left, "\pi"] \\
\underline{X}_0 \arrow[d, "\psi"'] \\
\underline{X}
\end{tikzcd}$$

As the action of $\mathrm{Gal}(K_1/K)$ on $\underline{X}_1$ is faithful and $\underline{X}_1$ is irreducible, there is a dense affine open subvariety of $\underline{X}_1$ on which $\mathrm{Gal}(K_1/K)$ acts freely. Replacing $\underline{X}_1$ by this subvariety and $\underline{X}$ and $\underline{X}_0$ by quotients of this subvariety by $\mathrm{Gal}(K_1/K)$ and $\mathrm{Gal}(K_1/K_0)$ respectively affects $o(q^{n \dim \underline{X}})$ of the $\mathbb{F}_{q^n}$-points of $\underline{X}$, $\underline{X}_0$, and $\underline{X}_1$, so without loss of generality, we may assume that $\mathrm{Gal}(K_1/K)$ acts freely on $\underline{X}_1$. Now

$$\psi(\underline{X}_0(\mathbb{F}_{q^n})) = \psi(\underline{X}_0(\mathbb{F}_{q^n}) \setminus \pi_0(\underline{X}_1(\mathbb{F}_{q^n}))) \cup \pi(\underline{X}_1(\mathbb{F}_{q^n})). \tag{2.4}$$

By Lang-Weil, $|\underline{X}_1(\mathbb{F}_{q^n})| = (1 + o(1))q^{n \dim \underline{X}}$, so

$$\begin{aligned}
|\pi_0(\underline{X}_1(\mathbb{F}_{q^n})))| &= ([K_1 : K_0]^{-1} + o(1))q^{n \dim \underline{X}}, \\
|\pi(\underline{X}_1(\mathbb{F}_{q^n}))| &= ([K_1 : K]^{-1} + o(1))q^{n \dim \underline{X}}.
\end{aligned}$$

By (2.4),

$$|\psi(\underline{X}_0(\mathbb{F}_{q^n}))| \leq (1 - [K_1 : K_0]^{-1} + [K_1 : K]^{-1} + o(1))q^{n \dim \underline{X}},$$

which implies (2.3). □

Lemma 2.2. *Let G be a finite group acting transitively on a set S with more than one element and H a normal subgroup of G such that every element of H has at least one fixed point in S. Then for all $s \in S$, $H\, \mathrm{Stab}_G(s)$ is a proper subgroup of G.*

Proof. By a classical theorem of Jordan, every non-trivial transitive permutation group contains a derangement, so H must act intransitively. Thus, the orbit of $H\, \mathrm{Stab}_G(s)$ containing s is a proper subset of S, which implies the lemma. □

Lemma 2.3. *Let K be a field, $\overline{K}$ a separable closure of K, and K_1 and K_2 finite extensions of K in $\overline{K}$. Suppose K_1 is Galois over K and $K_2 \neq K$. If $K_1 \cap K_2 = K$, then there exists an element of $\mathrm{Gal}(\overline{K}/K_1)$ which does not stabilize any K-embedding of K_2 in $\overline{K}$.*

Proof. Let K_3 be the Galois closure of K_2 in $\overline{K}$ and define $G := \mathrm{Gal}(K_1K_3/K)$. Thus G acts transitively on the set S of K-embeddings of K_2 in $\overline{K}$. Let $H = \mathrm{Gal}(K_1K_3/K_1)$, which is normal in G since K_1/K is Galois. If every element of $\mathrm{Gal}(\overline{K}/K_1)$ fixes at least one element of S, then by Lemma 2.2, $H\,\mathrm{Stab}_G(s)$ is a proper subgroup of G, where s denotes the identity embedding of K_2 in $\overline{K}$. If L is the fixed field of K_1K_3 under $H\,\mathrm{Stab}_G(s)$, then L is a non-trivial extension of K contained in both $(K_1K_3)^H = K_1$ and $(K_1K_3)^{\mathrm{Stab}_G(s)} = K_2$.

