Groups Geom. Dyn. 19 (2025), 899–912
DOI 10.4171/GGD/800

A height gap in $\mathrm{GL}_d(\bar{\mathbb{Q}})$ and almost laws

Lvzhou Chen, Sebastian Hurtado, and Homin Lee

Abstract. E. Breuillard showed that finite subsets F of matrices in $\mathrm{GL}_d(\bar{\mathbb{Q}})$ generating non-virtually solvable groups have normalized height $\hat{h}(F) \geq \varepsilon_d$, for some positive $\varepsilon_d > 0$. The normalized height $\hat{h}(F)$ is a measure of the arithmetic size of F and this result can be thought of as a non-abelian analog of Lehmer's Mahler measure problem. We give a new shorter proof of this result. Our key idea relies on the existence of particular word maps in compact Lie groups (known as almost laws) whose image lies close to the identity element.

1. Introduction

In [6], E. Breuillard proved what can be considered a non-abelian version of Lehmer's problem about the Mahler measure of an algebraic number. He showed that if a finite set $F \subset \mathrm{GL}_d(\bar{\mathbb{Q}})$ of invertible matrices with algebraic entries generates a group that is not virtually solvable, then its arithmetic height (a measure of its arithmetic complexity) is bounded below by an absolute constant $\varepsilon_d > 0$ independent of F. Some consequences of Breuillard's theorem include the existence of a lower bound for the exponential growth of a non-virtually solvable group of $\mathrm{GL}_d(\mathbb{C})$, a strong version of the classical Tits alternative about the existence of free subgroups in linear groups, lower bounds in the girth of finite groups of Lie type, and very recently some results about the geometry of arithmetic locally symmetric spaces [11]. See [7] for a discussion of some of these applications.

In this article, we will provide a more elementary proof of Breuillard's theorem. Our proof makes use of some of the results about products of matrices proved by Breuillard, for example, Lemma 2.3, Propositions 3.5 and 3.6, but avoids results in diophantine geometry such as Bilu's equidistribution theorem [3], or results of Zhang [16] about small points on algebraic tori, and also avoids the use of the geometry of Bruhat–Tits buildings. The key idea in our proof relies on the existence of *word maps* in compact Lie groups whose image lies close to the identity element. These words are known as *almost laws* after Andreas Thom [14]; see Section 2.2.

1.1. Definitions

Let $K \subset \bar{\mathbb{Q}}$ be a number field and let V_K be the set of absolute values on K up to equivalence, which are either archimedean (corresponding to real and complex embeddings of K

Mathematics Subject Classification 2020: 20G25 (primary); 20G30, 20E07 (secondary).
Keywords: height gap, almost law, word maps, arithmetic groups, and linear groups.

in $\mathbb{R}$ or $\mathbb{C}$) or non-archimedean (corresponding to prime ideals in the ring of integers $\mathcal{O}_K$ of K). For $v \in V_K$, let K_v be the corresponding completion of K and define $n_v := [K_v : \mathbb{Q}_p]$ if $v|p$ or $n_v := [K_v : \mathbb{R}]$ if $\mathbb{R} \subseteq K_v$. For a vector $x = (x_1, x_2, \ldots, x_d)$ in K_v^d, we define $\|x\|_v := \max_{i=1}^d |x_i|_v$ if v is non-archimedean and $\|x\|_v := \sqrt{|x_1|_v^2 + \cdots + |x_d|_v^2}$ if ν is archimedean.

For a matrix A in $\mathrm{GL}_d(K)$, let $\|A\|_v$ be the operator norm. For a finite set F of matrices in $\mathrm{GL}_d(K)$, let $\|F\|_v := \max_{A \in F} \|A\|_v$. The *height* of F is defined as

$$h(F) := \frac{1}{[K : \mathbb{Q}]} \sum_{v \in V_K} n_v \log^+ \|F\|_v,$$

where $\log^+(x) = \log \max(|x|, 1)$ for a real number x. It does not depend on K as long as F is a subset of $\mathrm{GL}_d(K)$. The *normalized height* is defined by

$$\hat{h}(F) := \lim_{n \to \infty} \frac{h(F^n)}{n},$$

where F^n consists of all products of n elements in F. The limit above exists because $h(F^n)$ is sub-additive, which also implies $h(F) \geq \hat{h}(F)$.

Our main result is a new proof of the following theorem due to Breuillard [6].

Theorem 1.1. *For any positive integer d, there exists $\varepsilon_d > 0$ such that for any symmetric finite set $F \subset \mathrm{GL}_d(\overline{\mathbb{Q}})$, either*

(1) *F generates a virtually solvable group, or*

(2) *$\hat{h}(F) \geq \varepsilon_d$.*

Remark 1.2. In Section 4 we give a brief discussion on the possible estimates one can get of ε_d by using our methods and ideas of E. Breuillard and A. Thom.

Remark 1.3. In [6], Breuillard proved this theorem without the assumption that F is symmetric[1]. Moreover he showed that F can be conjugated in such a way that their height and normalized height are comparable by a constant only depending on d. Our proof does not establish these results. Moreover we use the results of Eskin–Mozes–Oh [10] and some lemmas of E. Breuillard [6] that allow us to compare the height and normalized height.

1.2. Idea of the proof of Theorem 1.1

We illustrate the idea of the proof in a very particular case. Assume $d = 2$, $F \subset \mathrm{SL}_2(\overline{\mathbb{Q}})$ and assume that all the entries of elements of F are algebraic integers. In this case, the only contribution to the height comes from archimedean places. Let K be the number field generated by the entries of elements in F. Let $w \in \mathbb{F}_2$ be an almost law for $\mathrm{SU}(2)$ (see Section 2.2 for the definition). Suppose for simplicity there exist $A, B \in F$ such that

[1]Symmetric means that if $x \in F$, then $x^{-1} \in F$.

