COMPOSITIO MATHEMATICA

Words have bounded width in $\mathbf{SL}(n, \mathbb{Z})$

Nir Avni and Chen Meiri

Compositio Math. **155** (2019), 1245–1258.

doi:10.1112/S0010437X19007334

Compositio Math. 155 (2019) 1245–1258
doi:10.1112/S0010437X19007334

Words have bounded width in $\mathrm{SL}(n, \mathbb{Z})$

Nir Avni and Chen Meiri

Abstract

We prove two results about the width of words in $\mathrm{SL}_n(\mathbb{Z})$. The first is that, for every $n \geqslant 3$, there is a constant $C(n)$ such that the width of any word in $\mathrm{SL}_n(\mathbb{Z})$ is less than $C(n)$. The second result is that, for any word w, if n is big enough, the width of w in $\mathrm{SL}_n(\mathbb{Z})$ is at most 87.

1. Introduction

A word is an element in a free group. Given a word $w = w(x_1, \ldots, x_d) \in F_d$ and a group Γ, we have the word map $w : \Gamma^d \to \Gamma$ defined by substitution. The set of w-values is the set

$$w(\Gamma) := \{w(g_1, \ldots, g_d), w(g_1, \ldots, g_d)^{-1} \mid g_i \in \Gamma\}.$$

Sets of word values in many families of groups have been extensively studied. See the book [Seg09] and the references therein for results on free and hyperbolic groups, nilpotent groups, p-adic analytic groups, and general finite groups (the last part is the main ingredient in the proof by Nikolov and Segal of Serre's conjecture that any finite-index subgroup in a finitely generated pro-finite group is open). We briefly describe some of the results that are relevant to this work.

Sets of word values in algebraic groups are large. Borel proved in [Bor83] that if w is a non-trivial word and G is a connected simple algebraic group defined over an algebraically closed field k, then $w(G(k))$ contains a Zariski open dense set. For Lie groups, the situation is more complicated. For example, Thom [Tho13, Corollary 1.2] and Lindenstrauss (unpublished) proved that sets of word values in the unitary group U_n can have arbitrarily small radii. Nevertheless, Borel's theorem implies that, for any semisimple Lie group $\mathbf{G}$ and any non-trivial word w, the set of word values $w(\mathbf{G})$ contains an open ball. It follows that, if $\mathbf{G}$ is compact, there is a constant C (depending on $\mathbf{G}$ and w) such that any element of $\mathbf{G}$ is a product of at most C word values. For arithmetic groups, sets of word values are very mysterious, even for simple words. For example, for every $n \geqslant 3$, the question whether every element of $\mathrm{SL}_n(\mathbb{Z})$ is a commutator is wide open. We do, however, know that the set of commutators in $\mathrm{SL}_n(\mathbb{Z})$ is quite large: Dennis and Vasserstein proved in [DV88] that every element in $\mathrm{SL}_n(\mathbb{Z})$ is a product of at most six commutators if n is large enough.

A remarkable theorem of Larsen and Shalev [LS09] says that a stronger statement holds for finite simple groups: for every non-trivial word w, if Γ is a large enough finite simple group, then every element of Γ is a product of two word values.

Our first result generalizes the theorem of Dennis and Vasserstein in a form similar to the theorem of Larsen and Shalev.

Received 21 March 2018, accepted in final form 2 November 2018, published online 13 June 2019.
2010 Mathematics Subject Classification 20H05, 11E57 (primary).
Keywords: arithmetic groups, word width.

Theorem 1.1. *There is a constant C with the following property: for any word w, there is n_w such that, for all $n > n_w$, every element of $\mathrm{SL}_n(\mathbb{Z})$ is a product of at most C elements of $w(\mathrm{SL}_n(\mathbb{Z}))$. In fact, C can be taken to be equal to 87.*

In general, we cannot expect the subgroup generated by $w(\mathrm{SL}_n(\mathbb{Z}))$ to be equal to $\mathrm{SL}_n(\mathbb{Z})$. We define the width of w in $\mathrm{SL}_n(\mathbb{Z})$ to be the minimum of the numbers C such that any element of $\langle w(\mathrm{SL}_n(\mathbb{Z}))\rangle$ is a product of at most C elements of $w(\mathrm{SL}_n(\mathbb{Z}))$. If no such number exists, we say that the width of w is infinite.

Our next theorem provides uniform bounds for width, for a fixed n.

Theorem 1.2. *For any $n \geqslant 3$ there is an integer $C = C(n)$ such that, for any word w, the width of w in $\mathrm{SL}_n(\mathbb{Z})$ is less than C.*

Remark 1.3. Theorems 1.1 and 1.2 are optimal in the following sense: for every C there are infinitely many pairs (n, w) such that the width of w in $\mathrm{SL}_n(\mathbb{Z})$ is greater than C. This easily follows from [Lub14, Theorem 1].

We do, however, have the following result which is uniform in n and w.

Theorem 1.4. *There is a constant C such that, for every any non-trivial word w, there is $d = d(w) \in \mathbb{Z}$ such that for every $n \geqslant 3$, every element of the d-congruence subgroup $\mathrm{SL}_n(\mathbb{Z}; d)$ is a product of at most C elements of $w(\mathrm{SL}_n(\mathbb{Z}))$. If n is large enough, C can be taken to be equal to 80.*

Remark 1.5. Let O be the ring of integers in a number field, let S be a finite set of primes of O, and let O_S denote the localization of O by S. The proofs below also show similar bounds for $\mathrm{SL}_n(O_S)$, but the bounds obtained by these proofs depend on O_S. While we do not know whether widths of words in $\mathrm{SL}_n(O_S)$ are bounded uniformly in O_S, [MRS18, Corollary 4.6] gives some indication that this is indeed the case. In another direction, we do not even know whether words in other higher-rank non-uniform lattices (especially non-split) have finite width. We exclude lattices of rank 1 from the discussion since these include free groups and hyperbolic groups for which the the width of every non-trivial word is infinite; see [MN14].

Remark 1.6. Let Γ be an irreducible arithmetic lattice in a higher-rank semisimple group G, and assume that there exist a compact semisimple Lie group K and a dense embedding $\pi : \Gamma \hookrightarrow K$ (this implies that Γ is cocompcat in G). By the result of Thomas and Lindenstrauss mentioned above, there are words $w \in F_2$ such that $\pi(w(\Gamma))$ is contained in an arbitrarily small neighborhood of the identity. It follows that the width of w can be arbitrarily large. This means that the analog of Theorem 1.2 fails for Γ. Noting that the image under π of any finite-index subgroup of Γ is dense, we get that Theorem 1.4 also fails. We do not know whether every word has finite width in higher-rank cocompact lattices, nor whether the analog of Theorems 1.1 holds for the class of cocompact lattices.

We briefly sketch the proofs of the main theorems. For $n \geqslant 2$ and $q \in \mathbb{Z}$, denote by $U_n(\mathbb{Z}; q)$ the subgroup of all unipotent upper triangular matrices in $\mathrm{SL}_n(\mathbb{Z})$ whose off-diagonal entries are divisible by q. Denote similarly $L_n(\mathbb{Z}; q)$, replacing upper triangular by lower triangular. Finally, for a group G, a subset $X \subset G$, and a natural number n, we denote $X^n = \{x_1 \cdots x_n \mid x_i \in X \cup \{1\}\}$.

The main step is to prove the following theorem.

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

THEOREM 1.7. *There is a constant C such that, for any $n \geqslant 3$ and any $q \in \mathbb{Z}$, $(U_n(\mathbb{Z};q)L_n(\mathbb{Z};q))^C$ is a finite-index subgroup of $\mathrm{SL}_n(\mathbb{Z})$.*

Theorem 1.7 is proved by induction on n in § 2. The case $n = 3$ is essentially due to Carter, Keller, and Paige (see [Wit07] for an exposition of the proof). The argument for the induction step follows Dennis and Vaserstein [DV88].

Given Theorem 1.7, we deduce Theorem 1.4 without the explicit bound on C in § 3. A short argument implies that $w(\mathrm{SL}_3(\mathbb{Z}))^2$ contains an elementary matrix. Using various embeddings of $\mathrm{SL}_3(\mathbb{Z})$ into $\mathrm{SL}_n(\mathbb{Z})$, we show that $w(\mathrm{SL}_n(\mathbb{Z}))^{C'}$ contains $U_n(\mathbb{Z};q)$ and $L_n(\mathbb{Z};q)$ for some C and q. Theorems 1.1 and 1.2 follow from Theorem 1.4, a p-adic open mapping theorem, and the Larsen–Shalev theorem [LS09].

2. Proof of Theorem 1.7

In this section, we prove Theorem 1.7. We start by setting up the notation and recalling some facts.

DEFINITION 2.1. Let A be a commutative ring with unit, let I be an ideal in A, and let $n \geqslant 2$ be an integer.

(1) $\mathrm{SL}_n(A;I)$ is the subgroup of $\mathrm{SL}_n(A)$ consisting of the matrices which are congruent to the identity matrix modulo the ideal I. The subgroup $\mathrm{SL}_n(A;I)$ is called the I-congruence subgroup of $\mathrm{SL}_n(A)$.
(2) $U_n(A;I)$ is the subgroup of $\mathrm{SL}_n(A;I)$ consisting of unipotent upper triangular matrices.
(3) $L_n(A;I)$ is the subgroup of $\mathrm{SL}_n(A;I)$ consisting of unipotent lower triangular matrices.
(4) In the case where $A = \mathbb{Z}$ and $I = q\mathbb{Z}$ we sometimes write $\mathrm{SL}_n(\mathbb{Z};q)$, $U_n(\mathbb{Z};q)$ and $L_n(\mathbb{Z};q)$ instead of $\mathrm{SL}_n(\mathbb{Z};I)$, $U_n(\mathbb{Z};I)$ and $L_n(\mathbb{Z};I)$.