□

Proposition 2.4. *Let $\underline{X}$ be a variety over $\mathbb{F}_q$ with coordinate ring A with function field K. Let $K \subset K_0, K_2 \subset \overline{K}$, and let K_1 (resp. K_3) denote the Galois closure of K_0 (resp. K_2) in $\overline{K}$. Let A_i for $0 \le i \le 3$ denote the integral closure of A in K_i, and let $\underline{X}_i := \mathrm{Spec}\ A_i$. If K_1 and K_2 satisfy the hypotheses of Lemma 2.3, then there exists $\epsilon > 0$ so that for all sufficiently large integers n, there are at least $\epsilon q^{n \dim \underline{X}}$ elements of $\underline{X}(\mathbb{F}_{q^n})$ which lie in the image of $\underline{X}_i(\mathbb{F}_{q^n}) \to \underline{X}(\mathbb{F}_{q^n})$ for $i = 0$ but not for $i = 2$.*

Proof. Let $K_{13} = K_1K_3$, A_{13} denote the integral closure of A in K_{13}, and $\underline{X}_{13}$ denote Spec A_{13}. Let $G := \mathrm{Gal}(K_{13}/K)$. The action of G on $\underline{X}_{13}$ is faithful, and $\underline{X}_{13}$ is irreducible, so there exists a dense open affine subvariety $\underline{U}_{13} \subset \underline{X}_{13}$ on which G acts freely. Replacing $\underline{X}_{13}$, together with its quotients by subgroups of G, by $\underline{U}_{13}$ and its corresponding quotients affects only $o(q^{n \dim \underline{X}})$ $\mathbb{F}_{q^n}$-points of these quotients, and therefore does not affect the statement of the proposition. We may therefore assume that we are in the setting of [Se, Theorem 6] and can apply the Chebotarev density theorem for varieties.

By Lemma 2.3, there exists $g \in G$ such that g acts trivially on K_1 but acts without fixed points on the set of K-embeddings $K_2 \to \overline{K}$ or, equivalently, on the geometric points lying over any given geometric point of $\underline{X}$ for the covering map $\underline{X}_2 \to \underline{X}$. This implies that if $x \in \underline{X}(\mathbb{F}_{q^n})$ and g belongs to the Frobenius conjugacy class of x, then there is no q^n-Frobenius stable point lying over x on $\underline{X}_2 \to \underline{X}$, i.e., x does not lie in the image of $\underline{X}_2(\mathbb{F}_{q^n}) \to \underline{X}(\mathbb{F}_{q^n})$. On the other hand, every geometric point of $\underline{X}_0$ lying over x is stable by the q^n-Frobenius, so x lies in the image of $\underline{X}_0(\mathbb{F}_{q^n}) \to \underline{X}(\mathbb{F}_{q^n})$. By Chebotarev density [Se, Theorem 7], the proposition follows for every $\epsilon < |G|^{-1}$. □

The main technical result of this section is the following.

Proposition 2.5. *Let $\phi\colon \underline{Y} \to \underline{X}$ be a dominant morphism of normal varieties over $\mathbb{F}_q$. Then there exists a positive integer m and for every positive integer n, there exist subsets $X_{n,i} \subset \underline{X}(\mathbb{F}_{q^n})$, $1 \le i \le m$, with the following properties.*

(1) For each i from 1 to m, we have $\liminf_n \frac{|X_{n,i}|}{|\underline{X}(\mathbb{F}_{q^n})|} > 0$.

(2) If $\theta\colon \underline{Z} \to \underline{X}$ is any dominant morphism of normal varieties over $\mathbb{F}_q$ such that

(a) For all $n \geq 1$, $\theta(\underline{Z}(\mathbb{F}_{q^n})) \supset \phi(\underline{Y}(\mathbb{F}_{q^n}))$, and

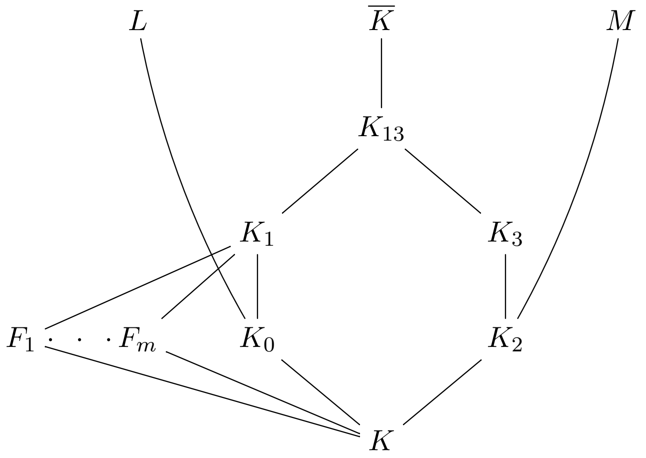
(b) there exists an integer $n_0 \geq 1$ such that $\theta(\underline{Z}(\mathbb{F}_{q^{n_0}})) \cap X_{n_0,i}$ is non-empty for each $i = 1, \ldots, m$,

then the generic fiber of θ is geometrically irreducible.