$\mathrm{Tr}(w(A,B)-Id) \neq 0$ (see Lemma 3.1), and let

$$Q = \prod_{i=1}^{k} \big(\mathrm{Tr}\big(w(A_i,B_i)-Id\big)\big), \tag{1}$$

where $k=[K:\mathbb{Q}]$, $A_i=\sigma_i(A)$ and $B_i=\sigma_i(B)$, and $\sigma_i : K \to \mathbb{C}$ are the different embeddings of K in $\mathbb{C}$.

By construction, Q is a nonzero integer and therefore $|Q| \geq 1$. Let $\delta > 0$ be small enough and let S be the set of i's such that either $\|A_i\|$ or $\|B_i\|$ is greater than e^{δ}. There are two cases: If $|S| \geq k/2$, then one of the heights $h(A)$ or $h(B)$ must be greater than $\frac{1}{2}\delta$ and so $h(F) \geq \frac{1}{2}\delta$ and we are done. If $S \leq k/2$ we have for $i \notin S$, A_i and B_i are close to a pair of elements of SU(2) and then for such i's, $\mathrm{Tr}(w(A_i,B_i)-Id)$ is smaller than say e^{-10} by choosing the almost law appropriately and δ small enough.

Therefore from $|Q| \geq 1$, we obtain

$$\Big|\prod_{i\in S}\mathrm{Tr}\big(w(A_i,B_i)-Id\big)\Big| \geq \Big|\prod_{i\notin S}\mathrm{Tr}\big(w(A_i,B_i)-Id\big)\Big|^{-1} \geq e^{10k/2}$$

and this implies that the height of $w(A,B)$ must be bounded below by a constant $c>0$ independent of k. Thus $h(F) \geq \frac{1}{|w|}h(F^{|w|}) \geq \frac{1}{|w|}h(w(A,B)) \geq \frac{c}{|w|}$, where $|w|$ is the length of w.

1.3. Organization of the article

In Section 2, we recall definitions and facts about heights, normalized heights, and almost laws. In Section 3, we prove Theorem 1.1. In Section 4 we discuss strategies to obtain explicit estimates about the constant ε_d in terms of d.

2. Background

In this section, we give detailed definitions and basic properties of the crucial notions in the statements and proofs of our results: *heights* and *almost laws*.

2.1. Heights

For a number field $K \leq \bar{\mathbb{Q}}$ and an absolute value $|\cdot|_v$ on K, define $n_v = [K_v : \mathbb{Q}_p]$ as the degree of the completion K_v of K over the closure $\mathbb{Q}_p$ of $\mathbb{Q}$ in K_v. In the case where v is archimedean $n_v = 1, 2$ depending on whether v comes from a real or complex embedding. When v is non-archimedean, we normalize the absolute value $|\cdot|_v$ on K_v so that its restriction to $\mathbb{Q}_p$ is the standard absolute value; i.e., $|p|_v = 1/p$. With such normalization, we have the following product formula.

Theorem 2.1 (Product formula). *Let K be a number field and V_K the set of equivalence classes of absolute values on K. Then, for every $x \in K$,*

$$\prod_{v\in V_K} |x|_v^{n_v} = 1.$$

Let k be a local field of characteristic 0. Let $\|\cdot\|_k$ be the standard norm on k^d as in the introduction. We use the same notation for the operator norm on the space $M_d(k)$ of d-by-d matrices with entries in k. We define below the quantities $\Lambda_k(F)$, $E_k(F)$ and $R_k(F)$ for a bounded set $F \subset M_d(k)$, which were defined in [6] but were also defined previously by other authors (see Rota–Strang [13], or Breuillard [8] and its references for a more detailed discussion).

Definition 2.2. Let F be a bounded subset of matrices in $M_d(k)$. We set

(1) the *norm* of F as

$$\|F\|_k = \sup_{g \in F} \|g\|_k,$$

(2) the *minimal norm* of F as

$$E_k(F) = \inf_{x \in \mathrm{GL}_d(\bar{k})} \|xFx^{-1}\|_k,$$

(3) the *maximal eigenvalue* of F as

$$\Lambda_k(F) = \max\left\{|\lambda|_k, \lambda \in \mathrm{spec}(q),\ q \in F\right\},$$

and

(4) the *spectral radius* of F as

$$R_k(F) = \lim_{n \to \infty} \|F^n\|_k^{1/n}.$$

For simplicity, when $k = \mathbb{C}$, we drop k in the notation and denote the above quantities using $\|\cdot\|$, E, Λ, and R, respectively. If k is understood, sometimes we will use the subscript v for $v \in V_k$, such as Λ_v, E_v, and R_v, instead of writing subscript k_v.

With the above definition, we can reformulate the normalized height for a finite set $F \subset M_d(k)$ as follows (see [6, Section 2.2]):

$$\hat{h}(F) = \frac{1}{[k:Q]} \sum_{v \in V_k} n_v \log^+ R_v(F).$$

We will often use two basic properties that follow directly from the definition:

(1) $\hat{h}(F^n) = n\hat{h}(F)$, and

(2) $\hat{h}(F') \leq \hat{h}(F)$ for any subset $F' \subset F$, especially when F' consists of a single matrix $A \in F$, in which case we simply write $\hat{h}(A)$ instead of $\hat{h}(\{A\})$.

We will also use Lemma 2.4 below in our estimates, which is based on the following lemma of Breuillard.

Lemma 2.3 ([6, Proposition 2.7]). $\Lambda_k(F) \leq R_k(F)$.