DEFINITION 2.2. Let A be a commutative ring with a unit, let I be an ideal in A, and let $n \geqslant 2$ be an integer.

(1) For $x \in A$ and $1 \leqslant i \neq j \leqslant n$, let $e_{i,j}(x)$ denote the $n \times n$ matrix with ones along the diagonal, x as (i,j)th entry, and zero in all other entries.
(2) Denote by $E(n,A;I)$ the subgroup generated by the elementary matrices $e_{i,j}(x)$, for $x \in I$. We will write $E(n,A)$ instead of $E(n,A;A)$.
(3) Denote by $E^{\triangleleft}(n,A;I)$ the normal subgroup of $E(n,A)$ generated by $E(n,A;I)$.
(4) In the case where $A = \mathbb{Z}$ and $I = q\mathbb{Z}$ we sometimes write $E(n,\mathbb{Z};q)$ and $E^{\triangleleft}(n,\mathbb{Z};q)$ instead of $E(n,\mathbb{Z};I)$ and $E^{\triangleleft}(n,\mathbb{Z};I)$.

The following result is [Tit76, Proposition 2].

PROPOSITION 2.3 (Tits). *If A is a commutative ring, I is an ideal of A, and $n \geqslant 3$, then $E^{\triangleleft}(n,A;I^2) \subseteq \langle U_n(A;I) \cup L_n(A;I)\rangle$.*

The following theorem is proved in [Wit07].

THEOREM 2.4 (Carter, Keller, and Paige). *There is a first-order statement φ in the language of rings with the following properties:*

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

(1) *φ holds in $\mathbb{Z}$;*
(2) *if A is a ring satisfying φ and I is an ideal of A, then* $[\mathrm{SL}_n(A;I) : E^{\triangleleft}(n,A;I)] \leqslant 2\cdot 8!$.

Remark 2.5. Theorem 2.4 is proved in [Wit07]. More precisely, if we take φ to be the conjunction of the conditions $SR_{1\frac{1}{2}}$, $\mathrm{Gen}(2\cdot 8!,1)$, $\mathrm{Exp}(2\cdot 8!,2)$ (see [Wit07, Definitions 2.10, 3.2, 3.6]), then [Wit07, Lemma 2.13, Corollary 3.5, Theorem 3.9] imply that $\mathbb{Z}$ satisfies φ and (2) is [Wit07, Theorem 3.12].

Corollary 2.6. *There is a constant $C = C(n)$ such that the following holds: for any $q \in \mathbb{N}^+$, there are $g_1,\dots,g_{2\cdot 8!} \in \mathrm{SL}_n(\mathbb{Z};q^2)$ such that $\mathrm{SL}_n(A;q^2)$ is contained in the union of the translations by $g_1,\dots,g_{2\cdot 8!}$ of the set $(U_n(\mathbb{Z};q)L_n(\mathbb{Z};q))^C$.*

Proof. Let A is a ring which is elementarily equivalent to $\mathbb{Z}$ (i.e. satisfies the same first-order sentences as $\mathbb{Z}$) and let I be an ideal of A. Proposition 2.3 and Theorem 2.4 imply that

$$[\mathrm{SL}_n(A;I^2) : \mathrm{SL}_n(A;I^2) \cap \langle U_n(A;I)L_n(A;I)\rangle] \leqslant 2\cdot 8!. \tag{1}$$

Assume the corollary is false. Then, for every $k \in \mathbb{N}$, there are $q_k \in \mathbb{N}^+$ and matrices $g_{k,1},\dots,g_{k,2\cdot 8!+1} \in \mathrm{SL}_n(\mathbb{Z};q_k^2)$ such that $(g_{k,i})^{-1}g_{k,j} \notin (U_n(\mathbb{Z};q_k)L_n(\mathbb{Z};q_k))^k$ if $i \neq j$.

Choose a non-principal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, and let A be the ultrapower of $\mathbb{Z}$ over $\mathcal{U}$. Then A is elementarily equivalent to $\mathbb{Z}$ and $\mathrm{SL}_n(A)$ is isomorphic to the ultrapower of $\mathrm{SL}_n(\mathbb{Z})$ over $\mathcal{U}$. Let I be the ideal of A represented by $\prod_k q_k\mathbb{Z}$, and for every $1 \leqslant i \leqslant k$, let $g_i \in \mathrm{SL}_n(A;I^2)$ be the element represented by $(g_{k,i})_k$. Then $g_1,\dots,g_{2\cdot 8!+1}$ belong to different cosets of $\langle U_n(A;I)L_n(A;I)\rangle$, contradicting (1). □

The following two technical lemmas will be needed in the proof of Proposition 2.9 below.

Lemma 2.7. *Let G be a group, and let $X \subset G$ be a symmetric set such that there are d translates of X that cover G. Then X^{4d+2} is a group.*

Proof. Denote $Y = X^2$. Then $1 \in Y$ and there are d translates of Y that cover G. Since $1 \in Y$, $Y^k \subseteq Y^{k+1}$ for every k. It is enough to show that $Y^k = Y^{k+1}$ for some $k \leqslant 2d+1$. Suppose that $G = \bigcup_{i=1}^d g_iY$ for some $g_1,\dots,g_d \in G$. We can assume that $g_1 = 1$. For every k, if $Y^k \neq Y^{k+1}$, choose $h \in Y^{k+1} \setminus Y^k$. By assumption, there is i such that $h \in g_iY$. Then $g_i \in Y^{k+2}$ but $g_i \notin Y^k$. By induction we see that if $Y^{2k-1} \neq Y^{2k}$ for some $1 \leqslant k$, then Y^{2k+1} contains at least k distinct g_i. This implies that $Y^{2d+1} = Y^{2d+2}$. □

Lemma 2.8. *Let $K \subseteq H \subseteq G$ be groups such that $[H:K] < \infty$. Let $X \subseteq G$ be a symmetric subset. Assume that $HX = G$ and that $K \subseteq X$. Then $X^{4[H:K]}$ is a subgroup.*

Proof. Since $1 \in K \subseteq X$, the sets $(X^nK \cap H) \subseteq H$ are non-decreasing. Hence, there is $n \leqslant 4[H:K] - 3$ such that

$$X^nK \cap H = X^{n+1}K \cap H = X^{n+2}K \cap H = X^{n+3}K \cap H = X^{n+4}K \cap H.$$

Since $HX = G$, we have $X^{n+3} \subseteq (X^{n+4} \cap H)X$. Thus,

$$X^{n+3} \subseteq (X^{n+4} \cap H)X \subseteq (X^{n+4}K \cap H)X \subseteq (X^nK \cap H)X \subseteq X^{n+2},$$

so X^{n+2} is a group. □

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

PROPOSITION 2.9. *There is a constant $D = D(n)$ such that, for any $q \in \mathbb{N}^+$, the set $(U_n(\mathbb{Z};q)L_n(\mathbb{Z};q))^D$ is a group, and, therefore, equal to $\langle U_n(\mathbb{Z};q)L_n(\mathbb{Z};q)\rangle$.*

Proof. For any k, the set $(L_n(\mathbb{Z};q)U_n(\mathbb{Z};q))^{k+1}$ contains the symmetric subset $(L_n(\mathbb{Z};q)U_n(\mathbb{Z};q))^k \cup (U_n(\mathbb{Z};q)L_n(\mathbb{Z};q))^k$. Corollary 2.6 and Lemma 2.7 imply that there is a constant D' such that $(L_n(\mathbb{Z};q)U_n(\mathbb{Z};q))^{D'}$ contains a subgroup $S(I)$ of $\mathrm{SL}_n(\mathbb{Z};q^2)$ of index at most $2\cdot 8!$.

Note that $\mathrm{SL}_n(\mathbb{Z},q)/\mathrm{SL}_n(\mathbb{Z},q^2)$ is abelian so $\mathrm{SL}_n(\mathbb{Z},q^2)L_n(\mathbb{Z};q)U_n(\mathbb{Z};q)$ is a subgroup of $\mathrm{SL}_n(\mathbb{Z})$. The desired result follows by applying Lemma 2.8 to $K = S(I)$, $H = \mathrm{SL}_n(\mathbb{Z};q^2)$, $G = \mathrm{SL}_n(\mathbb{Z},q^2)L_n(\mathbb{Z};q)U_n(\mathbb{Z};q)$, and $X = (L_n(\mathbb{Z};q)U_n(\mathbb{Z};q))^{D'} \cup (U_n(\mathbb{Z};q)L_n(\mathbb{Z} : q))^{D'} \subseteq (L_n(\mathbb{Z};q)U_n(\mathbb{Z};q))^{D'+1}$. □

In order to prove Theorem 1.7 we have to show that the constant $D(n)$ in Proposition 2.9 can be made independent of n. The following technical generalization of Proposition 2.3 is needed.

LEMMA 2.10. *Let $n \geqslant 3$ and let I be an ideal in a commutative ring A. Then $E^{\triangleleft}(n+1;A,I^2)$ is contained in the subgroup*

$$K(I) := \langle e_{i,j}(a) \mid 1 \leqslant i \neq j \leqslant n+1, \{i,j\} \neq \{1,n+1\}, a \in I\rangle.$$

Proof. We follow the proof of [Tit76, Proposition 2.3].