Proof. Let A, B, C denote the coordinate rings of $\underline{X}$, $\underline{Y}$, and $\underline{Z}$ respectively. Let K, L, and M be the fields of fractions of A, B, and C respectively. We regard B and C as A-algebras via ϕ and θ respectively, so L and M are extensions of K. Let K_0 and K_2 denote the separable closures of K in L and M respectively. As B and C are finitely generated $\mathbb{F}_q$-algebras, L and M are finitely generated K-extensions, and K_0 and K_2 are finite separable extensions of K. The claimed generic irreducibility of the generic fiber of θ amounts to the equality $K = K_2$. We define $\overline{K}$, K_1, K_3, and K_{13} as in Proposition 2.4.

Let $F_1, \ldots, F_m$ denote all subfields of K_1 over K, excluding K itself. Thus, we have the following diagram of fields:

(2.5)



For $0 \leq i \leq 3$, let A_i denote the integral closure of A in K_i and $\underline{X}_i =$ Spec A_i; likewise for A_{13} and $\underline{X}_{13}$. For $1 \leq i \leq m$, let D_i denote the integral closure of A in the field F_i, and let $\underline{W}_i :=$ Spec D_i. By (2.5), we have the following diagram of schemes:

$$\begin{tikzcd}
 & & \underline{Y} \arrow[dddr, bend right=10] & & & & \underline{Z} \arrow[dddl, bend left=10] \\
 & & & & \underline{X}_{13} \arrow[dl] \arrow[dr] & & \\
 & & & \underline{X}_1 \arrow[dlll] \arrow[dl] \arrow[d] & & \underline{X}_3 \arrow[d] & \\
\underline{W}_1 \arrow[drrrr] & \cdots & \underline{W}_m \arrow[drr] & \underline{X}_0 \arrow[dr] & & \underline{X}_2 \arrow[dl] & \\
 & & & & \underline{X} & &
\end{tikzcd}$$

Let $X_{n,i}$ denote the complement of the image of $\underline{W}_i(\mathbb{F}_{q^n})$ in $\underline{X}(\mathbb{F}_{q^n})$. By (2.3) and the Lang-Weil estimate, for $1 \leq i \leq m$,

$$|X_{n,i}| \geq \epsilon q^{\dim \underline{X}} > \frac{\epsilon}{2}|\underline{X}(\mathbb{F}_{q^n})| \tag{2.6}$$

if n is sufficiently large, which implies property (1).

Moreover, if $\theta\colon \underline{Z} \to \underline{X}$ is a dominant morphism satisfying condition (a), then for all $n \geq 1$, $\theta(\underline{Z}(\mathbb{F}_{q^n})) \supset \phi(\underline{Y}(\mathbb{F}_{q^n}))$, implying that

$$\begin{aligned}
&|\mathrm{im}(\underline{X}_1(\mathbb{F}_{q^n}) \to \underline{X}(\mathbb{F}_{q^n})) \setminus \mathrm{im}(\underline{X}_2(\mathbb{F}_{q^n}) \to \underline{X}(\mathbb{F}_{q^n}))| \\
&\qquad \leq |\mathrm{im}(\underline{X}_0(\mathbb{F}_{q^n}) \to \underline{X}(\mathbb{F}_{q^n})) \setminus \mathrm{im}(\underline{X}_2(\mathbb{F}_{q^n}) \to \underline{X}(\mathbb{F}_{q^n}))| \\
&\qquad = |\mathrm{im}(\underline{Y}(\mathbb{F}_{q^n}) \to \underline{X}(\mathbb{F}_{q^n})) \setminus \mathrm{im}(\underline{Z}(\mathbb{F}_{q^n}) \to \underline{X}(\mathbb{F}_{q^n}))| + o(q^{n \dim \underline{X}}) \\
&\qquad = |\phi(\underline{Y}(\mathbb{F}_{q^n})) \setminus \theta(\underline{Z}(\mathbb{F}_{q^n}))| + o(q^{n \dim \underline{X}}) \\
&\qquad = o(q^{n \dim \underline{X}}).
\end{aligned}$$