Lemma 2.4. *Let $|\cdot|_v$ be a non-archimedean place on a field K and equip $M_d(K)$ with the operator norm $\|\cdot\|_v$ induced by the norm $\|x\|_v := \max_i |x_i|_v$ on K^d. Then for any $A \in M_d(K)$ we have*

$$|\mathrm{Tr} A|_v \leq R_v(A).$$

Proof. Let L be a finite extension of K, such that all the eigenvalues $\lambda_1, \dots, \lambda_d$ of A belong to L. There is a unique completion $|\cdot|_w$ extending $|\cdot|_v$. Therefore $\mathrm{Tr}(A) = \lambda_1 + \cdots + \lambda_d$ and we have $|\mathrm{Tr}(A)|_v \le \max_i |\lambda_i|_w = \Lambda_w(A)$. By Lemma 2.3, we have that $\Lambda_w(A) \le R_w(A)$ and it is easily checked that $R_v(A) = R_w(A)$. ■

2.2. Almost laws

Given an element w in a free group $\mathbb{F}_n = \langle x_1, \dots, x_n \rangle$ of rank n, there is a natural *word map* associated with any group G

$$w\colon \prod_n G \to G,$$

defined as follows: Express w as a reduced word with alphabet $\{x_1^{\pm 1}, \dots, x_n^{\pm 1}\}$, and then for any $(g_1, \dots, g_n) \in \prod_n G$, substitute each x_i by g_i and x_i^{-1} by g_i^{-1}. For example, if $w = [x_1, x_2] = x_1 x_2 x_1^{-1} x_2^{-1} \in \mathbb{F}_2$, then $w(A, B) = [A, B]$.

Definition 2.5. For a group G, a *law* is a nontrivial element $w \in \mathbb{F}_n$ such that the image of w_G is the identity element 1_G. Given $\varepsilon > 0$ and a metric d on G, an *ε-almost law* is a nontrivial element $w \in \mathbb{F}_n$ such that the image of G lies in an ε-neighborhood of 1_G.

For instance, the word $w = [x_1, x_2] \in \mathbb{F}_2$ is a law in a group G if and only if G is abelian. In general, groups obeying a law have rather special properties; for instance, they contain no nonabelian free subgroups, and the stable commutator length vanishes [9]. Our proof of Theorem 1.1 relies on the following result due to A. Thom, which has also been attributed to E. Lindestrauss; see [1].

Theorem 2.6 (A. Thom [14]). *Let G be a compact Lie group and d_G a bi-invariant metric in G. For every $\varepsilon > 0$, there exists an ε-almost law $w_\varepsilon \in \mathbb{F}_2$ on G; that is, for all $A, B \in G$ we have*

$$d_G\big(w_\varepsilon(A, B), 1_G\big) < \varepsilon.$$

Remark 2.7. To calculate explicitly the constant ε_d in the main theorem, one needs to compute $|w_{\varepsilon/2}|$.

In $\mathrm{GL}_d(\mathbb{C})$, an almost law on the compact subgroup $\mathrm{U}(d)$ extends to a neighborhood. Below we denote the identity matrix as I_d, and let $d_{\mathrm{GL}_d(\mathbb{C})}(X, Y) := \|X - Y\|$, where $\|A\|$ is the operator norm with respect to the Hermitian metric for any $A \in M_d(\mathbb{C})$. Note that the restriction of $d_{\mathrm{GL}_d(\mathbb{C})}$ to $\mathrm{U}(d)$ is bi-invariant.

Corollary 2.8. *For every $d > 0$ and $\varepsilon \in (0, 1)$, there exists a nontrivial element $w \in \mathbb{F}_2$ and $\delta = \delta(d, \varepsilon) > 0$ such that for any $A, B \in \mathrm{GL}_d(\mathbb{C})$ satisfying $\|A\|, \|A^{-1}\|, \|B\|, \|B^{-1}\| < e^\delta$ we have*

$$d_{\mathrm{GL}_d(\mathbb{C})}\big(w(A, B), I_d\big) < \varepsilon.$$

Moreover, w can be taken as any $\frac{\varepsilon}{2}$-almost law on $\mathrm{U}(d)$, and $\delta = \frac{\varepsilon}{8|w|}$.

Proof. Let $w = w_{\varepsilon/2}$ be an $\frac{\varepsilon}{2}$-almost law as in Theorem 2.6 for $\mathrm{U}(d)$ with respect to the restricted metric. Then for any $A, B \in \mathrm{GL}_d(\mathbb{C})$, by the singular value decomposition we have $A = P_A \Lambda_A Q_A$ and $B = P_B \Lambda_B Q_B$ where $P_A, Q_A, P_B, Q_B \in \mathrm{U}(d)$ and Λ_A, Λ_B are positive diagonal matrices. If $\|A\|, \|A^{-1}\|, \|B\|, \|B^{-1}\| < e^\delta$ for some $\delta > 0$, then all diagonal entries of Λ_A, Λ_B lie in $(e^{-\delta}, e^\delta)$.

Let $A' = P_A Q_A$ and $B' = P_B Q_B$, which lie in $\mathrm{U}(d)$. Note that $\|A - A'\| = \|\Lambda_A - I_d\| \le e^\delta - 1$ and $\|A^{-1} - A'^{-1}\| = \|\Lambda_A^{-1} - I_d\| \le e^\delta - 1$, and similarly for B. By inserting intermediate words replacing one $A^{\pm 1}$ (resp. $B^{\pm 1}$) by $A'^{\pm 1}$ (resp. $B'^{\pm 1}$) at a time in the word $w_{\varepsilon/2}(A, B)$, we obtain from the triangle inequality that

$$\left\| w_{\varepsilon/2}(A, B) - w_{\varepsilon/2}(A', B') \right\| \le |w_{\varepsilon/2}| \cdot e^{\delta(|w_{\varepsilon/2}|-1)}(e^\delta - 1).$$