Let $1 \leqslant i \neq j \leqslant n+1$, $1 \leqslant r \neq s \leqslant n+1$, and $a,b \in A$. Recall the following relations:

$$\begin{cases} e_{r,s}(b)e_{i,j}(a)e_{r,s}(b)^{-1} = e_{i,j}(a)e_{i,s}(-ab) & \text{if } j = r \text{ and } i \neq s, \\ e_{r,s}(b)e_{i,j}(a)e_{r,s}(b)^{-1} = e_{i,j}(a)e_{r,j}(ab) & \text{if } j \neq r \text{ and } i = s, \\ e_{r,s}(b)e_{i,j}(a)e_{r,s}(b)^{-1} = e_{i,j}(a) & \text{if } j \neq r \text{ and } i \neq s. \end{cases} \tag{2}$$

For every $1 \leqslant i \neq j \leqslant n+1$, denote $F_{i,j}(I^2) := \langle e_{i,j}(a), e_{j,i}(a) \mid a \in I^2\rangle$. Let $F^{\triangleleft}_{i,j}(I^2)$ be the minimal normal subgroup of $F_{i,j} := F_{i,j}(A)$ which contains $F_{i,j}(I^2)$. Define $F^{\triangleleft}(I^2) := \langle F^{\triangleleft}_{i,j}(I^2) \mid 1 \leqslant i \neq j \leqslant n+1\rangle$. Equation (2) implies that for every $1 \leqslant i \neq j \leqslant n+1$ and every $a \in A$, $e_{i,j}(a)F^{\triangleleft}(I^2)e_{i,j}(a)^{-1} = F^{\triangleleft}(I^2)$. Thus $F^{\triangleleft}(I^2)$ is a normal subgroup of $E(n+1,A)$ containing all $e_{i,j}(a)$, $a \in I^2$, so it must be equal to $E^{\triangleleft}(n+1,A,I^2)$. Thus, in order to finish the proof it is enough to show that for every $1 \leqslant i < j \leqslant n+1$, $F^{\triangleleft}_{i,j}(I^2) \subseteq K(I)$.

Let $E^+(n,A;I)$ and $E^-(n,A;I)$ be the images of $E(n,A,I)$ in $\mathrm{SL}_{n+1}(A)$ under the embeddings $M \mapsto \left(\begin{smallmatrix} M & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $M \mapsto \left(\begin{smallmatrix} 1 & 0 \\ 0 & M \end{smallmatrix}\right)$. By applying Proposition 2.3 with respect to $E^+(n,A;I)$ and $E^-(n,A;I)$, we see that $K(I)$ contains $F^{\triangleleft}_{i,j}(I^2)$ for every $1 \leqslant i < j \leqslant n+1$ such that $(i,j) \neq (1,n+1)$.

Equation (2) implies that $K(I)$ in normalized by $e_{1,n+1}(a)$ and $e_{n+1,1}(a)$ for every $a \in R$. For every $a,b \in I$, $e_{1,n+1}(ab) = [e_{1,2}(a), e_{2,n+1}(-b)] \in K(I)$ and $e_{n+1,1}(ab) = [e_{n+1,2}(a), e_{2,1}(-b)] \in K(I)$. Thus, $F^{\triangleleft}_{1,n+1}(I^2) \leqslant K(I)$. □

The next lemma is the key ingredient in the proof of Theorem 1.7.

LEMMA 2.11. *Let $n \geqslant 3$ and $q \in \mathbb{N}^+$ and assume that $(U_n(\mathbb{Z};q)L_n(\mathbb{Z};q))^D = \langle U_n(\mathbb{Z};q)L_n(\mathbb{Z};q)\rangle$. Then for every $m \geqslant n$, $(U_m(\mathbb{Z};q)L_m(\mathbb{Z};q))^D = \langle U_m(\mathbb{Z};q)L_m(\mathbb{Z};q)\rangle$.*

Proof. The proof follows [DV88, proof of Lemma 7] and is by induction on m. The base case $m = n$ is clear. It remains to show that if the claim is true for some $m \geqslant 3$ then it is also true for $m+1$.

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

Let $T := (U_{m+1}(\mathbb{Z};q)L_{m+1}(\mathbb{Z};q))^D$ and $H = \{g \in \mathrm{SL}_{m+1}(\mathbb{Z}) \mid gT = T\}$. Since H is a group, it is enough to prove that H contains both $U_{m+1}(\mathbb{Z};q)$ (which is clear) and $L_{m+1}(\mathbb{Z};q)$.

We embed $L_m(\mathbb{Z};q)$ and $U_m(\mathbb{Z};q)$ in $\mathrm{SL}_{m+1}(\mathbb{Z};q)$ by the embedding $M \mapsto \left(\begin{smallmatrix} M & 0 \\ 0 & 1 \end{smallmatrix}\right)$. We denote the abelian group $\langle e_{i,m+1}(a) \mid 1 \leqslant i \leqslant m,\ a \in q\mathbb{Z}\rangle$ by $C_{m+1}(\mathbb{Z};q)$ and the abelian group $\langle e_{m+1,i}(q) \mid 1 \leqslant i \leqslant m,\ a \in q\mathbb{Z}\rangle$ by $R_{m+1}(\mathbb{Z};q)$. We have that $U_{m+1}(\mathbb{Z};q) = U_m(\mathbb{Z};q) \ltimes C_m(\mathbb{Z};q)$, that $L_{m+1}(\mathbb{Z};q) = L_m(\mathbb{Z};q) \ltimes R_m(\mathbb{Z};q)$, and that $U_m(\mathbb{Z};q)$ and $L_m(\mathbb{Z};q)$ each normalize both $C_m(\mathbb{Z};q)$ and $R_m(\mathbb{Z};q)$. The induction hypothesis implies that

$$\begin{aligned}
&L_m(\mathbb{Z}:q)(U_{m+1}(\mathbb{Z}:q)L_{m+1}(\mathbb{Z};q))^D \\
&\quad = L_m(\mathbb{Z}:q)(U_m(\mathbb{Z}:q)C_m(\mathbb{Z}:q)L_m(\mathbb{Z}:q)R_m(\mathbb{Z}:q))^D \\
&\quad = L_m(\mathbb{Z}:q)(U_m(\mathbb{Z}:q)L_m(\mathbb{Z}:q))^D \cdot (C_m(\mathbb{Z}:q)R_m(\mathbb{Z}:q))^D \\
&\quad = (U_m(\mathbb{Z}:q)L_m(\mathbb{Z}:q))^D \cdot (C_m(\mathbb{Z}:q)R_m(\mathbb{Z}:q))^D \\
&\quad = (U_m(\mathbb{Z}:q)C_m(\mathbb{Z}:q)L_m(\mathbb{Z}:q)R_m(\mathbb{Z}:q))^D \\
&\quad = (U_{m+1}(\mathbb{Z}:q)L_{m+1}(\mathbb{Z}:q))^D.
\end{aligned}$$

Hence, $L_m(\mathbb{Z}:q) \subseteq H$, that is, for every $1 \leqslant i < j \leqslant m$ and $a \in I$, we have $e_{j,i}(a) \in H$. Arguing similarly using the embedding $M \mapsto \left(\begin{smallmatrix} 1 & 0 \\ 0 & M \end{smallmatrix}\right)$, we get that $e_{j,i}(a) \in H$, for every $2 \leqslant i < j \leqslant m+1$ and $a \in I$. It remains to show that for every $a \in I$, $e_{m+1,1}(a) \in H$.

The main theorem of [Men65] says that $E^{\triangleleft}(n,\mathbb{Z},k) = \mathrm{SL}_n(\mathbb{Z},k)$ for every $k \in \mathbb{N}^+$. Thus, Lemma 2.10 implies that $\mathrm{SL}_{m+1}(\mathbb{Z};q^2) = E^{\triangleleft}(m+1,\mathbb{Z};q^2) \subseteq H$. Since $\mathrm{SL}_{m+1}(\mathbb{Z};q)/\mathrm{SL}_{m+1}(\mathbb{Z};q^2)$ is abelian, $e_{m+1,1}(a)U_{m+1}(\mathbb{Z}:q)e_{m+1,1}(a)^{-1} \subseteq \mathrm{SL}_{m+1}(\mathbb{Z};q^2) \cdot U_{m+1}(\mathbb{Z}:q)$, for every $a \in I$. It follows that, for every $a \in I$,

$$\begin{aligned}
&e_{m+1,1}(a)(U_{m+1}(\mathbb{Z}:q)L_{m+1}(\mathbb{Z}:q))^D \\
&\quad = e_{m+1,1}(a)U_{m+1}(\mathbb{Z}:q)e_{m+1,1}(a)^{-1}e_{m+1,1}(a)L_{m+1}(\mathbb{Z}:q)(U_{m+1}(\mathbb{Z}:q)L_{m+1}(\mathbb{Z}:q))^{D-1} \\
&\quad \subseteq \mathrm{SL}_{m+1}(\mathbb{Z};q^2) \cdot U_{m+1}(\mathbb{Z}:q)L_{m+1}(\mathbb{Z}:q)(U_{m+1}(\mathbb{Z}:q)L_{m+1}(\mathbb{Z}:q))^{D-1} \\
&\quad = (U_{m+1}(\mathbb{Z}:q)L_{m+1}(\mathbb{Z}:q))^D.
\end{aligned}$$

In particular, for every $a \in I$, $e_{m+1,1}(a) \in H$. Hence, $L_{m+1}(\mathbb{Z}:q) \subseteq H$ as desired. □

Proof of Theorem 1.7. Proposition 2.9 implies that there is a constant $C = D(3)$ such that $(U_3(\mathbb{Z};q)L_3(\mathbb{Z};q))^C = \langle U_3(\mathbb{Z};q)L_3(\mathbb{Z};q)\rangle$. Lemma 2.11 implies that $(U_n(\mathbb{Z};q)L_n(\mathbb{Z};q))^C = \langle U_n(\mathbb{Z};q)L_n(\mathbb{Z};q)\rangle$ for every $n \geqslant 3$. Proposition 2.3 implies that $(U_n(\mathbb{Z};q)L_n(\mathbb{Z};q))^C$ contains the congruence subgroup $\mathrm{SL}_n(\mathbb{Z};q^2)$ and this subgroup has a finite index in $\mathrm{SL}(n,\mathbb{Z})$. □

3. Proof of Theorem 1.4 without an explicit bound

We will need the following lemma, which we state without a proof.

