A NON CONCENTRATION ESTIMATE FOR RANDOM MATRIX PRODUCTS

EMMANUEL BREUILLARD

Abstract. Using the theory of random matrix products, we show that random walks on semisimple Lie groups do not concentrate on proper algebraic subgroups.

1. introduction

Let k be a local field of characteristic zero, i.e. $\mathbb{R}$, $\mathbb{C}$, a p-adic field. The goal of this note is the following result:

Theorem 1.1. *Let $\mathbb{G}$ be a connected semisimple algebraic group over k. Let μ be a probability measure on $\mathbb{G}(k)$ with a finite exponential moment. Assume that the closed subgroup G_μ generated by the support of μ is Zariski-dense and not compact in $\mathbb{G}(k)$. Then there is a positive constant $c = c(\mu) > 0$ such that for every integer $n \geqslant 1$, and every proper closed algebraic subgroup $\mathbb{H}$ of $\mathbb{G}$,*

$$\mu^n(\mathbb{H}) < e^{-cn}.$$

Recall that a probability measure μ on $\mathbb{G}(k) \subset \mathrm{GL}_d(k)$ is said to have a finite exponential moment if $\max\{\|g\|^\varepsilon, \|g^{-1}\|^\varepsilon\}$ is μ-integrable for some $\varepsilon > 0$ and some choice of operator norm on the matrix ring $M_d(k)$. This notion is easily seen to be independent of the linear embedding of $\mathbb{G}(k)$.

Here μ^n denotes the n-th fold convolution product of the probability measure μ. It is the distribution of the product of n independent $\mathbb{G}(k)$-valued random variables distributed according to μ.

Using the fact that a proper algebraic subgroup must fix a line in a suitable linear representation of $\mathbb{G}$, the proof of the theorem is reduced to the following proposition in which we take $y = x$.

Proposition 1.2. *(Probability of fixing a line) Assume that the support of μ generates a subgroup which is not relatively compact in projection to $PGL_d(k)$ and does not preserve any finite union of proper vector subspaces of k^d. Then there is $c = c(\mu) > 0$ such that for every $n \geqslant 1$ and every lines $x, y \in \mathbb{P}(k^d)$,*

$$\mu^n(\{g \in \mathrm{GL}_d(k); g(x) = y\}) < e^{-cn}.$$

1991 *Mathematics Subject Classification.* 20G40, 20N99.

This proposition is well-known in the case when the action of the group generated by the support of μ on k^d is proximal, i.e. when it admits elements with a single eigenvalue of maximal modulus. In that case, it is known that the probability of the event $\{g(x) \in V\}$ decays exponentially fast, for any given point $x \in \mathbb{P}(K^d)$ and proper projective subspace V, with a rate of decay uniform in x and V (see e.g. [16] or [7, Chap. VI.]). Hence the main novelty of our proposition is the removal of the proximality assumption. The cost to pay is that we have to restrict the class of events to the case when V is a point. In fact, there is some room to allow somewhat larger subspaces V, and the exact condition $\{g(x) \in V\}$ can be replaced with $\{d(g(x), V) < e^{-c'n}\}$ for a suitable $c' > 0$. A precise statement is formulated in Proposition 5.1 below.

Using a suitable field embedding, we obtain the following Corollary of Theorem 1.1.

Corollary 1.3. *(Probability of return to a subgroup) Let $\mathbb{G}$ be a connected semisimple algebraic group over an arbitrary field K of characteristic zero, and $\Gamma \leqslant \mathbb{G}(K)$ a Zariski-dense subgroup generated by a finite set S. Let μ be a probability measure on S with $\mu(s) > 0$ for each $s \in S$. Then there is a positive constant $c > 0$ such that for every integer $n \geqslant 1$,*

$$\mu^n(\mathbb{H}) < e^{-cn}, \tag{1.1}$$

uniformly for every proper closed algebraic subgroup $\mathbb{H}$ of $\mathbb{G}$.

Note the difference in the assumptions in this corollary, compared to Theorem 1.1: the measure is assumed to be finitely supported, but the field K is arbitrary and in particular no assumption is made on the group generated by the support of μ besides Zariski density.