For $\delta = \frac{\varepsilon}{8|w_{\varepsilon/2}|} < 1$, we have $e^{\delta(|w_{\varepsilon/2}|-1)} < e^{\varepsilon/8} < 2$ and $e^\delta - 1 < 2\delta = \frac{\varepsilon}{4|w_{\varepsilon/2}|}$. Hence

$$\left\| w_{\varepsilon/2}(A, B) - w_{\varepsilon/2}(A', B') \right\| \le |w_{\varepsilon/2}| e^{\delta(|w_{\varepsilon/2}|-1)}(e^\delta - 1) \le \frac{\varepsilon}{4} \cdot e^{\frac{\varepsilon}{8}} \le \frac{\varepsilon}{2},$$

and

$$\left\| w_{\varepsilon/2}(A, B) - I_d \right\| \le \left\| w_{\varepsilon/2}(A, B) - w_{\varepsilon/2}(A', B') \right\| + \left\| w_{\varepsilon/2}(A', B') - I_d \right\| \le \varepsilon$$

using the almost law. ∎

Corollary 2.9. *For every $d > 0$ and $\varepsilon \in (0, 1)$, there exists a nontrivial element $w \in \mathbb{F}_2$ and $\delta = \delta(d, \varepsilon) > 0$ such that for any $A, B \in \mathrm{GL}_d(\mathbb{C})$ satisfying $\|A\|, \|A^{-1}\|, \|B\|, \|B^{-1}\| < e^\delta$ we have*

$$\mathrm{Tr}\big(w(A, B) - I_d\big) < \varepsilon.$$

Moreover, w can be taken as any $\frac{\varepsilon}{2d}$-almost law on $\mathrm{U}(d)$, and $\delta = \frac{\varepsilon}{8|w|}$.

Proof. This easily follows from Corollary 2.8 applied to ε/d. ∎

3. Proof of Theorem 1.1

We need to prove the uniform lower bound $\hat{h}(F) \ge \varepsilon_d$ whenever $F \subset \mathrm{GL}_d(\overline{\mathbb{Q}})$ generates a group that is not virtually solvable.

To make use of this assumption on F, we need the following lemma, which relies on [10, Proposition 3.2] or [6, Lemma 4.2].

Lemma 3.1. *Given any non-trivial word $w \in \mathbb{F}_2$, there is some n_w (only depending on w and d) such that for any set $F \subset \mathrm{GL}_d(\mathbb{C})$ containing the identity and generating a non-virtually solvable group, there exists $A, B \in F^{n_w}$ with $\mathrm{Tr}(w(A, B) - I_d) \ne 0$.*

Remark 3.2. The assumption that F contains the identity might be unnecessary, but the results of Eskin–Mozes–Oh [10] are written using balls instead of spheres. For our application, the set F is symmetric so we can replace F by F^2, if necessary. The same applies for Lemma 3.3 below.

Proof. Let Γ be the group generated by F in $\mathrm{GL}_d(\mathbb{C})$ and H the Zariski closure of Γ. We assumed that Γ is not virtually solvable, so H is not virtually solvable. We first show that there exists $X, Y \in \Gamma$ such that $\mathrm{Tr}(w(X,Y) - I_d) \neq 0$.

As H is an algebraic group defined over $\mathbb{C}$, we have the Levi decomposition $H = L \ltimes U$, where U is the unipotent radical of H and L is a reductive Levi subgroup of H. Furthermore, we have a semisimple algebraic group $S = [L, L]$ so that L can be written as an almost direct product of S and the central torus T; i.e., $L = S \cdot T$ and $|S \cap T| < \infty$. Let S^0 be the connected component of identity in S.

Assume that $\mathrm{Tr}(w(X,Y) - I_d) = 0$ for all $X, Y \in \Gamma$. We claim that S^0 is trivial. Indeed, by Borel's theorem [5], the restriction of the word map to S^0, $w\colon S^0 \times S^0 \to S^0$ is dominant; i.e., the image is Zariski dense in S^0. As $R = \{X \in H : \mathrm{Tr}(X) = d\}$ is a Zariski closed subset of H, the image of the word map w and hence S^0 are contained in R. It is easy to see that if a matrix X satisfies $\mathrm{Tr}(X^m) = d$ for all m, then X is a unipotent element. This implies that every element in S^0 is unipotent so it must be trivial as S^0 is semisimple. Therefore, H is virtually solvable, as virtually it is an extension of the solvable group U by the abelian group T. This contradicts our assumption. Hence there are $X, Y \in \Gamma$ such that $\mathrm{Tr}(w(X,Y) - I_d) \neq 0$.

For the rest of proof, we think of $\mathrm{GL}_d(\mathbb{C}) \times \mathrm{GL}_d(\mathbb{C})$ as a subgroup in $\mathrm{GL}_{2d}(\mathbb{C})$ diagonally. Let H' be the Zariski closure of $\Gamma \times \Gamma$ in $\mathrm{GL}_{2d}(\mathbb{C})$. Then the subset

$$\mathcal{X} = \big\{(A,B) \in \mathrm{GL}_d(\mathbb{C}) \times \mathrm{GL}_d(\mathbb{C}) : \mathrm{Tr}\big(w(A,B) - I_d\big) = 0\big\} \cap H' \subset \mathrm{GL}_{2d}(\mathbb{C})$$

is Zariski closed in H'. As we saw above, $\mathcal{X}$ is a proper Zariski closed subset in H'. Using the fact that $\Gamma \times \Gamma$ is generated by $F \times F$, [10, Proposition 3.2] or [6, Lemma 4.2] says that there exists $n_w \geq 1$ only depending on d and w such that we have

$$\big\{(A,B) \in \Gamma \times \Gamma : A, B \in F^n \subset \Gamma\big\} \not\subseteq \mathcal{X}$$

as desired. ∎

Let K be a number field so that $F \subset \mathrm{GL}_d(K)$. Let $k = [K : \mathbb{Q}]$ and let $\sigma_1, \sigma_2, \ldots, \sigma_k$ be all the embeddings of K in $\mathbb{C}$. Fix any $\varepsilon \in (0,1)$, choose an $\frac{\varepsilon}{2}$-almost law $w = w_{\varepsilon/2}$ on $\mathrm{U}(d)$, and let $\delta = \frac{\varepsilon}{8|w|}$ as in Corollary 2.9. For $w = w(x,y) \in \mathbb{F}_2 = \langle x, y\rangle$, let $w' = w([x,y],[x^{-1},y]) \in \mathbb{F}_2$. Note that w' is still nontrivial since $[x,y]$ and $[x^{-1},y]$ generate a free subgroup of $\mathbb{F}_2$. Denote the word length of w by $|w|$.