Lemma 3.1. *All upper-triangular matrices $g \in U_n(\mathbb{Z};q)$ such that $g_{i,i+1} = q$, for all i, are conjugate.*

Proof of Theorem 1.4 without an explicit bound. Identify $\mathrm{SL}_2(\mathbb{Z})$ with its image in $\mathrm{SL}_3(\mathbb{Z})$ under the embedding $M \mapsto \left(\begin{smallmatrix} M & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Since $\mathrm{SL}_2(\mathbb{Z})$ contains a non-abelian free group there exists $\pm I_2 \neq g \in w(\mathrm{SL}_2(\mathbb{Z}))$. There exists $h \in \langle e_{1,3}(1), e_{2,3}(1)\rangle$ such that $[g,h] = g^{-1}h^{-1}gh$ is a non-trivial element and this element is conjugate to $e_{1,3}(q)$ for some positive $q \in \mathbb{N}$. For the chosen g and h, we have $[g,h]^n = [g,h^n] \in w(\mathrm{SL}_3(\mathbb{Z}))^2$. Since $w(\mathrm{SL}_3(\mathbb{Z}))^2$ is a normal subset, $\langle e_{1,3}(q)\rangle \subseteq w(\mathrm{SL}_3(\mathbb{Z}))^2$. We will show that the statement of Theorem 1.4 holds with respect to $d = q^2$.

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

We claim that for any integers $a_1,\dots,a_{n-1}$, there is $g \in w(\mathrm{SL}_n(\mathbb{Z}))^8 \cap U_n(\mathbb{Z};q)$ such that for every i, $g_{i,i+1} = qa_i$. Using two different embeddings of the group $\mathrm{SL}_3(\mathbb{Z}) \times \cdots \times \mathrm{SL}_3(\mathbb{Z})$ ($\lfloor n/3 \rfloor$ times) into $\mathrm{SL}_n(\mathbb{Z})$ as block-diagonal matrices, we get that there is a matrix $g^1 \in w(\mathrm{SL}_n(\mathbb{Z}))^4 \cap U_n(\mathbb{Z};q)$ such that $g^1_{i,i+1} = qa_i$ if $i \equiv 1 \pmod 3$ and $g^1_{i,i+1} = 0$ otherwise. Using one embedding of the group $\mathrm{SL}_3(\mathbb{Z}) \times \cdots \times \mathrm{SL}_3(\mathbb{Z})$ ($\lfloor n/3 \rfloor$ times) into $\mathrm{SL}_n(\mathbb{Z})$ as block-diagonal matrices, we get that there is a matrix $g^2 \in w(\mathrm{SL}_n(\mathbb{Z}))^2 \cap U_n(\mathbb{Z};q)$ such that $g^2_{i,i+1} = qa_i$ if $i \equiv 2 \pmod 3$ and $g^2_{i,i+1} = 0$ otherwise. Similarly, there is $g^3 \in w(\mathrm{SL}_n(\mathbb{Z}))^2 \cap U_n(\mathbb{Z};q)$ such that $g^3_{i,i+1} = qa_i$ if $i \equiv 3 \pmod 3$ and $g^3_{i,i+1} = 0$ otherwise. The matrix $g = g^0 g^1 g^2 \in w(\mathrm{SL}_n(\mathbb{Z}))^8 \cap U_n(\mathbb{Z};q)$ satisfies $g_{i,i+1} = qa_i$. The proof of the claim in now complete.

It follows from Lemma 3.1 that $w(\mathrm{SL}_n(\mathbb{Z}))^8$ contains all elements $g \in U_n(\mathbb{Z};q)$ such that $g_{i,i+1} = q$ for every i.

Next, we claim that $U_n(\mathbb{Z};q) \subseteq w(\mathrm{SL}_n(\mathbb{Z}))^{16}$. Indeed, let $h \in U_n(\mathbb{Z};q)$. There is an element $f \in w(\mathrm{SL}_n(\mathbb{Z}))^8 \cap U_n(\mathbb{Z};q)$ such that for every i, $f_{i,i+1} = q - h_{i,i+1}$. Then $hf \in U_n(\mathbb{Z};q)$ and, for every i, $(fh)_{i,i+1} = q$, so $fh \in w(\mathrm{SL}_n(\mathbb{Z}))^8$. Since $w(\mathrm{SL}_n(\mathbb{Z}))^8$ is symmetric, it follows that $h \in w(\mathrm{SL}_n(\mathbb{Z}))^{16}$. Similarly, $L_n(\mathbb{Z};q) \subseteq w(\mathrm{SL}_n(\mathbb{Z}))^{16}$.

By Theorem 1.7, there is a constant C (independent of q) such that

$$\langle U_n(\mathbb{Z};q)U_n(\mathbb{Z};q)\rangle = (U_n(\mathbb{Z};q)U_n(\mathbb{Z};q))^C \subseteq w(\mathrm{SL}_n(\mathbb{Z}))^{32C}.$$

Propositon 2.3 implies that $\mathrm{SL}_n(\mathbb{Z},q^2) \leqslant \langle U_n(\mathbb{Z};q)U_n(\mathbb{Z};q)\rangle$. □

4. Proof of Theorems 1.2 and 1.1

In order to deduce Theorems 1.1 and 1.2 from Theorem 1.4, we need to study word values in $\mathrm{SL}_n(\mathbb{Z}/q\mathbb{Z})$ uniformly in q. Equivalently, we need to study word values in $\mathrm{SL}_n(\widehat{\mathbb{Z}})$ where $\widehat{\mathbb{Z}}$ is the pro-finite completion of $\mathbb{Z}$. We will use a version of the open mapping theorem which is well known, but for which we were unable to find a reference.

For $a \in \mathbb{Z}_p^n$, denote $\|a\| = \max\{|a_i|_p\}$, where $|a|_p$ is the p-adic valuation of a. The function $d(a,b) = \|a-b\|$ is a metric on $\mathbb{Z}_p^n$. Let $X \subset \mathbb{A}^n_{\mathbb{Z}_p}$ be an affine $\mathbb{Z}_p$-scheme, that is, the zero locus of a collection of polynomials in $\mathbb{Z}_p[x_1,\dots,x_n]$. We denote the set of solutions of X with coordinates in $\mathbb{Z}_p$ by $X(\mathbb{Z}_p)$. The restriction of d to $X(\mathbb{Z}_p)$ is a metric on $X(\mathbb{Z}_p)$.[1] Let $\mathbb{Z}_p[X]$ be the ring of regular functions on X (the restrictions of polynomials with $\mathbb{Z}_p$ coefficients to X). For $f \in \mathbb{Z}_p[X]$, we define $\mathrm{val}_p(f) = \max\{k \mid f \in p^k\mathbb{Z}_p[X]\}$. More generally, if $f : X \to Y$ is a map of affine $\mathbb{Z}_p$-schemes, we define $\mathrm{val}_p(f)$ as the minimum of the valuations of its coordinates. Note that if $\mathrm{val}(f) \geqslant k$, then $d(f(a),f(b)) \leqslant p^{-k}d(a,b)$, for every $a,b \in X(\mathbb{Z}_p)$.

Recall that X is called smooth at $a \in X(\mathbb{Z}_p)$ if there are $\psi_1,\dots,\psi_c \in \mathbb{Z}_p[x_1,\dots,x_n]$ such that X is the common zero locus of ψ_i and the reductions of $\nabla\psi_i(a)$ modulo p are linearly independent. In this case $n-c$ is called the dimension of X at a.

LEMMA 4.1. *Let $X \subseteq \mathbb{A}^n_{\mathbb{Z}_p}$ be a $\mathbb{Z}_p$-scheme and $a \in X(\mathbb{Z}_p)$. Assume that X is smooth in a. Then there is a subset $S \subset \{1,\dots,n\}$ such that the coordinate projection $\pi : \mathbb{Z}_p^n \to \mathbb{Z}_p^S$ satisfies the following statements:*

(1) *the restriction of π to $X(\mathbb{Z}_p) \cap B(a,p^{-1})$ is one-to-one, where $B(a,p^{-1})$ is the closed ball of radius p^{-1} around a;*
(2) $\pi(T_aX(\mathbb{Z}_p)) = \mathbb{Z}_p^S$.

[1] This metric is independent of the affine embedding, but we will not use this fact.