Earlier instances of this corollary were known in a number of cases. For example if $\mathbb{G} = \mathrm{SL}_d$ and S belongs to $\mathrm{SL}_d(\mathbb{Z})$, then a result of Goldsheid and Margulis [13] allows to exhibit proximal elements in associated representations and (1.1) then follows from the known estimates from random matrix theory mentioned earlier. Unfortunately the Goldsheid-Margulis result fails for p-adic fields (see e.g. [19]).

In [1], R. Aoun proved the non-uniform version of Theorem 1.1 (i.e. with a exponential rate of decay depending on $\mathbb{H}$) when $\mathbb{G}$ is a semisimple real algebraic group. This argument does not give uniformity readily and it does not extend to other local fields, due to its reliance the Golsheid-Margulis theorem.

Recently, in their work on the spectral gap for dense subgroups of compact Lie groups, Benoist and Saxcé [3, Prop. 3.2] showed this result in the case when $K = \mathbb{R}$ and $\mathbb{G}(\mathbb{R})$ is compact, again by exhibiting suitable proximal representations.

Non concentration estimates such as (1.1) are crucial for establishing spectral gaps for the action of a finite set S on various unitary representations of $\langle S \rangle$. It is one of the ingredients of the Bourgain-Gamburd method (see e.g. [12, §3]) developed in [9] to exhibit a wide family of Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$, p prime, with uniform spectral gap (i.e. a family of expander graphs).

One of our motivations for proving Corollary 1.3 comes from the survey paper [11], in which we use Corollary 1.3 to give an alternate proof of the main non concentration estimate needed in the proof of the super-strong approximation theorem [11, Theorem 1.2] for congruence quotients modulo a prime. The super-strong approximation theorem asserts that when $K = \mathbb{Q}$, the Cayley graphs obtained by reducing modulo p a given Zariski dense subgroup of $\mathbb{G}(\mathbb{Q})$ form a family of expanders. This result (and the stronger version when quotients modulo an arbitrary square free integer are allowed) was proved by Salehi-Golsefidy and Varjú [14] following the Bourgain-Gamburd method and using a ping-pong argument à la Tits instead of Corollary 1.3 for establishing the required non-concentration estimate.

Unsurprisingly Corollary 1.3 can be deduced from the super-strong approximation theorem (as in [11, Thm 7.2] or [10, Cor 1.1]) by reducing modulo a suitable prime. However we see Corollary 1.3 as an ingredient rather than a by-product of the proof of the super-strong approximation theorem.

Remark. It is interesting to study the dependence in μ of the rate of exponential decay $c(\mu)$ is Corollary 1.3. We will show in a subsequent work that it can be taken to be independent of the generating set S and the field K, provided $\mu(s)$, $s \in S$, is bounded away from zero.

2. Notation

In this note k denotes a local field of characteristic zero, V a finite dimensional k-vector space and $\mathbb{P}(V)$ its projective space. Now μ will denote a probability on $\mathrm{GL}(V)$. It is said to have a *finite exponential moment* if there is $\tau > 0$ such that

$$\int \max\{||g||, ||g^{-1}||\}^{\tau} d\mu(g) < +\infty.$$

Here $||\cdot||$ denotes the operator norm on the endomorphisms of V, which is induced by a norm on V, namely $||g|| = \sup\{||gx||; x \in V, ||x|| = 1\}$. This definition is independent of the norm. In this paper we will only consider norms on V of the following special type: given a k-basis $(e_1, \dots, e_d)$ of V and $x = \sum x_i e_i$, we take the ℓ^2 norm $||x|| := (\sum |x_i|^2)^{1/2}$ if k is archimedean (i.e. $k = \mathbb{R}$ or $\mathbb{C}$) and we take the ℓ^∞-norm $||x|| := \max |x_i|$ if k is non archimedean.

$S_n := Y_1 \cdot \ldots \cdot Y_n$ will denote the right random walk on $GL(V)$ associated with μ. Namely the Y_i's are i.i.d. random variables distributed according to μ.

G_μ denotes the smallest closed subgroup of $GL(V)$ (for the Hausdorff topology induced by k) containing the support of μ, i.e. the set of points $g \in GL(V)$ such that $\mu(B_x) > 0$ for every open set B_x containing x.

We say that G_μ is *strongly irreducible* if there is no finite union of proper k-subspaces of V which is invariant under G_μ.