For technical reasons later, we need the following improvement of Lemma 3.1 to ensure that $A, B \in \mathrm{SL}_d(\mathbb{C})$.

Lemma 3.3. *In the above setting, there is $n = n_w$ only depending on w and d such that for any finite set $F \subset \mathrm{GL}_d(\mathbb{C})$ containing the identity and generating a subgroup that is not virtually solvable, there exists $A, B \in F^{n_w} \cap \mathrm{SL}_d(\mathbb{C})$ with $\mathrm{Tr}(w(A,B) - I_d) \neq 0$.*

Proof. Applying Lemma 3.1 to w' we described above, we obtain some n' (relying only on w and d) and $A_1, B_1 \in F^{n'}$ such that $\mathrm{Tr}(w'(A_1,B_1) - I_d) \neq 0$. Let $n = n_w := 4n'$,

$A = [A_1, B_1]$ and $B = [A_1^{-1}, B_1]$. Then $A, B \in F^n \cap \mathrm{SL}_d(\mathbb{C})$, and $w(A,B) = w'(A_1, B_1)$ by definition. Hence $\mathrm{Tr}(w(A,B) - I_d) \neq 0$. ∎

Throughout this section, let $n = n_w$ and $A, B \in F^{n_w} \cap \mathrm{SL}_d(\mathbb{C})$ be as in Lemma 3.3. Consider the following quantity:

$$Q := \prod_{i=1}^{k} \left(\mathrm{Tr}\big(w(A_i, B_i) - I_d\big)\right), \tag{2}$$

where $A_i = \sigma_i(A)$ and $B_i = \sigma_i(B)$.

This is a nonzero rational number since $\mathrm{Tr}(w(A,B) - \mathrm{Id}) \neq 0$. Let

$$\varepsilon_2 = \frac{\log \frac{1}{\varepsilon}}{\log \frac{2d}{\varepsilon} + \frac{|w|\delta^2}{16d \log \frac{1}{c}}} \quad \text{and} \quad \varepsilon_1 = \frac{1}{2}\left(\log \frac{1}{\varepsilon} - \varepsilon_2 \log \frac{2d}{\varepsilon}\right),$$

where $\delta = \frac{\varepsilon}{8|w|}$ as in Corollary 2.9 and $0 < c < 1$ is the constant (only depending on d) from Proposition 3.6, which can be chosen as $c = \frac{1}{2d}$. Note that $\varepsilon_2 < \log \frac{1}{\varepsilon} / \log \frac{2d}{\varepsilon} < 1$ and thus $\varepsilon_1 > 0$. Let

$$\varepsilon_d := \frac{\varepsilon_2 \delta^2}{32nd \log \frac{1}{c}} = \frac{\delta^2 \log \frac{1}{\varepsilon}}{32nd \log \frac{2d}{\varepsilon} \log \frac{1}{c} + 2n|w|\delta^2}, \tag{3}$$

which is at least at the order of $\frac{1}{n(d^2|w|^2 + |w|)}$ as $d \to \infty$ by setting $\varepsilon \in (0,1)$ independent of d. The constants ε_1 and ε_2 are chosen so that

$$\varepsilon_d = \frac{\varepsilon_2 \delta^2}{32nd \log \frac{1}{c}} = \frac{-\varepsilon_1 + (1 - \varepsilon_2) \log \frac{1}{\varepsilon} - \varepsilon_2 \log 2d}{n|w|} = \frac{\varepsilon_1}{n|w|},$$

which are the lower bounds of the normalized height in the analysis below.

We consider the two possibilities in the following two subsections.

3.1. Case 1: $|Q| \geq e^{-\varepsilon_1 k}$

In this case, we will use the estimate

$$\hat{h}(F) = \frac{1}{k} \sum_{v \in V_K} n_v \log^+ R_v(F) \geq \frac{1}{k} \sum_{i=1}^{k} \log^+ R_i(F), \tag{4}$$

where $R_i(F) = R(\sigma_i(F))$ and we simply ignore the non-archimedean places.

Partition the index set $I := \{1, 2, \ldots, k\}$ as $I = I_S \sqcup I_M \sqcup I_L$, where

$$I_S := \left\{i : \left|\mathrm{Tr}\big(w(A_i, B_i) - I_d\big)\right| < \varepsilon\right\}, \quad I_L := \left\{i : \left|\mathrm{Tr}\big(w(A_i, B_i) - I_d\big)\right| > 2d\right\},$$

and $I_M = I \setminus (I_S \sqcup I_L)$. Note that $I_S \cap I_L = \emptyset$ since $\varepsilon < 1$.

We further consider two subcases depending on the size of I_S.

3.1.1. Case 1a: $|I_S| > (1-\varepsilon_2)k$. In this case, we have

$$\prod_{i\notin I_S} \left|\mathrm{Tr}\big(w(A_i,B_i)-I_d\big)\right| = \frac{|Q|}{\left|\prod_{i\in I_S} \mathrm{Tr}\big(w(A_i,B_i)-I_d\big)\right|} \geq e^{-\varepsilon_1 k}\cdot \varepsilon^{-|I_S|}.$$

Taking log on both sides, we have

$$\sum_{i\notin I_S} \log\left|\mathrm{Tr}\big(w(A_i,B_i)-I_d\big)\right| \geq -\varepsilon_1 k + |I_S|\log\frac{1}{\varepsilon} \geq -\varepsilon_1 k + (1-\varepsilon_2)k\log\frac{1}{\varepsilon}.$$