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

Proof. Let ψ_i be as in the definition of smoothness. After permutation of the indices, we can assume that the $c \times c$ matrix $\left(\frac{\partial \psi_i}{\partial x_j}(a)\right)$ is invertible over $\mathbb{Z}_p$. For any $f \in \mathbb{Z}_p[x_1, \ldots, x_n]$ and any $a, b \in \mathbb{Z}_p^n$ with $0 < d(a,b) < 1$, we have $|f(a) - f(b) - \langle \nabla f(a), a - b\rangle| \leqslant \|a-b\|^2 < \|a-b\|$. If $a, b \in X(\mathbb{Z}_p)$ and $d(a,b) < 1$, we have $\psi_i(a) = \psi_i(b) = 0$, so $|\langle \nabla \psi_i(a), a-b\rangle| < \|a-b\|$. If, in addition, $\pi(a) = \pi(b)$, write $a - b = (v, 0)$, where $v \in p\mathbb{Z}_p^c$ and then

$$\left\| \left(\frac{\partial \psi_i}{\partial x_j}(a) \right) v \right\| = \max\{|\langle \nabla \psi_i(a), a-b\rangle|\} < \|v\|.$$

Since invertible matrices do not decrease norm, this is a contradiction. This completes the proof of statement (1). Denoting $S := \{c+1, \ldots, n\}$, statement (2) readily follows form the assumption that $\left(\frac{\partial \psi_i}{\partial x_j}(a)\right)$ is invertible. □

Lemma 4.2. *Let X, Y be affine $\mathbb{Z}_p$-schemes. Let $f : X \to Y$ be a morphism, let $a \in X(\mathbb{Z}_p)$, and let $k \geqslant 0$ be an integer. Suppose that the following statements hold:*

(1) $\mathrm{val}_p(f) \geqslant k$;
(2) $df(a)(T_a X(\mathbb{Z}_p)) \supseteq p^k T_{f(a)} Y(\mathbb{Z}_p)$;
(3) *X is smooth at a and Y is smooth at $f(a)$.*

Then $f(X(\mathbb{Z}_p))$ contains the closed ball of radius p^{-k-1} around $f(a)$.

Proof. We first reduce the claim to the case where X is an affine space. Suppose that $X \subset \mathbb{A}^n$ is d-dimensional. By smoothness, it is given as the zero locus of $\varphi_1, \ldots, \varphi_{n-d} \in \mathbb{Z}_p[x_1, \ldots, x_n]$ such that the reductions modulo p of $\nabla \varphi_i(a)$ are linearly independent. Consider the map $F : \mathbb{A}^n \to Y \times \mathbb{A}^{n-d}$ given by $x \mapsto (f(x), p^k \varphi_1(x), \ldots, p^k \varphi_{n-d}(x))$. Then F satisfies the conditions of the lemma. If the claim holds for F, then it holds for f.

Next, we reduce the claim to the case where X and Y are affine spaces. Indeed, let e be the dimension of Y at $f(a)$. Item (1) of Lemma 4.1 allows us to assume that the coordinate projection $\pi : Y \to \mathbb{A}^e$ is one-to-one on $B(f(a), p^{-1})$. Item (2) of Lemma 4.1 implies that the function $\pi \circ f$ satisfies the conditions of the lemma, and the claim for $\pi \circ f$ implies the claim for f.

Finally, we prove the claim in the case $X = \mathbb{A}^n$ and $Y = \mathbb{A}^m$. We can assume that $a = 0$ and $f(a) = 0$. Since the coefficients of f are in $\mathbb{Z}_p$, we have that $df(a')(\mathbb{Z}_p^n) \supseteq p^k \mathbb{Z}_p^m$, for any $a' \in p\mathbb{Z}_p^n$. Let $b \in p^{k+1}\mathbb{Z}_p^m$. We will construct a sequence $a_\ell \in p\mathbb{Z}_p^n$ such that $\|f(a_\ell) - b\| < p^{-k-\ell}$. Taking a limit point of the a_ℓ, we get that $b \in f(\mathbb{Z}_p^n)$.

The sequence a_ℓ is defined by recursion starting from $a_0 = 0$. Given a_ℓ, the assumptions imply that there is $\epsilon \in p^{\ell+1}\mathbb{Z}_p^n$ such that $df(a_\ell)(\epsilon) = b - f(a_\ell)$. We have

$$\begin{aligned} \|f(a_\ell + \epsilon) - b\| &= \|f(a_\ell + \epsilon) - f(a_\ell) - df(a_\ell)(\epsilon) + df(a_\ell)(\epsilon) + f(a_\ell) - b\| \\ &= \|f(a_\ell + \epsilon) - f(a_\ell) - df(a_\ell)(\epsilon)\| \leqslant p^{-k}\|\epsilon\|^2 < p^{-k-\ell-1}, \end{aligned}$$

since the function $x \mapsto f(a_\ell + x) - f(a_\ell) - df(a_\ell)(x)$ is a polynomial without constant or linear term and its coefficients are divisible by p^k. □

Definition 4.3. For elements $g, h \in \mathrm{SL}_n$, let $\Phi_{g,h} : \mathrm{SL}_n \times \mathrm{SL}_n \to \mathrm{SL}_n$ be the map $\Phi_{g,h}^R(x,y) = g^x h^y$.

Lemma 4.4. *Let $n \geqslant 3$ and assume that $a, b \in \mathrm{SL}_n(\mathbb{F}_q)$ generate $\mathrm{SL}_n(\mathbb{F}_p)$ where $\mathbb{F}_q$ is a finite field of order q. Then the differential of $\Phi_{a,b}$ at $(1,1)$ is onto.*

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

Proof. After identifying $T_{ab}\,\mathrm{SL}_n = ab + ab\mathfrak{sl}_n$ and $\mathfrak{sl}_n$, the differential of $\Phi_{a,b}$ is $(X,Y) \mapsto (X - X^a)^b + (Y - Y^b)$. Let $\varphi \in \mathrm{M}_n(\mathbb{F}_q)^*$ and assume it vanishes on the image of $d\Phi_{a,b}$. Then there is $A \in \mathrm{M}_n(\mathbb{F}_q)$ such that $\varphi(X) = \mathrm{tr}(AX)$. For every $Y \in \mathfrak{sl}_n(\mathbb{F}_p)$, $\varphi(Y - Y^b) = 0$, so $\mathrm{tr}(Y \cdot (A^{b^{-1}} - A)) = \mathrm{tr}(A(Y - Y^b)) = 0$. Thus, $A^{b^{-1}} - A$ is a scalar. Similarly, using the assumption that $\varphi((X - X^a)^b) = 0$, we get that $(A^{b^{-1}})^{a^{-1}} - A^{b^{-1}}$ is a scalar. Using the fact that $A^{b^{-1}} - A$ is a scalar, we get that $A^{a^{-1}} - A$ is also a scalar. The set $X = \{g \in \mathrm{SL}_n(\mathbb{F}_q) \mid A^g - A \text{ is a scalar}\}$ is closed under multiplication. Since $a^{-1}, b^{-1} \in X$, we get $X = \mathrm{SL}_n(\mathbb{F}_q)$. Since $\mathrm{SL}_n(\mathbb{F}_p)$ is perfect and the function $g \mapsto A^g - A$ is a homomorphism between $\mathrm{SL}_n(\mathbb{F}_q)$ and $\mathbb{F}_q\,\mathrm{Id}$, we get that this homomorphism must be trivial. Hence, A commutes with $\mathrm{SL}_n(\mathbb{F}_q)$, so it must be scalar. It follows that the restriction of φ to $\mathfrak{sl}_n(\mathbb{F}_q)$ is zero. □

LEMMA 4.5. *For every non-trivial word w, there is n_0 such that, for any integer $n \geqslant n_0$, we have $w(\mathrm{SL}_n(\widehat{\mathbb{Z}}))^7 = \mathrm{SL}_n(\widehat{\mathbb{Z}})$.*

Proof. By [LS09], there is n_0 such that, if $n \geqslant n_0$ and p is any prime, then $w(\mathrm{SL}_n(\mathbb{F}_p))^2$ contains all non-scalar matrices. In particular, $w(\mathrm{SL}_n(\mathbb{F}_p))^3 = \mathrm{SL}_n(\mathbb{F}_p)$. Fix a prime p. Choosing generators $a, b \in \mathrm{SL}_n(\mathbb{F}_p)$ (which are not scalars), there are $g, h \in w(\mathrm{SL}_n(\mathbb{Z}_p))^2$ such that the reductions of g, h modulo p are a, b respectively. We get that $w(\mathrm{SL}_n(\mathbb{Z}_p))^4 \supset \Phi_{g,h}(\mathrm{SL}_n(\mathbb{Z}_p) \times \mathrm{SL}_n(\mathbb{Z}_p))$. It is well known that SL_n and thus also $\mathrm{SL}_n \times \mathrm{SL}_n$ are smooth at every point. Lemmas 4.4 and 4.2 imply that $w(\mathrm{SL}_n(\mathbb{Z}_p))^4$ contains the coset $gh\,\mathrm{SL}_n(\mathbb{Z}_p; p)$. Hence, $w(\mathrm{SL}_n(\mathbb{Z}_p))^7 = \mathrm{SL}_n(\mathbb{Z}_p)$.

Since $w(\mathrm{SL}_n(\widehat{\mathbb{Z}})) = \prod_p w(\mathrm{SL}_n(\mathbb{Z}_p))$, the claim follows. □

Proof of Theorem 1.1 (without an explicit bound). By Theorem 1.4 and Lemma 4.5. □

We move on to the proof of Theorem 1.2.

LEMMA 4.6. *For every $n \geqslant 2$ there is a constant C such that the following holds: if p is a prime and $A \in \mathfrak{sl}_n(\mathbb{F}_p)$ is non-central, then every element of $\mathfrak{sl}_n(\mathbb{F}_p)$ is equal to the sum of at most C elements of $\{x^{-1}Ax \mid x \in \mathrm{SL}_n(\mathbb{F}_p)\}$.*

Proof. It is well known that the only non-trivial $\mathrm{SL}_n(\mathbb{F}_p)$-invariant subspace of $\mathfrak{sl}_n(\mathbb{F}_p)$ is the subset consisting of scalar matrices. Hence, for every p, there is a constant C_p such that every element of $\mathfrak{sl}_n(\mathbb{F}_p)$ is equal to the sum of at most C_p elements of $\{x^{-1}Ax \mid x \in \mathrm{SL}_n(\mathbb{F}_p)\}$. Therefore, in order to find a uniform C, we can and will assume that p is large. In particular, we assume that $p \neq 2$.