The *proximal dimension* (also called *index*) of G_μ is the smallest rank p of an endomorphism in the closure of the set $\{\frac{g}{||g||}, g \in G_\mu\}$ inside the unit sphere $\{m \in End(V); ||m|| = 1\}$. Note that the proximal dimension is always non zero. It is equal to $d = \dim V$ if and only if G_μ has compact image in $PGL(V)$. We say that G_μ is proximal if its proximal dimension is 1.

Assuming μ has a finite first moment (i.e. $\mathbb{E}(|\log(||g^{\pm 1}||)|) < +\infty$), we denote by $\lambda_1 \geqslant ... \geqslant \lambda_d$ the *Lyapunov exponents* of μ. They are defined by the formula:

$$\lim_{n \to +\infty} \frac{1}{n}\mathbb{E}(\log ||\Lambda^k S_n||) = \lambda_1 + \ldots + \lambda_k,$$

where $\Lambda^k g$ is the image of $g \in GL(V)$ under the k-th exterior representation $\Lambda^k V$. Note that a choice of a basis of V induced a basis of $\Lambda^k V$ and thus a norm on V (associated to a basis as above) induces a norm on $\Lambda^k V$; this will always be our choice of norm on $\Lambda^k V$.

According to a fundamental result of Guivarc'h-Raugi [17], we have $\lambda_1 = \ldots = \lambda_p > \lambda_{p+1}$ (see [7, Prop. III.6.2.]). The proof is given there in the case $k = \mathbb{R}$, but it extends verbatim to any local field.

3. Large deviations

In this section we assume that μ is a probability measure on $\mathrm{GL}(k^d)$ with finite exponential moment and that the closed subgroup G_μ generated by its support acts strongly irreducibly on k^d. We now recall the following large deviations estimate for random matrix products.

Theorem 3.1. *(Large deviations) Given $\varepsilon > 0$, there is $c_1 = c_1(\varepsilon, \mu) > 0$ such that for all $v \in k^d \setminus \{0\}$ and all $n \geqslant 1$,*

$$\mathbb{P}(|\frac{1}{n}\log\frac{||S_n \cdot v||}{||v||} - \lambda_1| > \varepsilon) \leqslant e^{-c_1 n}, \tag{3.1}$$

$$\mathbb{P}(|\frac{1}{n}\log ||S_n|| - \lambda_1| > \varepsilon) \leqslant e^{-c_1 n}, \tag{3.2}$$

In the case when $k = \mathbb{R}$ and G_μ acts proximally (i.e. with proximal dimension $p = 1$), this estimate was first proved by Le Page [18], using his spectral gap result for the action on Holder functions on projective space. His proofs extend without difficulty to the case of an arbitrary local field; details were worked out by Guimier [15], and more recently Aoun [2] and Benoist-Quint [5, Theorems 8.1, 9.16]. However these arguments require a proximality assumption. To deal with the general case, one needs to reduce to the proximal case by considering the wedge power representation $\Lambda^p k^d$, which is proximal but no longer strongly irreducible. This was worked out by Bougerol in [7, Theorems 6.1, 6.2] in the real case.

In this section we detail Bougerol's argument for Theorem 3.1 (and its extension to local fields), because it relies on Lemma 3.2 below, a statement that will be useful in the proof of Propostion 1.2, where a similar difficulty occurs.

In the next lemma $G \leqslant \mathrm{GL}(k^d)$ denotes a closed subgroup acting strongly irreducibly on k^d. Its proximal dimension is defined as the minimal $p \geqslant 1$ such that there are matrices of rank p among the limit points of the matrices $\frac{g}{||g||}$, $g \in G$. If $p = d$, then G is relatively compact in projection to $\mathrm{PGL}(k^d)$, so we assume $p < d$. The limit points of rank p form a subset of $M_d(k)$ denoted by Π. Let $W \leqslant \Lambda^p k^d$ be the span of the lines $Im(\Lambda^p \pi)$, where π ranges over Π.