As $|\mathrm{Tr}(w(A_i,B_i)-I_d)| \leq 2d$ for all $i\in I_M$, it follows that

$$\begin{aligned}\sum_{i\in I_L} \log\left|\mathrm{Tr}\big(w(A_i,B_i)-I_d\big)\right| &\geq \sum_{i\notin I_S} \log\left|\mathrm{Tr}\big(w(A_i,B_i)-I_d\big)\right| - |I_M|\log 2d\\ &\geq -\varepsilon_1 k + (1-\varepsilon_2)k\log\frac{1}{\varepsilon} - |I_M|\log 2d. \qquad (5)\end{aligned}$$

Lemma 3.4. *For any $X\in \mathrm{GL}_d(\mathbb{C})$, if $|\mathrm{Tr}(X-I_d)|>d$, then the spectral radius $\Lambda(X)\geq \frac{1}{d}|\mathrm{Tr}(X-I_d)|-1$. Moreover, if $|\mathrm{Tr}(X-I_d)|>2d$, then $\Lambda(X)\geq \frac{1}{2d}|\mathrm{Tr}(X-I_d)|$.*

Proof. Let $r=\frac{1}{d}|\mathrm{Tr}(X-I_d)|$ and let $\lambda_1,\dots,\lambda_d\in\mathbb{C}$ be the eigenvalues of X. Then for the average $\bar{\lambda}=\frac{1}{d}\sum_{j=1}^d \lambda_j$ we have $r=|\bar{\lambda}-1|$.

It follows that some λ_i lies outside the open disk D of radius r around $1\in\mathbb{C}$ since otherwise their average $\bar{\lambda}$ lies in D by convexity, contradicting $r=|\bar{\lambda}-1|$.

Since $r>1$ by our assumption, the disk D contains the open disk B of radius $r-1$ around $0\in\mathbb{C}$. Thus $\lambda_i\notin B$ and $\Lambda(X)\geq|\lambda_i|\geq r-1$.

In the case $r>2$, we further have $\Lambda(X)\geq r-1\geq r-\frac{r}{2}=\frac{r}{2}$. ∎

Applying Lemma 3.4 to the left-hand side of equation (5) we obtain

$$\begin{aligned}\sum_{i\in I_L}\log\big(2d\Lambda\big(w(A_i,B_i)\big)\big) &\geq -\varepsilon_1 k + (1-\varepsilon_2)k\log\frac{1}{\varepsilon} - |I_M|\log 2d,\\ \sum_{i\in I_L}\log\big(\Lambda\big(w(A_i,B_i)\big)\big) &\geq -\varepsilon_1 k + (1-\varepsilon_2)k\log\frac{1}{\varepsilon} - \big(|I_M|+|I_L|\big)\log 2d\\ &> -\varepsilon_1 k + (1-\varepsilon_2)k\log\frac{1}{\varepsilon} - \varepsilon_2 k\log 2d\end{aligned}$$

since $|I_M|+|I_L|=|I|-|I_S|<\varepsilon_2 k$.

Combining this with equation (4) and Lemma 2.3, as $w(A,B)\in F^{n|w|}$, we conclude

$$\begin{aligned}n|w|\cdot\hat{h}(F)=\hat{h}(F^{n|w|}) &\geq \frac{1}{k}\sum_{i\in I_L}\log\big(R_i\big(w(A,B)\big)\big)\\ &\geq \frac{1}{k}\sum_{i\in I_L}\log\big(\Lambda\big(w(A_i,B_i)\big)\big)\\ &\geq -\varepsilon_1 + (1-\varepsilon_2)\log\frac{1}{\varepsilon} - \varepsilon_2\log 2d = n|w|\varepsilon_d.\end{aligned}$$

Thus $\hat{h}(F)\geq\varepsilon_d$ as desired in this situation.

3.1.2. Case 1b: $|I_S| \le (1-\varepsilon_2)k$. Recall that for any $i \notin I_S$, we have

$$\left|\mathrm{Tr}\big(w(A_i, B_i)\big) - I_d\right| \ge \varepsilon.$$

Since trace is conjugate-invariant, we know that $|\mathrm{Tr}(w(CA_iC^{-1}, CB_iC^{-1})) - I_d| \ge \varepsilon$ for all $C \in \mathrm{GL}_d(\mathbb{C})$. Thus by Corollary 2.9, for any $i \notin I_S$ and any $C \in \mathrm{GL}_d(\mathbb{C})$, at least one of the four matrices $CA_i^{\pm 1}C^{-1}, CB_i^{\pm 1}C^{-1}$ has norm no less than e^{δ}. Thus we have $E(T_i) \ge e^{\delta}$ for $T_i := \{A_i^{\pm 1}, B_i^{\pm 1}\}$.

In particular, this implies that for one element among $\{A^{\pm 1}, B^{\pm 1}\}$, let us say A, we have $\|A_i\| \ge e^{\delta}$ for $\frac{1}{4}\varepsilon_2 k$ different i's and that implies the lower bound

$$h(F^n) \ge h(A) \ge \frac{1}{4}\varepsilon_2\delta,$$

which implies

$$h(F) \ge \frac{\frac{1}{4}\varepsilon_2\delta}{n}.$$

As we want to obtain a bound for the normalized height and not only for the height, we need a way to compare them. To do this we will make use of the following two propositions, which are due to Breuillard. Recall that our A, B lie in $\mathrm{SL}_d(\mathbb{C})$.

Proposition 3.5 ([6, Proposition 2.9]). *For any positive integer M and a finite subset $F \subset \mathrm{SL}_d(\mathbb{C})$, we have*

$$E(F^M) \ge E(F)^{\sqrt{\frac{M}{4d}}}.$$

Proposition 3.6 ([6, Lemma 2.1(b)]). *There are uniform constants $c = c(d)$ and $N(d)$ such that for any finite subset $F \subset \mathrm{GL}_d(\mathbb{C})$ and any positive integer M there is $q \in [1, N(d)]$ with*

$$\Lambda(F^{qM}) \ge c^q E(F^M)^q.$$

The latter proposition can be deduced from an inequality of Bochi [4, Thereom B], which was also recently proved by Breuillard in [8] with a good choice of $c = c(d)$ (at the cost of increasing $N(d)$); see [8, Theorem 5]. The inequality of Breuillard together with [8, Lemma 1] implies that one can take $c = \frac{1}{2d}$.