For $1 \leqslant i \neq j \leqslant n$, let $E_{i,j}(a)$ be the matrix whose (i,j)th entry is a and with all other entries zero. Note that $E_{1,2}(a)$ is conjugate to $E_{1,2}(ab^2)$, for every $b \in \mathbb{F}_p$. Since every element in $\mathbb{F}_p$ is a sum of two squares, we get that, for any $a \in \mathbb{F}_p^\times$, any element of the form $E_{1,2}(b)$ is the sum of at most two conjugates of $E_{1,2}(a)$. In particular, there exists a one-dimensional linear subspace of $\mathfrak{sl}_n(\mathbb{F}_p)$ such that all its elements are sums of two conjugates of $E_{1,2}(a)$. Using the fact that the only $\mathrm{SL}_n(\mathbb{F}_p)$-invariant subspace of $\mathfrak{sl}_n(\mathbb{F}_p)$ is the subset consisting of scalar matrices once again, we see that if $a \neq 0$, then every matrix in $\mathfrak{sl}_n(\mathbb{F}_p)$ is the sum of at most $2(n^2 - 1)$ conjugates of $E_{1,2}(a)$. Therefore, it is enough to prove that there is a constant C such that, for some $a \in \mathbb{F}_p$, the matrix $E_{1,2}(a)$ is a sum of C conjugates of A. We divide the proof into several steps.

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

Step A. Assume that A is nilpotent. By using Jordan's normal form we see that A is conjugate to a block-diagonal matrix and each block-diagonal matrix has the from

$$\begin{pmatrix} 0 & a & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}, \tag{3}$$

where $0 \neq a \in \mathbb{F}_p$ (we cannot assume that $a = 1$ since we are conjugating with a matrix in $\mathrm{SL}_n(\mathbb{F}_p)$ and not $\mathrm{GL}_n(\mathbb{F}_p)$). A straightforward argument implies that it is enough to deal with the case where there is just one block. Clearly, we can assume that the dimension of this block is at least 3. Then there exists $\varepsilon \in \{-1, 1\}$ such that the diagonal matrix $\mathrm{diag}(\varepsilon, 1, -1, \ldots, (-1)^n)$ belongs to $\mathrm{SL}_n(\mathbb{F}_p)$. Denote $B := \mathrm{diag}(\varepsilon, 1, -1, \ldots, (-1)^n) + E_{2,n}(1)$. Then the nth coordinate of the first row of $A + B^{-1}AB$ is non-zero while all the other rows equal zero. Thus, $A + B^{-1}AB$ is conjugate to $E_{1,2}(a)$ for some non-zero a.

Step B. Assume that $n = 2$. If A is nilpotent, it is conjugate to $E_{1,2}(a)$, for some a, and the claim holds. In general, since A is not scalar, it is conjugate to $\left(\begin{smallmatrix} 0 & a \\ b & 0 \end{smallmatrix}\right)$. We claim that, if $p \geqslant 11$, there are $x, y, z \in \mathbb{F}_p^\times$ such that $x^2 + y^2 + z^2 = 0$ and $x^{-2} + y^{-2} + z^{-2} \neq 0$. If this claim holds, then

$$\begin{pmatrix} x & \\ & x^{-1} \end{pmatrix} A \begin{pmatrix} x^{-1} & \\ & x \end{pmatrix} + \begin{pmatrix} y & \\ & y^{-1} \end{pmatrix} A \begin{pmatrix} y^{-1} & \\ & y \end{pmatrix} + \begin{pmatrix} z & \\ & z^{-1} \end{pmatrix} A \begin{pmatrix} z^{-1} & \\ & z \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 \\ b(x^{-2} + y^{-2} + z^{-2}) & 0 \end{pmatrix},$$

which is nilpotent.

To prove the claim, let X be the projective curve defined by $x^2 + y^2 + z^2 = 0$. Then X has $p + 1$ points over $\mathbb{F}_p$, and at most six of them have a zero in some coordinate. At most four of the points of X satisfy the equation $x^{-2} + y^{-2} + z^{-2} = 0$ (because these points satisfy $1 = (y^2 + z^2)(y^{-2} + z^{-2})$). In particular, if $p \geqslant 11$, the claim is true.

Step C. Assume $n > 2$ and the claim is true for all numbers smaller than n. We consider the following cases.

Case C1. Assume that $\det A = 0$. By conjugating A we can assume that it is of the form

$$\begin{pmatrix} 0 & * \\ 0 & B \end{pmatrix}, \tag{4}$$

where $B \in \mathfrak{sl}_{n-1}(\mathbb{F}_p)$. If $B = 0$ then A is a nilpotent matrix and we are done by step 1. Otherwise, we can assume that $p > n - 1$ so B is a non-scalar matrix since its trace is equal to zero. Then by the induction hypothesis the sum of a bounded number of conjugates of A is a non-zero nilpotent matrix and we are done by step 1.

Case C2. Assume now that $\det A \neq 0$. By using the rational canonical normal form we see that there exist non-zero $a, b \in \mathbb{F}_p$ such that A is conjugate to a block-diagonal matrix and one of the blocks of A (for notational ease, assume it is the first) is of the form

$$\begin{pmatrix} 0 & a & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ b & * & \cdots & * & * \end{pmatrix}. \tag{5}$$

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

Denote $B := \mathrm{diag}(-1,-1,1,\dots,1)$. Then $A + B^{-1}AB$ is a non-zero singular matrix and we are done by case C1. □

Remark 4.7. The proof of Lemma 4.6 can be adapted to work over all finite fields of characteristic different than 2. For fields of characteristic 2, the argument of step B should be replaced.

LEMMA 4.8. *For any $n \geqslant 3$ there is C such that the following statements hold.*

(1) *For any p, if $X \subseteq \mathrm{SL}_n(\mathbb{Z}_p)$ is symmetric and invariant to conjugation, then $X^C = \langle X \rangle$.*

(2) *For any non-trivial word w, the width of w in $\mathrm{SL}_n(\widehat{\mathbb{Z}})$ is less than C.*

Proof. (1) Let $k = \min\{i \mid (\exists g \in X) g \text{ is not central modulo } p^{k+1}\}$. Clearly, $\langle X \rangle \subseteq Z(\mathrm{SL}_n(\mathbb{Z}_p)) \cdot \mathrm{SL}_n(\mathbb{Z}_p; p^k)$. We will show that there is C such that $\mathrm{SL}_n(\mathbb{Z}_p; p^k) \subseteq X^C$, and it will follow that $X^{C+|Z(\mathrm{SL}_n(\mathbb{Z}_p))|} = \langle X \rangle$, which proves the claim since $|Z(\mathrm{SL}_n(\mathbb{Z}_p))| \leqslant n$.

Case 1: $k = 0$. Let $\overline{X} \subseteq \mathrm{SL}_n(\mathbb{F}_p)$ be the reduction of X modulo p. By assumption, $\overline{X}$ is non-central, so [LS01, Corollary 1.9] implies that there is C_1, depending only on n, such that $\overline{X}^{C_1} = \mathrm{SL}_n(\mathbb{F}_p)$. Let $a, b \in X^{C_1}$ such that their reductions modulo p generate $\mathrm{SL}_n(\mathbb{F}_p)$. By Lemma 4.4, the differential of the map $\Phi_{\overline{a},\overline{b}}$ at $(1,1)$ is onto, which implies that $d\Phi_{a,b}(\mathfrak{sl}_n^2(\mathbb{Z}_p)) = \mathfrak{sl}_n(\mathbb{Z}_p)$, since $d\Phi_{a,b}$ is $\mathbb{Z}_p$-linear. By Lemma 4.2, $ab\,\mathrm{SL}_n(\mathbb{Z}_p; p) \subseteq \Phi_{a,b}(\mathrm{SL}_n(\mathbb{Z}_p) \times \mathrm{SL}_n(\mathbb{Z}_p)) \subseteq X^{2C_1}$. It follows that $X^{3C_1} = \mathrm{SL}_n(\mathbb{Z}_p)$.

Case 2: $k > 0$. Let $g \in X$ be such that g is not a scalar modulo p^{k+1}. Since g is a scalar modulo p^k, there exists $h \in \mathrm{SL}_n(\mathbb{Z}_p)$ such that $g^{|Z(\mathrm{SL}_n(\mathbb{Z}_p))|-1} h^{-1} g h \in X^{|Z(\mathrm{SL}_n(\mathbb{Z}_p))|}$ belongs to $\mathrm{SL}_n(\mathbb{Z}_p; p^k)$ and is not a scalar modulo p^{k+1}. Since $\mathrm{SL}_n(\mathbb{Z}_p; p^k)/\mathrm{SL}_n(\mathbb{Z}_p; p^{k+1}) = \mathfrak{sl}_n(\mathbb{F}_p)$ as $\mathrm{SL}_n(\mathbb{Z}_p)$-modules, Lemma 4.6 implies that there is a constant C, independent of X, such that $X^C \cdot \mathrm{SL}_n(\mathbb{Z}_p; p^{k+1}) \supseteq \mathrm{SL}_n(\mathbb{Z}_p; p^k)$. Let $\overline{a}$ be a maximal nilpotent Jordan block and let $\overline{b} = \overline{a}^T$. Note that the intersection of the centralizers of $\overline{a}, \overline{b}$ in M_n is the collection of scalar matrices. Choose $a, b \in X^C \cap \mathrm{SL}_n(\mathbb{Z}_p; p^k)$ whose images in $\mathrm{SL}_n(\mathbb{Z}_p; p^k)/\mathrm{SL}_n(\mathbb{Z}_p; p^{k+1}) = \mathfrak{sl}_n(\mathbb{F}_p)$ are $\overline{a}$ and $\overline{b}$. We will show that $\Phi_{a,b} : \mathrm{SL}_n \times \mathrm{SL}_n \to \mathrm{SL}_n$ satisfies the conditions of Lemma 4.2.