We note that the set Λ_Π of lines of the form $Im(\Lambda^p\pi)$, $\pi \in \Pi$, is a closed G-invariant subset of $\mathbb{P}(\Lambda^p k^d)$ and that the G-action on it is minimal, i.e. every orbit is dense in Λ_Π: indeed if $\pi, \pi' \in \Pi$, there is $g \in G$ and $\{g_n\} \in G$ such that $gIm(\pi) \nsubseteq \ker \pi'$, i.e. $\pi' g \pi \neq 0$ and $\pi' = \lim \frac{g_n}{||g_n||}$. Hence $Im(\Lambda^p \pi') = \lim Im(\Lambda^p g_n g \pi)$. Moreover we have:

Lemma 3.2. *(Decomposition of $\Lambda^p k^d$) Let U_0 be the G-invariant subspace $\bigcap_{\pi\in\Pi} \ker \Lambda^p \pi$. We have a direct sum decomposition:*

$$\Lambda^p k^d = W \oplus U_0. \tag{3.3}$$

The action of G on W is strongly irreducible and proximal, and there is $C_1 = C_1(G) > 0$ such that for all $g \in G$,

$$C_1^{-1}||g||^p \leqslant ||(\Lambda^p g)_{|W}|| \leqslant ||g||^p. \tag{3.4}$$

This lemma is taken from Benoist-Quint [4, Lemma 4.13] with (3.3) in extra.

Proof. Suppose $U \leqslant \Lambda^p k^d$ is a G-invariant subspace. First we prove the following claim: either U contains W, or it is contained in U_0.

For this, assume that for some $\pi \in \Pi$, $Im(\Lambda^p\pi) \nsubseteq U$. This implies that $U \leqslant \ker \Lambda^p \pi$, because otherwise, picking $v \in U$ not in $\ker \Lambda^p\pi$, we would get $Im(\Lambda^p \pi) = k\Lambda^p \pi v \leqslant U$, because U is invariant under $\Lambda^p \pi$.

Now every other $\omega \in \Pi$ will satisfy $Im(\Lambda^p\omega) \nsubseteq U$, and hence $U \leqslant \ker \Lambda^p \omega$. Indeed, pick $\gamma \in G$ such that $\pi\gamma\omega \neq 0$. This is possible, since otherwise $G \cdot Im(\omega) \subset \ker \pi$, contradicting the irreducibility of G on k^d. Now by minimality of p, the rank of $\pi\gamma\omega$ must be equal to p. Since U is preserved by $\Lambda^p\gamma$ and $\Lambda^p\pi$, if $Im(\Lambda^p)\omega$ were contained in U, so would be $Im(\Lambda^p\pi)$, contrary to our hypothesis. It follows that either $Im(\Lambda^p\pi) \leqslant U$ for all $\pi \in \Pi$, and U contains W, or U is contained in $U_0 := \cap \ker \Lambda^p \pi$. The claim is proved.

The two alternatives are mutually exclusive, because W is not contained in U_0: indeed if $\pi \in \Pi$, as above by irreducibility of G on k^d there is $\gamma \in G$ such that $\pi\gamma\pi \neq 0$, and by minimality of p this implies that $\Lambda^p\gamma\pi$ has rank one, and $Im(\Lambda^p\gamma\pi) \nsubseteq \ker \Lambda^p\pi = \ker \Lambda^p\gamma\pi$.

Now we claim that the same is true of every G_0-invariant subspace U, where G_0 is any finite index subgroup of G. Indeed, first note that since G acts strongly irreducibly so does G_0, and then observe that the span W_0 of the images of the elements $\Lambda^p\pi_0$, where π_0 is a rank p element in the closure of kG_0, coincides with W. Indeed pick $\pi \in \Pi$ such that $Im(\pi) \nsubseteq \ker(\pi)$, then $Im(\pi^n) = Im(\pi)$ and similarly for all $g\pi g^{-1}$, $g \in G$. Since W is spanned by the images of the $\Lambda^p(g\pi g^{-1})$, this shows that $W = W_0$. It follows that the first part of the proof applies verbatim with G_0 and W_0 in place of G and W. This shows our claim.

Now, note that the Zariski-closure $\mathbb{G}$ of G in $GL_d(k)$ is a reductive group, because the action on k^d is irreducible. Let G_0 be the intersection of G with the connected component of the identity in $\mathbb{G}$. Since k has characteristic zero, every linear representation of G_0 is completely reducible. In particular $W \cap U_0$ has a G_0-invariant complement say U_1 in W. By the above U_1 must be contained in U_0 or contain W. This forces $W \cap U_0$ to be trivial. This argument uses characteristic zero in a key way, there are examples in positive characteristic where $W \cap U_0$ is non trivial.