If $K_2 \neq K$, Proposition 2.4 implies that $K_1 \cap K_2$ must be a non-trivial extension of K, so $F_i \subset K_2$ for some $i \in [1, m]$. Thus, for n_0 as in (b),

$$\theta(\underline{Z}(\mathbb{F}_{q^{n_0}})) \subset \mathrm{im}(\underline{X}_2(F_{q^{n_0}}) \to \underline{X}(\mathbb{F}_{q^{n_0}})) \subset \mathrm{im}(\underline{W}_i(\mathbb{F}_{q^{n_0}}) \to \underline{X}(\mathbb{F}_{q^{n_0}})),$$

contrary to the assumption that $\theta(\underline{Z}(\mathbb{F}_{q^{n_0}})) \cap X_{n_0,i}$ is non-empty for each i. We conclude that $K_2 = K$, and the proposition follows.

□

3. Random walks

This section does not claim any original results. Its goal is to present well known ideas in probability theory in the form needed for the proof of Theorem 1.1.

For any positive integer d and non-negative integer l, we define $\mathsf{X}_{d,l}$ to be the convolution of l i.i.d. random variables on $\mathbb{Z}^d$, each uniformly distributed over the $2d$-element set $\{\pm e_1, \ldots, \pm e_d\}$, where $e_1, \ldots, e_d$ are the standard generators of $\mathbb{Z}^d$. When $d = 2$, we write X_l for short.

The main result in this section is the following.

Proposition 3.1. *For all $d \geq 2$ and $\epsilon > 0$, there exist M and N such that for $l \geq N$,*

$$\mathbf{P}[\mathsf{X}_{d,l} \in \bigcup_{i>M} i\mathbb{Z}^d] < \epsilon.$$

We begin with a general result.

Lemma 3.2. *Let G be a finite group and S a (not necessarily symmetric) set of generators. Let $\mathsf{S}_1, \mathsf{S}_2, \ldots$ be i.i.d. random variables on G with support S. Let $\mathsf{G}_l = \mathsf{S}_1 \cdots \mathsf{S}_l$. Suppose that there does not exist a homomorphism from G to any non-trivial cyclic group C mapping S to a single element. Then the limit as $l \to \infty$ of the distribution of G_l is the uniform distribution on G.*

Proof. Consider the Markov chain with state space G in which the probability of a transition from g to hg is $\mathbf{P}[\mathsf{S}_i = h]$. Since the uniform distribution is stationary, it suffices to check that this Markov chain is irreducible and periodic [LPW, Theorem 4.9]. Irreducibility is immediate from the condition that S generates G. If the Markov chain is periodic, then for some proper subset $X \subset G$ and some integer j, $s_1 \cdots s_j \in \mathrm{Stab}_G(X)$ for all $s_i \in S$. Let G_j denote the subgroup of G generated by

$$\{s_1 \ldots s_j \mid s_1, \ldots, s_j \in S\}.$$

As $G_j \subset \mathrm{Stab}_G(X) \subsetneq G$, G_j is a proper subgroup of G.

Consider the subgroup $\tilde{G}$ of $G \times \mathbb{Z}/j\mathbb{Z}$ generated by $\{(s, 1) \mid s \in S\}$. By definition, the kernel of projection on the second factor is G_j. By Goursat's Lemma, $\tilde{G}$ is the pullback to $G \times \mathbb{Z}/j\mathbb{Z}$ of the graph of an isomorphism between G/G_j and a quotient of $\mathbb{Z}/j\mathbb{Z}$. This identifies G/G_j with a non-trivial cyclic group C, and all elements of S map to the same generator of C, contrary to hypothesis. □

The remaining results in this section are needed for the proof of Proposition 3.1.

Lemma 3.3. *Let $p > 2$ be prime, k a positive integer, and $\epsilon > 0$. For l sufficiently large,*

$$\mathbf{P}[\mathsf{X}_l \in p^k\mathbb{Z}^2] < \frac{1+\epsilon}{p^{2k}}.$$

Proof. The image under (mod p^k) reduction of our random walk on $\mathbb{Z}^2$ is a random walk on $G = (\mathbb{Z}/p^k\mathbb{Z})^2$ with generating set $S = \{\pm 1, 0), (0, \pm 1)\}$. As differences between elements of S generate G, there is no proper coset of G which contains S. By Lemma 3.2, X_l becomes uniformly distributed (mod p^k) in the limit $l \to \infty$, which implies the lemma.