Since $a-1$ is divisible by p^k, we have $\mathrm{val}_p(x \mapsto x^{-1}ax - a) \geqslant k$. It follows that the derivative of this map also has p-valuation at least k. Similarly, $\Phi_{a,b}$ satisfies the first condition of Lemma 4.2.

Note that $d\Phi_{a,b}(\mathfrak{sl}(\mathbb{Z}_p)^2) \subset p^k \mathfrak{sl}(\mathbb{Z}_p)$. In order to show the reverse containment, it is enough to show that the composition of $d\Phi_{a,b}$ and the reduction map $p^k\mathfrak{sl}_n(\mathbb{Z}_p) \to p^k\mathfrak{sl}_n(\mathbb{Z}_p)/p^{k+1}\mathfrak{sl}_n(\mathbb{Z}_p)$ is onto. This composition is the map $(X, Y) \mapsto [\overline{X}, \overline{a}] + [\overline{Y}, \overline{b}]$ (where $[x, y]$ is the Lie bracket), so we need to show that there is no non-zero linear functional that vanishes on all elements of the form $[\overline{X}, \overline{a}]$ and $[\overline{X}, \overline{b}]$, for $\overline{X} \in \mathfrak{sl}_n(\mathbb{F}_p)$. Any such functional has the form $\mathrm{tr}(A\cdot)$ for some $A \in \mathfrak{sl}_n(\mathbb{F}_p)$. Since $\mathrm{tr}(A[B, C]) = \mathrm{tr}([A, B]C)$ for every three matrices A, B and C, the assumption that $\mathrm{tr}(A[\overline{a}, \overline{X}]) = 0$ for all $\overline{X} \in \mathfrak{sl}_n(\mathbb{F}_p)$ implies that $[A, \overline{a}] = \alpha I$, for some α. Similarly, $[A, \overline{b}] = \beta I$, for some β. From $[A, \overline{a}] = \alpha I$ we get (by induction) that $A_{i+1,i} = -i\alpha$, whereas from $[A, \overline{b}] = \beta I$ we get that $A_{i+1,i} = A_{i+2,i+1}$. Since $n \geqslant 3$, we get $\alpha = 0$. Similarly, $\beta = 0$. Consequently, A commutes with $\overline{a}$ and $\overline{b}$, so $A = 0$, a contradiction.

Applying Lemma 4.2 to $\Phi_{a,b}$, we get that any element in $ab\,\mathrm{SL}_n(\mathbb{Z}_p; p^{k+1})$ is in $\Phi_{a,b}(\mathrm{SL}_n(\mathbb{Z}_p)^2)$, so, in particular, $ab\,\mathrm{SL}_n(\mathbb{Z}_p; p^{k+1}) \subset X^{2C}$ and $\mathrm{SL}_n(\mathbb{Z}_p; p^{k+1}) \subset X^{4C}$. Since $X^C\,\mathrm{SL}_n(\mathbb{Z}_p; p^{k+1}) \supseteq \mathrm{SL}_n(\mathbb{Z}_p; p^k)$, we get that $X^{5C} \supseteq \mathrm{SL}_n(\mathbb{Z}_p; p^k)$, proving the claim in this case.

(2) Since $w(\mathrm{SL}_n(\widehat{\mathbb{Z}})) = \prod w(\mathrm{SL}_n(\mathbb{Z}_p))$, the claim follows from the first claim. □

Proof of Theorem 1.2. By Theorem 1.4 and Lemma 4.8. □

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

5. Proof of Theorems 1.1 and 1.4 with explicit bounds

The goal of this section is to prove the explicit bound of Theorem 1.1. The proof follows the arguments in [DV88].

Lemma 5.1. *Let $q, m \in \mathbb{N}^+$ and denote $n := 3m$. Assume that $g_1, \dots, g_m \in \mathrm{SL}_3(\mathbb{Z}; q)$ and $g_1 \cdots g_m = e$. Then $g := \mathrm{diag}(g_1, \dots, g_m) \in L_n(\mathbb{Z}; q)\tilde{U}_n(\mathbb{Z}; q)U_n(\mathbb{Z}; q)$ where $\tilde{U}_n(\mathbb{Z}; q) := \{hkh^{-1} \mid k \in U_n(\mathbb{Z}; q) \wedge h \in \mathrm{SL}_n(\mathbb{Z})\}$.*

Proof. Let I_3 be the identity matrix of $\mathrm{SL}_3(\mathbb{Z})$ and identify $M_n(\mathbb{Z})$ with $M_m(M_3(\mathbb{Z}))$, where $M_k(R)$ is the ring of $k \times k$ matrices over the ring R. Let l_1 be the matrix of $M_m(M_3(\mathbb{Z}))$ with I_3 on the diagonal, g_i^{-1} on the $(i+1, i)$th entry for every $1 \leqslant i \leqslant m-1$, and zero elsewhere. Let l_2 be the matrix of $M_m(M_3(\mathbb{Z}))$ with I_3 on the diagonal, I_3 on the $(i+1, i)$th entry for every $1 \leqslant i \leqslant m-1$, and zero elsewhere. Let u_1 be the matrix of $M_m(M_3(\mathbb{Z}))$ with I_3 on the diagonal, $1 - g_1 \cdots g_i$ on the $(i, i+1)$th entry for every $1 \leqslant i \leqslant m-1$, and zero elsewhere. Let u_2 be the matrix of $M_m(M_3(\mathbb{Z}))$ with I_3 on the diagonal, $(1 - g_1 \cdots g_i)g_{i+1}$ on the $(i, i+1)$th entry for every $1 \leqslant i \leqslant m-1$, and zero elsewhere. Then $g = l_1^{-1}u_1^{-1}l_2u_2 = (l_1^{-1}l_2)(l_2^{-1}u_1^{-1}l_2)u_2 \in L_n(\mathbb{Z}; q)\tilde{U}_n(\mathbb{Z}; q)U_n(\mathbb{Z}; q)$. □

Lemma 5.2. *Let $q, m \in \mathbb{N}^+$ and denote $n := 3m$. Assume that $g_1, \dots, g_m \in U_3(\mathbb{Z}; q)L_3(\mathbb{Z}; q)$. Denote $g := \mathrm{diag}(g_1 \cdots g_m, I_3, \dots, I_3) \in \mathrm{SL}_n(\mathbb{Z}; q)$ and $\tilde{U}_n(\mathbb{Z}; q) := \{hkh^{-1} \mid k \in U_n(\mathbb{Z}; q) \wedge h \in \mathrm{SL}_n(\mathbb{Z})\}$. Then $g \in L_n(\mathbb{Z}; q)\tilde{U}_n(\mathbb{Z}; q)U_n(\mathbb{Z}; q)L_n(\mathbb{Z}; q)$.*

Proof. Define $h := \mathrm{diag}(g_m, g_{m-1}, \dots, g_1) \in U_n(\mathbb{Z}; q)L_n(\mathbb{Z}; q)$. Lemma 5.1 implies that $gh^{-1} \in L_n(\mathbb{Z}; q)\tilde{U}_n(\mathbb{Z}; q)U_n(\mathbb{Z}; q)$. Thus,

$$g \in L_n(\mathbb{Z}; q)\tilde{U}_n(\mathbb{Z}; q)U_n(\mathbb{Z}; q)h \subseteq L_n(\mathbb{Z}; q)\tilde{U}_n(\mathbb{Z}; q)U_n(\mathbb{Z}; q)L_n(\mathbb{Z}; q). \qquad \square$$

Lemma 5.3. *Let $n \geqslant m \geqslant 3$ and $q \geqslant 1$. Denote $E^*(m, \mathbb{Z}; q) := \{\mathrm{diag}(1, \dots, 1, g) \in \mathrm{SL}_n(\mathbb{Z}) \mid g \in E(m, \mathbb{Z}; q)\}$. Then $E(n, \mathbb{Z}; q) = L_n(\mathbb{Z}; q)U_n(\mathbb{Z}; q)L_n(\mathbb{Z}; q)E^*(m, \mathbb{Z}; q)U_n(\mathbb{Z}; q)$.*

Proof. Let $q \geqslant 1$. The proof is by induction on n. The base case $n = m$ is clear. Assume that the statement is true for some $n \geqslant m$. We have to show that the statement is true also for $n+1$. Let $U_n^-(\mathbb{Z}; q)$ and $L_n^-(\mathbb{Z}; q)$ be the images in $\mathrm{SL}_{n+1}(\mathbb{Z})$ of $U_n(\mathbb{Z}; q)$ and $L_n(\mathbb{Z}; q)$ under the map $M \mapsto \mathrm{diag}(1, M)$. Denote $C_n^-(q) := \langle e_{j,1}(q) \mid 2 \leqslant j \leqslant n+1 \rangle$ and $R_n^-(q) := \langle e_{1,j}(q) \mid 2 \leqslant j \leqslant n+1 \rangle$. Finally, recall that the main theorem of [Men65] implies that for every $k \geqslant 3$,

$$E(k, \mathbb{Z}; q) = \{g \in \mathrm{SL}_k(\mathbb{Z}; q) \mid \forall 1 \leqslant i \leqslant k,\ g_{i,i} = 1 (\mathrm{mod}\ q^2)\}.$$