Similarly we get that the action on W is strongly irreducible: any G_0-invariant proper subspace of W must be contained in U_0, hence is trivial. Its proximal dimension is one of course.

By the same token, we get that $W + U_0$ is all of $\Lambda^p k^d$, because otherwise we would find a G_0-invariant complement, which neither contains W nor is contained in U_0. This establishes (3.3).

It remains to prove the lower bound in (3.4). By contradiction, if no such C_1 existed we would find a sequence $\{g_n\} \in G$ such that $||(\Lambda^p g_n)_{|W}||/||g_n||^p$ tends to 0. Pick a limit π of $g_n/||g_n||$ in $M_d(k)$. It must have rank p and $(\Lambda^p g_n)/||g_n||^p$ converges to $\Lambda^p\pi$. Changing $\{g_n\}_n$ into $\{\gamma g_n\}_n$ for some suitably chosen $\gamma \in G$ if necessary, we may assume that $Im(\pi)$ is not contained in $\ker \pi$. Now if $v \in Im(\Lambda^p\pi) \setminus \{0\}$, then $\Lambda^p\pi v \neq 0$. However v belongs to W and so $(\Lambda^p g_n v)/||g_n||^p$ tends to 0, which is a contradiction. This ends the proof of the lemma. □

Remark. In positive characteristic the decomposition (3.3) is no longer true. In fact $U_0 \cap W$ may be non trivial, as in the case of adjoint representation of $\mathrm{SL}_n(k)$ when $n = char(k)$ (the identity matrix spans U_0 and W is the whole Lie algebra). See the remark at the end of this section for more on positive characteristic.

Lemma 3.3. *There is a constant $C_2 = C_2(G) > 0$ such that given any $v \in k^d \setminus \{0\}$, one can find a p-plane P containing v such that*

$$||\pi_W(v_P)|| \geqslant \frac{1}{C_2}||v_P||, \tag{3.5}$$

where $v_P := \Lambda^p P \notin U_0$ and π_W denotes the projection onto W in (3.3). *Similarly, given a hyperplane H of k^d, one may find a p-plane P inside H such that* (3.5) *holds.*

Proof. We may assume that $||v|| = 1$ without loss of generality. Arguing by contradiction, we would then obtain a sequence of vectors $\{v_n\}_{n\geqslant 1}$ of norm 1 such that for every p-plane P containing v_n, $||\pi_W(v_P)|| \leqslant \frac{1}{n}$, where we set v_P to be the norm 1 vector in $\Lambda^p k^d$ representing P. Taking a limit we obtain a vector $v \in k^d$ of norm 1 such that every p-plane containing it has non trivial intersection with U_0. However by irreducibility of G, given $\pi \in \Pi$, we may find $g \in G$ such that $g(v) \notin \ker \pi$, hence $v \notin \ker \pi'$, where $\pi' = g^{-1}\pi \in \Pi$. Any complement P of $\ker \pi'$ containing v will intersect U_0 trivially. This ends the proof. The case of the hyperplane is analogous. □

Proof of Theorem 3.1. As already mentioned, in the case when the proximal dimension of G_μ is one, then the theorem is well-known. See for example [5, Theorem 8.1(iii)] for a proof. It is also well-known if the proximal dimension is $p = d$, because then G_μ is relatively compact in projection to $\mathrm{PGL}(k^d)$ and thus $||g||^d/|\det(g)|$ is bounded above and below uniformly in $g \in G$. Then $\log|\det(S_n)|$ is a sum of i.i.d random variables, and (3.2) follows from the classical large deviation estimate for i.i.d real random variables. Similarly (3.1) follows since $||g \cdot v||/(||g|| \cdot ||v||)$ is bounded uniformly in $g \in G$ and $v \in k^d \setminus \{0\}$.