□

Lemma 3.4. *Let k be a positive integer, and $\epsilon > 0$. For l sufficiently large,*

$$\mathbf{P}[\mathsf{X}_l \in 2^k\mathbb{Z}^2] < \frac{2+\epsilon}{4^k}.$$

Proof. If l is odd, the probability that $\mathsf{X}_l \in 2\mathbb{Z}^2$ is zero. We therefore assume $l = 2l_0$, so X_l is the sum of l_0 i.i.d. random variables supported on

$$\{(\pm 2, 0), (0, \pm 2), (\pm 1, \pm 1), (0, 0)\}.$$

Reducing (mod 2^k), we obtain an irreducible aperiodic random walk on $\ker(\mathbb{Z}/2^k\mathbb{Z})^2 \to \mathbb{Z}/2\mathbb{Z}$, and the argument proceeds as before by Lemma 3.2. □

Proposition 3.5. *For all $\epsilon > 0$, there exist M and N such that for $l \geq N$,*

$$\mathbf{P}[\mathsf{X}_l \in \bigcup_{i>M} i\mathbb{Z}^2] < \epsilon.$$

Proof. By [LST, Proposition 3.2], if $p > 2$ is prime,

$$\mathbf{P}[\mathsf{X}_l \in p\mathbb{Z}^2 \setminus \{(0,0)\}] < \frac{4}{(p+1)^2}.$$

We choose $s \geq 2$ large enough that

$$\sum_{p>s} \frac{4}{(p+1)^2} < \frac{\epsilon}{2}$$

and choose k such that $\frac{3s}{4^k} < \frac{\epsilon}{2}$, so that if l is sufficiently large, the total probability that $\mathsf{X}_l \in p^k\mathbb{Z}^2$ for some $p \leq s$ is less than $\epsilon/2$. Note that this includes the probability that $\mathsf{X}_l = (0,0)$. Let M be larger than $s\prod_{p\leq s} p^k$. If $i > M$, then either i has a prime factor greater than s or a prime factor $\leq s$ with multiplicity at least k. The probability that there exists $i > M$ such that $G \in i\mathbb{Z}^2$ is therefore less than ϵ. □

We can now prove Proposition 3.1

Proof. The projection of a random walk on $\mathbb{Z}^d$ onto the first two coordinates gives a random walk on $\mathbb{Z}^2$ where each of the four possible non-zero steps are equally likely, but a zero step is also possible in the projection if $d > 2$. Since the projection of an element of $i\mathbb{Z}^d$ is an element of $i\mathbb{Z}^2$, the conditional probability that $\mathsf{X}_{d,l} \in \bigcup_{i>M} i\mathbb{Z}^d$ if we condition on at least l_0 steps which are non-zero in the projection is less than $\epsilon/2$ if l_0 is large enough. Given l_0 the probability that there are less than l_0 steps non-zero in the projection goes to 0 as l goes to infinity, so it can be taken to be less than $\epsilon/2$, implying that $\mathbf{P}[\mathsf{X}_{d,l} \in \bigcup_{i>M} i\mathbb{Z}^d] < \epsilon$. □

4. Proof of Theorem 1.1

We now prove the main theorem.

Proof. Fix a simple, simply-connected algebraic group $\underline{G}$ over a finite field $\mathbb{F}_q$. We will apply Proposition 2.5 in the case $\underline{X} = \underline{G}$, $\underline{Y} = \underline{G}$, $\underline{Z} = \underline{G}^d$, ϕ is the kth power map for some positive integer k, and θ is the evaluation map w for some $w \in F_d$ for which $\overline{w} = (a_1, \ldots, a_d)$ and $\gamma(a_1, \ldots, a_d) = k$. Given w, there exist integers $b_1, \ldots, b_d$ for which $k = a_1b_1 + \cdots + a_db_d$, so that

$$w_{\underline{G}(\mathbb{F}_{q^n})}(g^{b_1}, \ldots, g^{b_d}) = g^k$$

for all n and all $g \in \underline{G}(\mathbb{F}_{q^n})$, so $\phi(\underline{G}(\mathbb{F}_{q^n})) \subset \theta(\underline{G}(\mathbb{F}_{q^n}))$ for all $n \geq 1$.