These propositions imply that for every $i \notin I_S$, there is some q with

$$\Lambda(T_i^{qM}) \ge c^q E(T_i^M)^q \ge c^q E(T_i)^{q\sqrt{\frac{M}{4d}}} \ge c^q e^{q\delta\sqrt{\frac{M}{4d}}}.$$

Note that for $T := \{A^{\pm 1}, B^{\pm 1}\}$, we have $T \subset F^n$ and thus $T_i^{qM} \subset (\sigma_i F)^{nqM}$. Therefore, using the estimate above, we obtain

$$R_i(F)^{nqM} = R_i(F^{nqM}) = R\big((\sigma_i F)^{nqM}\big) \ge R(T_i^{qM}) \ge \Lambda(T_i^{qM}) \ge c^q e^{q\delta\sqrt{\frac{M}{4d}}}.$$

That is,

$$\log R_i(F) \ge \frac{1}{nM}\left(\log c + \delta\sqrt{\frac{M}{4d}}\right).$$

Here $\log(c) < 0$, so the right-hand side is maximized to $\frac{\delta^2}{16nd\log(1/c)}$ when

$$M = (16d\log^2 c)/\delta^2.$$

As M is an integer, we should take $M = \lfloor (16d\log^2 c)/\delta^2 \rfloor$ or $\lceil (16d\log^2 c)/\delta^2 \rceil$. Since $(16d\log^2 c)/\delta^2$ is sufficiently large, in either case we get

$$\log R_i(F) \geq \frac{\delta^2}{32nd\log\frac{1}{c}}.$$

Hence by equation (4), we have

$$\hat{h}(F) \geq \frac{1}{k}\sum_{i\notin I_S} \log^+ R_i(F) \geq \frac{\varepsilon_2\delta^2}{32nd\log\frac{1}{c}} = \varepsilon_d.$$

3.2. Case 2: $|Q| < e^{-\varepsilon_1 k}$

In this case, for $\alpha := \mathrm{Tr}(w(A,B) - I_d)$, we can apply the product formula Theorem 2.1 so that

$$\prod_{v\in V_K^f} |\alpha|_v^{n_v} = \frac{1}{|Q|} > e^{\varepsilon_1 k},$$

where V_K^f denotes the non-archimedean places of K.

Since $|\cdot|_v$ is an ultrametric for all $v \in V_K^f$, by Lemma 2.4 we have

$$|\alpha|_v \leq \max\left\{\left|\mathrm{Tr}\big(w(A,B)\big)\right|_v, \left|\mathrm{Tr}(I_d)\right|_v\right\} \leq \max\left\{R_v\big(w(A,B)\big), 1\right\}.$$

Hence for all $v \in V_K^f$ we have

$$\log|\alpha|_v \leq \log^+ R_v\big(w(A,B)\big).$$

Therefore, this inequality and the product formula together imply

$$\begin{aligned}
\hat{h}(F) = \frac{1}{n|w|}\hat{h}(F^{n|w|}) &\geq \frac{1}{n|w|}\hat{h}\big(w(A,B)\big) \\
&\geq \frac{1}{n|w|k}\sum_{v\in V_K^f} \log^+ R_v\big(w(A,B)\big) \\
&\geq \frac{1}{n|w|k}\sum_{v\in V_K^f} \log|\alpha|_v \\
&\geq \frac{1}{n|w|k}(\varepsilon_1 k) \\
&= \frac{\varepsilon_1}{n|w|} \\
&= \varepsilon_d.
\end{aligned}$$

4. Remarks about explicit estimates of the height gap ε_d

One can construct examples showing that the height gap $\varepsilon_d \leq \frac{c}{d}$ for some $c > 0$ independent of d. A natural question is to determine the actual order of ε_d and Breuillard has asked the following.

Question 4.1. *Is $\varepsilon_d \geq \frac{1}{Cd^C}$ for some $C > 0$?*

We will explain how our method might lead to obtain some explicit bounds for this constant.

From equation (3), we have that

$$\varepsilon_d \geq \frac{\delta^2 \log \frac{1}{\varepsilon}}{32 n_w d \log \frac{2d}{\varepsilon} \log \frac{1}{c} + 2n_w |w| \delta^2},$$

where c is the constant appearing in Proposition 3.6 and $\delta = \frac{\varepsilon}{8|w|}$. Recently Breuillard proved that c can be taken to $\frac{1}{2d}$ by [8, Theorem 5 and Lemma 1]. Therefore by taking $\varepsilon = 1/4$, we see that ε_d is at least at the order of

$$\frac{1}{n_w \left(d^2 |w|^2 + |w|\right)},$$

where $|w|$ is the word length of a $\frac{1}{4}$-almost law w on $\mathrm{U}(d)$ and the constant n_w comes from the escape of the hypersurface $\mathcal{X}$ defined by

$$\mathrm{Tr}\big(w(X,Y) - \mathrm{Id}\big) = 0$$

as in Lemma 3.1.

One can obtain a bound of the order of 10^{10^d} for $|w|$ by considering the almost laws described by Thom in [14]. It is nonetheless likely that a much shorter almost law w exists. Thom and Breuillard pointed out to us the fact that Kozma and Thom showed in [12] that for the symmetric group S_d (which in some ways behave similarly to $\mathrm{SU}(d)$ when d is large), there exists a law of order $e^{C \log^4(d) \log\log(d)}$, which assuming Babai's conjecture could be improved to be of the order $e^{C \log(d) \log\log(d)}$, which is quite close to be polynomial.