Hence we may assume that the proximal dimension is $p < d$. One needs to reduce to the proximal case by considering the wedge power representation $\Lambda^p k^d$. Note first that (3.2) follows immediately from (3.4) and the fact that the action of G_μ on W is strongly irreducible and proximal. To prove (3.1), we will use Lemma 3.3. From (3.2) and the upper bound $||S_n \cdot v|| \leqslant ||S_n|| \cdot ||v||$, it is enough to establish that a similar lower bound holds with the desired probability. We apply Lemma 3.3 to v and find a p-plane P containing v with (3.5). We may find vectors $v_1, \ldots, v_{p-1}$ in P with $||v_1 \wedge \ldots \wedge v_{p-1}|| = 1$, $v_P = v \wedge v_1 \wedge \ldots \wedge v_{p-1}$ and $||v_P|| = ||v||$. We then get that $||S_n||^{p-1}||S_n \cdot v|| \geqslant ||\Lambda^p S_n \cdot v_P|| \geqslant \frac{1}{C'}||\Lambda^p S_n \cdot \pi_W(v_P)||$ (where $C' = C'(G) > 0$ satisfies $\forall u \in \Lambda^p k^d$, $||u|| \geqslant \frac{1}{C'}\max\{||\pi_W(u)||, ||\pi_{U_0}(u)||\}$). Hence

$$\frac{||S_n \cdot v||}{||v||} \geqslant \frac{1}{C_1 C_2 C'} \frac{||\Lambda^p S_n \cdot \pi_W(v_P)||}{||(\Lambda^p S_n)_{|W}|| \cdot ||\pi_W(v_P)||}||S_n||.$$

Since the action of G_μ on W is proximal and strongly irreducible, the large deviation estimate holds there uniformly in $\pi_W(v_P)$ and we are now done. □

Remark about positive characteristic. In this note we have assumes throughout that k has characteristic zero, however we believe that Theorem 3.1 continues to hold for local fields of positive characteristic. The proof however will have to go back to more foundational results in random matrix products, which are not available to date. In particular, due to the lack of complete reducibility of reductive groups in characteristic p, one needs to establish first the basic theorems of Furstenberg (as well as the large deviation estimate) without the irreducibility assumption, because one just cannot reduce to this case in general. For example using the non-vanishing of the first cohomology group, one can build examples of proximal indecomposable, yet non irreducible, modules which are non trivial extensions of an irreducible module by the trivial module: in such cases W can be a proper subspace, while U_0 is trivial. Besides $\Lambda^r V$ is not completely reducible in general, even if V is.

We have assumed in Theorem 3.1 that the action of G_μ on V is strongly irreducible. We will need to relax this hypothesis somewhat:

Theorem 3.4. *Let μ be a probability measure with finite exponential moment on* $\mathrm{GL}(V)$*. Assume that G_μ is completely reducible. Then given $\varepsilon > 0$ there is $c_1 > 0$ such that for all $n \geqslant 1$,*

$$\mathbb{P}(|\frac{1}{n}\log||S_n|| - \lambda_1| > \varepsilon) \leqslant e^{-c_1 n},$$

Proof. It is enough to consider the case when V is irreducible. Then the connected component of the identity $\mathbb{G}$ of the Zariski closure of G_μ is a reductive group, hence it acts completely reducibly on V. Let as before G_0 denote the intersection of G_μ with $\mathbb{G}$, which is a subgroup of finite index in G. Consider the n-th return time τ_n to G_0. Then S_{τ_n} is a random product of length n whose increments are distributed according to μ_0, the law of S_{τ_1}. R. Aoun proved in [2, Lemmas 4.40, 4.42] that μ_0 has finite exponential moment, that $\mathbb{E}(\tau_1) < \infty$ and $\lambda_1(\mu_0) = \mathbb{E}(\tau_1)\lambda_1(\mu)$, and that for all $\varepsilon > 0$, the event $|\tau_{[\frac{n}{\mathbb{E}(\tau_1)}]} - n| > \varepsilon^2 n$ has exponentially small probability. The estimate (3.4) then follows from (3.2) applied to μ_0, because

$$\big|\log||S_{\tau_{[\frac{n}{\mathbb{E}(\tau_1)}]}}|| - \log||S_n||\big| \leqslant \max_{k \leqslant \varepsilon^2 n} \big|\sum_1^k \log||Y_i||\big|$$

and the right hand side, being a sum of i.i.d variables, is $< \varepsilon n$ with probability tending to 1 exponentially fast in n if ε is chosen small enough. □

4. Distance in the Grassmannian and Cartan decomposition

In this section we prove Lemma 4.2 below, which is an extension to Grassmannians of what is done in [8, Theorem 4.4] and [5, Lemma 12.2] for the projective space.