By the main theorem of [MT], for every finite simple group Γ, there exists a 2-element generating set of Γ^N whenever $N \leq 2\sqrt{|\Gamma|}$. Let n_0 be any positive integer. Defining $N_0 := q^{n_0}$ and applying this to $\Gamma := \underline{G}(\mathbb{F}_{q^{n_0}})/Z(\underline{G}(\mathbb{F}_{q^{n_0}}))$, we see that Γ^{N_0} is d-generated. As $G := \underline{G}(\mathbb{F}_{q^{n_0}})^{N_0}$ is a perfect central extension of Γ^{N_0}, lifting any set of d generators of the latter to the former, we again obtain a generating set.

We denote by

$$S = \{(g_{i1}, \ldots, g_{iN_0}) \mid 1 \leq i \leq d\}$$

a generating set of G and consider an l-step random walk on this group with generating set S. By Lemma 3.2, for all $\delta > 0$, if l sufficiently large, the probability that the walk ends in any subset $T \subset G$ is at least

$$(1 - \delta/2)|T|/|G|.$$

We define $T := T_0 \cup \cdots \cup T_{\lfloor N_0/m \rfloor - 1}$, where

$$T_i := \underline{G}(\mathbb{F}_{q^{n_0}})^{im} \times X_{n_0,1} \times \cdots \times X_{n_0,m} \times \underline{G}(\mathbb{F}_{q^{n_0}})^{N_0-(i+1)m},$$

and $X_{n_0,i}$ are defined as in Proposition 2.5.

To estimate the probability that a uniformly randomly chosen element of G lies in T, we note that membership in the T_i are independent conditions. The probability of membership in each T_i is

$$\prod_{j=1}^{m} \frac{|X_{n_0,j}|}{|\underline{G}(\mathbb{F}_{q^{n_0}})|} \geq \frac{\epsilon^m}{2^m}$$

by (2.6). Therefore, the probability of membership in T for a uniformly chosen element of G is at least

$$1 - (1 - \epsilon^m/2^m)^{\lfloor N_0/m \rfloor}.$$

Taking n_0 (and therefore N_0) sufficiently large, we can guarantee this exceeds $1 - \delta/2$. Thus, the probability that the random walk ends in T is greater than $1 - \delta$.

For $1 \leq j \leq N_0$, let $\mathbf{g}_j = (g_{1j}, \ldots, g_{dj})$. We have seen that for a random word w of length n, the probability that $(w(\mathbf{g}_1), \ldots, w(\mathbf{g}_{N_0})) \in T$ is greater than $1 - \delta$. Membership in T implies membership in some T_i, which implies

$$w(\mathbf{g}_{im+1}) \in X_{n_0,1}, \ldots, w(\mathbf{g}_{im+m}) \in X_{n_0,m},$$

and therefore, by Proposition 2.5, if $\gamma(\overline{w}) = k$, then w is geometrically almost uniform for $\underline{G}$.

Thus, for each k, the probability is $\leq \delta$ that a random word w of length l satisfies $\gamma(\overline{w}) = k$ and that w is not geometrically almost uniform. By Proposition 3.1, for each fixed $\epsilon > 0$, there exists M such that if l is large enough, then the probability that $\gamma(\overline{w})$ is zero or greater than M for a word of length l is less than ϵ. Therefore, the probability that w is not geometrically almost uniform for $\underline{G}$ is less than $\epsilon + M\delta$. Choosing first ϵ and then δ, we can make this quantity as small as we wish, proving the theorem. □

We remark that the proof also shows that almost all words w are almost uniform for the family of groups $\{\underline{G}(\mathbb{F}_{q^n}) \mid n \geq 1\}$. The proof, together with that of [LST, Theorem 2], implies that w is almost always uniform for all finite simple groups with fixed root system and characteristic. For instance, almost all w are almost uniform for the Suzuki and Ree groups.

5. Questions

Question 5.1. If $\mathcal{G}$ is a simple, simply connected group scheme over $\mathbb{Z}$, does the probability that a random word is almost uniform for all simple groups of the form $\mathcal{G}(\mathbb{F}_q)/Z(\mathcal{G}(\mathbb{F}_q))$ go to 1?

It seems likely that the methods of this paper will allow one to prove this for all characteristics satisfying some Chebotarev-type condition, but can one do it for all characteristics simultaneously, or even a density one set of characteristics? Even more optimistically, one can ask:

Question 5.2. Does the probability that a random word is geometrically almost uniform for all simple, simply connected algebraic groups over finite fields go to 1?