The constant n_w when computed from the generalized Bezout theorem as in [10] seems to be quite large. It was suggested by Breuillard that there seems to be another way of estimating n_w by considering an appropriate finite quotient of the group generated by F (for example by moding out a prime ideal of the ring where F is defined), and then showing that in this finite quotient one can escape from the hypersurface $\mathcal{X}$ quickly. The advantage is that there are various results about the diameter of simple groups. Michael Larsen suggested the use of a strong approximation theorem by Weisfeiler [15] to find the appropriate quotient.

To implement this idea, one must make sure the reduction of $\mathcal{X}$ to this finite group of Lie type is a proper subset. The dimension of the variety $\mathcal{X}$ is bounded by $d|w|$ and so the reduction of $\mathcal{X}$ can be proved to be a proper subset via the Schwartz–Zippel lemma, or via the Lang–Weil estimates, provided that the finite quotient is large enough (the prime ideal giving rise to the finite quotient must have covolume larger than the degree of $\mathcal{X}$, at least).

Assuming Babai's conjecture about the diameter of finite groups one then might expect that n_w could be taken of the order $(|d \log(|w|))^{O(1)}$ and using results towards Babai conjecture this can possibly be taken to be of the order $e^{(|d \log(|w|))^{O(1)}}$; see [2].

A further complication arises in order to use the previous approach. Consider, for example, the subset of F given by $\left[\begin{smallmatrix} 1 & n! \\ 0 & 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 & 0 \\ n! & 1 \end{smallmatrix}\right]$, where n is a large positive integer. In this case one must reduce $\mathcal{X}$ over a prime larger than n, and so the prime cannot be independent of F as we wanted, so one has to deal with these cases in a different way.

Acknowledgments. We would like to thank Emannuel Breuillard, Michael Larsen, and Andreas Thom for helpful discussions and suggestions. We particularly would like to thank Mikolaj Fraczyk, who noticed the connection between a previous version of our results and the work of Breuillard, and also Vladimir Finkelstein for multiple discussions. We are very grateful to thank Michael Larsen for pointing out a mistake in Lemma 3.1 from an early draft and for many discussions about Strong Approximation. HL would like to thank Dylan Thurston for his interest and various discussions. We would like to thank the anonymous referee for numerous helpful suggestions.

References

[1] N. Avni and C. Meiri, Words have bounded width in $\mathrm{SL}(n, \mathbb{Z})$. *Compos. Math.* **155** (2019), no. 7, 1245–1258 Zbl 1454.20097 MR 3963488

[2] J. Bajpai, D. Dona, and H. A. Helfgott, Growth estimates and diameter bounds for classical Chevalley groups. [v1] 2021, [v2] 2022, arXiv:2110.02942v2

[3] Y. Bilu, Limit distribution of small points on algebraic tori. *Duke Math. J.* **89** (1997), no. 3, 465–476 Zbl 0918.11035 MR 1470340

[4] J. Bochi, Inequalities for numerical invariants of sets of matrices. *Linear Algebra Appl.* **368** (2003), 71–81 Zbl 1031.15023 MR 1983195

[5] A. Borel, On free subgroups of semisimple groups. *Enseign. Math. (2)* **29** (1983), no. 1-2, 151–164 Zbl 0533.22009 MR 0702738

[6] E. Breuillard, A height gap theorem for finite subsets of $\mathrm{GL}_d(\bar{\mathbb{Q}})$ and nonamenable subgroups. *Ann. of Math. (2)* **174** (2011), no. 2, 1057–1110 Zbl 1243.11071 MR 2831113

[7] E. Breuillard, Heights on SL_2 and free subgroups. In *Geometry, rigidity, and group actions*, pp. 455–493, Chicago Lectures in Math., University Chicago Press, Chicago, IL, 2011 Zbl 1261.20047 MR 2807840

[8] E. Breuillard, On the joint spectral radius. In *Analysis at large*, edited by A. Avila, M. T. Rassias, and Y. Sinai, pp. 1–16, Springer, Cham, 2022 Zbl 07672481

[9] D. Calegari, Quasimorphisms and laws. *Algebr. Geom. Topol.* **10** (2010), no. 1, 215–217 Zbl 1185.20031 MR 2580432

[10] A. Eskin, S. Mozes, and H. Oh, On uniform exponential growth for linear groups. *Invent. Math.* **160** (2005), no. 1, 1–30 Zbl 1137.20024 MR 2129706

[11] M. Frączyk, S. Hurtado, and J. Raimbault, Homotopy type and homology versus volume for arithmetic locally symmetric spaces. 2022, arXiv:2202.13619v1

[12] G. Kozma and A. Thom, Divisibility and laws in finite simple groups. *Math. Ann.* **364** (2016), no. 1-2, 79–95 Zbl 1344.20024 MR 3451381

[13] G.-C. Rota and G. Strang, A note on the joint spectral radius. *Nederl. Akad. Wet. Proc. Ser. A. Indag. Math.* **22** (1960), 379–381 Zbl 0095.09701 MR 0147922

[14] A. Thom, Convergent sequences in discrete groups. *Canad. Math. Bull.* **56** (2013), no. 2, 424–433 Zbl 1276.54028 MR 3043070

[15] B. Weisfeiler, Strong approximation for Zariski-dense subgroups of semisimple algebraic groups. *Ann. of Math. (2)* **120** (1984), no. 2, 271–315 Zbl 0568.14025 MR 0763908

[16] S. Zhang, Small points and adelic metrics. *J. Algebraic Geom.* **4** (1995), no. 2, 281–300 Zbl 0861.14019 MR 1311351

Received 10 November 2022.

Lvzhou Chen
Department of Mathematics, Purdue University, 150 N University Street, West Lafayette, IN 47096, USA; lvzhou@purdue.edu

Sebastian Hurtado
Department of Mathematics, Yale University, 219 Prospect St, Office 713, New Haven, CT 06511, USA; sebastian.hurtado-salazar@yale.edu

Homin Lee
Department of Mathematics, Northwestern University, 2033 Sheridan Road, Evanston, IL 60208, USA; homin.lee@northwestern.edu