Recall that k is a local field and V is a finite dimensional k-vector space. A norm on V associated to a basis (see the paragraph of notation above) induces by duality a norm on the dual V^*, namely the norm associated to the dual basis. Similarly they induce norms on the exterior powers $\Lambda^k V$ and $\Lambda^k(V^*)$. We now identify V with k^d endowed with its canonical basis.

Given subspaces P, Q in V define:

$$d(P,Q) = \frac{||v_P \wedge v_Q||}{||v_P|| \cdot ||v_Q||},$$

where $v_P = v_1 \wedge \ldots \wedge v_p$ for some basis $v_1, \ldots v_p$ of P, and similarly for v_Q. Clearly $d(P,Q)$ is independent of the choice of basis used to define v_P and v_Q. Note that $d(P,Q) = 0$ iff P and Q have non trivial intersection. So $d(P,Q)$ is not a distance, in fact it does not satisfy the triangle inequality, except on $P(V)$, that is when P and Q are lines. Indeed it is well-known (see e.g. [6, Prop. 2.8.18]) that $d(x,y)$ is a distance on the projective space for $x, y \in \mathbb{P}(V)$.

Lemma 4.1. *Given two subspaces $P, Q \leqslant V$,*

$$d(P,Q) \leqslant d(x,Q) = \min_y d(x,y) \tag{4.1}$$

for every line x in P, where the y's range over all lines contained in Q.

Proof. The archimedean and non-archimedean cases have to be treated separately. The inequality $||v \wedge w|| \leqslant ||v|| \cdot ||w||$ holds for all $v, w \in \Lambda^* V$. To see the left hand side of (4.1), let $v_1 \in V$ of norm 1 representing x, and complete it into a basis of P using vectors $v_2, \ldots, v_p$ such that $||v_1 \wedge \ldots \wedge v_p|| = ||v_2 \wedge \ldots \wedge v_p|| = 1$. In the archimedean case one achieves this by choosing the v_i's in an orthogonal complement of v_1 in P. In the non archimedean case one notes that $GL_d(\mathcal{O}_k)$ acts transitively on p-planes for each p, so one can assume that $v_1 = e_1$ and P is the span of the first $\dim P$ elements of the canonical basis of k^d. Then $||v_P \wedge v_Q|| \leqslant ||v_1 \wedge v_Q||$ and hence $d(P,Q) \leqslant d(x,Q)$ as desired. The right hand side of (4.1) is similar and left to the reader.

□

For a matrix $g \in GL(k^d)$, let $g = ua\ell$ be its Cartan decomposition, where u and ℓ belong to the maximal compact subgroup K (K is the stabilizer of the norm associated to the canonical basis $e_1, \ldots, e_d$ on k^d, i.e. $K = U(d, \mathbb{C}), O(d, \mathbb{R})$ or $GL_d(\mathcal{O}_k)$), and where a is diagonal with entries $a_1, \ldots, a_d$ with $|a_1| \geqslant \ldots \geqslant |a_d|$. Fix an integer r between 1 and d. Let $V_g^+ = u\langle e_1, \ldots, e_r \rangle$, $V_g^- = \ell^{-1}\langle e_{r+1}, \ldots, e_d \rangle$. Note that $(V_g^-)^\perp = (V^*_{{}^tg})^+$.

Lemma 4.2. *Let $V = k^d$. For every r-dimensional subspaces $P \leqslant V$ and $R \leqslant V^*$, every line $x \in \mathbb{P}(V)$, and every $g = ka\ell \in GL(k^d)$, we have*

(i) $d(P, V_g^-) \leqslant \frac{||\Lambda^r g v_P||}{||\Lambda^r g|| \cdot ||v_P||} \leqslant d(P, V_g^-) + \frac{|a_{r+1}|}{|a_r|}$,

(ii) $d(V_g^+, R^\perp) \leqslant \frac{||\Lambda^r \ {}^t g v_R||}{||\Lambda^r \ {}^t g|| \cdot ||v_R||} \leqslant d(V_g^+, R^\perp) + \frac{|a_{r+1}|}{|a_r|}$,

(iii) $d(gx, V_g^+) \cdot d(x, V_g^-) \leqslant \frac{|a_{r+1}|}{|a_r|}$.