Given an e-tuple of words $w_1, \ldots, w_e \in F_d$, for each G we can define a function $G^d \to G^e$, and we can ask about almost uniformity. In geometric families, this reduces again to the question of the geometric irreducibility of the generic fiber of the morphism $\underline{G}^d \to \underline{G}^e$ for simple, simply connected algebraic groups over finite fields. In the case that

$$\mathbb{Z}^d/\mathrm{Span}_{\mathbb{Z}}(\overline{w}_1, \ldots, \overline{w}_e) \cong \mathbb{Z}^{d-e},$$

the function $\underline{G}(\mathbb{F}_{q^n})^d \to \underline{G}(\mathbb{F}_{q^n})^e$ is surjective. Geometric irreducibility for such words follows as before.

Question 5.3. For $e < d$, does the probability that a random e-tuple of elements of F_d of length n is geometrically almost uniform go to 1 as $n \to \infty$?

Question 5.2 has an analogue for simple, simply connected compact Lie groups. As a special case, one can ask:

Question 5.4. Does the probability that for a random word w of length n

$$\lim_{m\to\infty} \|w_*U_{\mathrm{SU}(m)^d} - U_{\mathrm{SU}(m)}\| = 0$$

go to 1 as $n \to \infty$?

References

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

[GS] Shelly Garion and Aner Shalev: Commutator maps, measure preservation, and T-systems, *Trans. Amer. Math. Soc.* **361** (2009), 4631–4651.

[EGA IV$_2$] Alexandre Grothendieck, Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas. II. Inst. Hautes Études Sci. Publ. Math. No. **24**, 1965.

[EGA IV$_3$] Alexandre Grothendieck, Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas. III. Inst. Hautes Études Sci. Publ. Math. No. **28**, 1966.

[H] Philip Hall: The Eulerian Function of a Group, *Quart. J. Math. Oxford*, Vol. os-7, No. 1 (1936), 134–151.

[L] Michael Larsen: How random are word maps? Thin groups and superstrong approximation, 141–149, Math. Sci. Res. Inst. Publ., 61, Cambridge Univ. Press, Cambridge, 2014.

[LS1] Michael Larsen and Aner Shalev: Fibers of word maps and some applications, *J. Algebra* **354** (2012), 36–48.

[LS2] Michael Larsen and Aner Shalev: On the distribution of values of certain word maps, *Trans. Amer. Math. Soc.* **368** (2016), 1647–1661.

[LST] Michael Larsen, Aner Shalev, and Pham Huu Tiep: Probabilistic Waring problems for finite simple groups, *Annals of Math.*, to appear.

[LPW] David Levin, Yuval Peres, and Elizabeth Wilmer: Markov chains and mixing times. With a chapter by James G. Propp and David B. Wilson. American Mathematical Society, Providence, RI, 2009.

[LiS] Martin Liebeck and Aner Shalev: Fuchsian groups, finite simple groups, and representation varieties *Invent. Math.* **159** (2005), 317–367.

[LM] Alexander Lubotzky and Chen Meiri: Sieve methods in group theory I: Powers in linear groups. *J. Amer. Math. Soc.* **25** (2012), no. 4, 1119–1148.

[MT] Attila Maróti and Maria Chiara Tamburini: A solution to a problem of Wiegold. *Comm. Algebra* **41** (2013), no. 1, 34–49.

[P] Nicolas Pouyanne: On the number of permutations admitting an m-th root. *Electron. J. Combin.* **9** (2002), no. 1, Research Paper 3, 12 pp.

[PS] Ori Parzanchevski and Gili Schul: On the Fourier expansion of word maps. *Bull. Lond. Math. Soc.* **46** (2014), no. 1, 91–102.

[Se] Jean-Pierre Serre: Zeta and L functions. 1965 Arithmetical Algebraic Geometry (Proc. Conf. Purdue Univ., 1963), pp. 82–92, Harper & Row, New York.

[Sh] Aner Shalev: Some results and problems in the theory of word maps. Erdős centennial, 611–649, Bolyai Soc. Math. Stud., 25, János Bolyai Math. Soc., Budapest, 2013.

[T] Terence Tao: The Lang-Weil Bound, https://terrytao.wordpress.com/2012/08/31/the-lang-weil-bound/.

Email address: mjlarsen@indiana.edu

Department of Mathematics, Indiana University, Bloomington, IN 47405, U.S.A.