Let $g \in E(n+1, \mathbb{Z}; q)$. Then $\gcd(g_{1,1}, g_{2,1}, \dots, g_{n,1}) = 1$ and $\gcd(qg_{1,1}, g_{2,1}, \dots, g_{n,1}) = q$. Recall that $\mathbb{Z}$ satisfies the following stable range condition: if $m \geqslant 3$ and $a_1, \dots, a_m \in \mathbb{Z}$ then there exist $t_2, \dots, t_m \in \mathbb{Z}$ such that $\gcd(a_1, \dots, a_m) = \gcd(a_2 - t_2a_1, \dots, a_n - t_na_1)$. Thus, there exists $h \in C_n^-(\mathbb{Z}; q)g$ such that $\gcd(h_{2,1}, \dots, h_{n,1}) = q$. Since $h \in E(n, \mathbb{Z}; q)$, we have $h_{1,1} = 1$ modulo q^2 so there exists $h' \in R_n^-(\mathbb{Z}; q)h$ such that $h'_{1,1} = 1$. Finally, there exists $h'' \in C_n^-(q)h'R_n^-(q)$ such that $h'' = \mathrm{diag}(1, g')$ for some $g' \in \mathrm{SL}_n(\mathbb{Z}; q)$. Note that $g' \in E(n, \mathbb{Z}; q)$ since its diagonal entries are equal to 1 modulo q^2. Thus, the induction hypothesis implies that $h'' \in L_n^-(\mathbb{Z}; q)U_n^-(\mathbb{Z}; q)L_n^-(\mathbb{Z}; q)E^*(m, \mathbb{Z}; q)U_n^-(\mathbb{Z}; q)$. It follows that g belongs to

$$C_n^-(\mathbb{Z}; q)R_n^-(\mathbb{Z}; q)C_n^-(\mathbb{Z}; q)L_n^-(\mathbb{Z}; q)U_n^-(\mathbb{Z}; q)L_n^-(\mathbb{Z}; q)E^*(m, \mathbb{Z}; q)U_n^-(\mathbb{Z}; q)R_n^-(\mathbb{Z}; q).$$

Since both $U_n^-(\mathbb{Z}; q)$ and $L_n^-(\mathbb{Z}; q)$ normalize $C_n^-(\mathbb{Z}; q)$ and $R_n^-(\mathbb{Z}; q)$, we have

$$\begin{aligned} &C_n^-(\mathbb{Z}; q)R_n^-(\mathbb{Z}; q)C_n^-(\mathbb{Z}; q)L_n^-(\mathbb{Z}; q)U_n^-(\mathbb{Z}; q)L_n^-(\mathbb{Z}; q)E^*(m, \mathbb{Z}; q)U_n^-(\mathbb{Z}; q)R_n^-(\mathbb{Z}; q) \\ &= C_n^-(\mathbb{Z}; q)L_n^-(\mathbb{Z}; q)R_n^-(\mathbb{Z}; q)U_n^-(\mathbb{Z}; q)C_n^-(\mathbb{Z}; q)L_n^-(\mathbb{Z}; q)E^*(m, \mathbb{Z}; q)U_n^-(\mathbb{Z}; q)R_n^-(\mathbb{Z}; q) \\ &= L_{n+1}(\mathbb{Z}; q)U_{n+1}(\mathbb{Z}; q)L_{n+1}(\mathbb{Z}; q)E^*(m, \mathbb{Z}; q)U_{n+1}(\mathbb{Z}; q). \end{aligned}$$ □

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press

COROLLARY 5.4. *For every $n \geqslant 3$ denote $\tilde{U}_n(\mathbb{Z};q) := \{hkh^{-1} \mid k \in U_n(\mathbb{Z};q) \wedge h \in \mathrm{SL}_n(\mathbb{Z})\} = \{hkh^{-1} \mid k \in L_n(\mathbb{Z};q) \wedge h \in \mathrm{SL}_n(\mathbb{Z})\}$. There exists an integer N such that, for every $n \geqslant N$ and every $q \in \mathbb{Z}$,*

$$E(n,\mathbb{Z};q) \subseteq L_n(\mathbb{Z};q)(\tilde{U}_n(\mathbb{Z};q))^3 U_n(\mathbb{Z};q).$$

Proof. Proposition 2.9 implies that there exists a constant D such that for every $q \in \mathbb{Z}$, $(U_3(\mathbb{Z};q)L_3(\mathbb{Z};q))^D = \langle U_3(\mathbb{Z};q)L_3(\mathbb{Z};q)\rangle$. Denote $E^*(3,\mathbb{Z};q) := \{\mathrm{diag}(1,\dots,1,g) \in \mathrm{SL}_{3D}(\mathbb{Z}) \mid g \in E(3,\mathbb{Z};q)\}$. Lemma 5.2 shows that, for any q, $E^*(3,\mathbb{Z};q) \subseteq L_{3D}(\mathbb{Z};q)\tilde{U}_{3D}(\mathbb{Z};q)U_{3D}(\mathbb{Z};q)L_{3D}(\mathbb{Z};q)$. Lemma 5.3 implies that

$$E(n,\mathbb{Z};q) \subseteq L_n(\mathbb{Z};q)U_n(\mathbb{Z};q)L_n(\mathbb{Z};q)\tilde{U}_n(\mathbb{Z};q)U_n(\mathbb{Z};q)L_n(\mathbb{Z};q)U_n(\mathbb{Z};q).$$

Since $L_n(\mathbb{Z};q)U_n(\mathbb{Z};q)L_n(\mathbb{Z};q) \subset L_n(\mathbb{Z};q)\tilde{U}_n(\mathbb{Z};q)$ and $U_n(\mathbb{Z};q)L_n(\mathbb{Z};q)U_n(\mathbb{Z};q) \subset \tilde{U}_n(\mathbb{Z};q)U_n(\mathbb{Z};q)$, we get the result. □

Proof of Theorems 1.1 and 1.4 (with explicit bounds). Let $n \geqslant 3$. The proof of Theorem 1.4 shows that there is a non-zero $q \in \mathbb{Z}$ such that $U_n(\mathbb{Z};q)$ and $L_n(\mathbb{Z};q)$ are contained in $w(\mathrm{SL}_n(\mathbb{Z}))^{16}$. Since $w(\mathrm{SL}_n(\mathbb{Z}))$ is a normal subset, the set $\tilde{U}_n(\mathbb{Z};q) := \{hkh^{-1} \mid k \in U_n(\mathbb{Z};q) \wedge h \in \mathrm{SL}_n(\mathbb{Z})\}$ is contained in $w(\mathrm{SL}_n(\mathbb{Z}))^{16}$. Corollary 5.4 implies that if n is large enough then $E(n,\mathbb{Z};q) \subseteq w(\mathrm{SL}_n(\mathbb{Z}))^{16\cdot 5}$. Since $E(n,\mathbb{Z};q)$ contains a congruence subgroup, we have proved the bound in Theorem 1.4. Finally, Lemma 4.5 implies that if n is large enough then $\mathrm{SL}_n(\mathbb{Z}) \subseteq w(\mathrm{SL}_n(\mathbb{Z}))^{16\cdot 5+7}$. □

ACKNOWLEDGEMENTS

The authors thank Andrei Rapinchuk and Ariel Karelin for helpful conversations. They are also grateful to the anonymous referee for improving the bounds of Theorems 1.1 and 1.4 and for simplifying the proof of Lemma 4.6. N.A. was partially supported by NSF grant DMS-1303205 and BSF grant 2012247. C.M. was partially supported by BSF grant 2014099 and ISF grant 662/15.

REFERENCES

Bor83 A. Borel, *On free subgroups of semisimple groups*, Enseign. Math. (2) **29** (1983), 151–164.

DV88 R. K. Dennis and L. N. Vaserstein, *On a question of M. Newman on the number of commutators*, J. Algebra **118** (1988), 150–161.

LS09 M. Larsen and A. Shalev, *Word maps and Waring type problems*, J. Amer. Math. Soc. **22** (2009), 437–466.

LS01 M. Liebeck and A. Shalev, *Diameters of finite simple groups: sharp bounds and applications*, Ann. of Math. (2) **154** (2001), 383–406.

Lub14 A. Lubotzky, *Images of word maps in finite simple groups*, Glasg. Math. J. **56** (2014), 465–469.

Men65 J. L. Mennicke, *Finite factor groups of the unimodular group*, Ann. of Math. (2) **81** (1965), 31–37.

MN14 A. Myasnikov and A. Nikolaev, *Verbal subgroups of hyperbolic groups have infinite width*, J. Lond. Math. Soc. (2) **90** (2014), 573–591.

MRS18 A. V. Morgan, A. S. Rapinchuk and B. Sury, *Bounded generation of SL_2 over rings of S-integers with infinitely many units*, Algebra Number Theory **12** (2018), 1949–1974.

Seg09 D. Segal, *Words: notes on verbal width in groups*, London Mathematical Society Lecture Note Series, vol. 361 (Cambridge University Press, Cambridge, 2009).

Tho13 A. Thom, *Convergent sequences in discrete groups*, Canad. Math. Bull. **56** (2013), 424–433.

Tit76 J. Tits, *Systèmes générateurs de groupes de congruence*, C. R. Acad. Sci. Paris Ser. A-B **283** (1976), A693–A695.

Wit07 D. Witte Morris, *Bounded generation of* $\mathrm{SL}(n,A)$ *(after D. Carter, G. Keller and E. Paige)*, New York J. Math. **13** (2007), 383–421.

Nir Avni avni.nir@gmail.com
Department of Mathematics, Northwestern University, Evanston IL, USA

Chen Meiri chenm@technion.ac.il
Department of Mathematics, Technion, Haifa, Israel

https://doi.org/10.1112/S0010437X19007334 Published online by Cambridge University Press