Proof. First note that we may assume that $g = a$ is a diagonal matrix. Then $||\Lambda^r a|| = |a_1 \cdot \ldots \cdot a_r|$. Writing $v_P = v_1 \wedge \ldots \wedge v_r$ for some basis $v_1, \ldots, v_r$ of P, decomposing $v_i = v_i^+ + v_i^-$ with $v_i^+ \in \langle e_1, \ldots e_r \rangle$, and $v_i^- \in \langle e_{r+1}, \ldots, e_d \rangle$, and expanding the wedge product, we see that:

$$||v_1^+ \wedge \ldots \wedge v_r^+|| \leqslant \frac{||\Lambda^r a \cdot v_P||}{|a_1 \cdot \ldots a_r|} \leqslant ||v_1^+ \wedge \ldots \wedge v_r^+|| + \frac{|a_r|}{|a_{r+1}|}||v_P||.$$

Since $v_1^+ \wedge \ldots \wedge v_r^+ \wedge e_{r+1} \wedge \ldots \wedge e_d = v_P \wedge v_{\langle e_{r+1}, \ldots, e_d \rangle}$, we obtain the first item.

The second item follows from the first by duality since $d_V(V_g^+, R^\perp) = d_{V^*}(R, (V^*_{{}^tg})^-)$. For the third item, write $x = x_1 + x_2$ with $x_1 \in \langle e_1, \ldots, e_r \rangle$ and $x_2 \in \langle e_{r+1}, \ldots, e_d \rangle$ and note that $d(ax, \langle e_1, \ldots, e_r \rangle) \leqslant \frac{||ax_2||}{||ax||}$, while $d(x, \langle e_{r+1}, \ldots, e_d \rangle) = \frac{||x_1||}{||x||}$. The inequality is then a simple check (in both the archimedean and non archimedean cases). □

5. Proof of Proposition 1.2

The proof is a combination of the large deviation estimate of Theorem 3.1 with Lemmas 4.2 and 3.3. As above p denotes the proximal dimension.

Given points x, y in projective space $\mathbb{P}(V)$, we apply Lemma 3.3 and find a p-plane P in V and a p-plane R in the dual V^* such that P contains x (resp. such that y is contained in $R^\perp := \cap_{f \in R} \ker f$) and

$$||\pi_W(v_P)|| \geqslant \frac{1}{C_2}||v_P|| \tag{5.1}$$

(resp. $||\pi_{W'}(v_R)|| \geqslant \frac{1}{C_2}||v_R||$). Here π_W denotes as before the projection map associated to the decomposition (3.3) from $\Lambda^p V$ onto W (resp. $\pi_{W'}$ the analogous projection onto the corresponding subspace of $\Lambda^p V^*$ which we denoted by W'), and $C_2 > 0$ is a constant independent of x, y.

Now we apply Lemma 4.2 with $r = p$ to $g = S_n$ the random product of length n. By the Guivarc'h-Raugi theorem, we know that $\lambda_{p+1} < \lambda_p$, hence according to Theorem 3.4 (note that all $\Lambda^k V$ are completely reducible, because V is), if $0 < \varepsilon < (\lambda_p - \lambda_{p+1})/10$ say,

$$\frac{|a_{p+1}|}{|a_p|} = \frac{||\Lambda^{p+1} S_n||}{||\Lambda^p S_n||} \leqslant e^{-4\varepsilon n}$$

with probability $> 1 - \rho^n$ for some $\rho = \rho(\varepsilon) > 0$ and for all $n \geqslant n_0(\varepsilon)$. Similarly using Theorem 3.1, the fact that W is strongly irreducible, and (5.1), we get that

$$\frac{||\Lambda^p S_n v_P||}{||\Lambda^p S_n|| \cdot ||v_P||} \geqslant e^{-\varepsilon n}$$

with probability $> 1-\rho^n$. By Lemma 4.2 (i) it follows that

$$d(x, V^-_{S_n}) \geqslant e^{-2\varepsilon n}$$

for all n larger than some $n_1 = n_1(\varepsilon) > 0$ independent of x. Similarly using Lemma 4.2 (ii), we prove that

$$d(y, V^+_{S_n}) \geqslant e^{-2\varepsilon n}$$

with probability $> 1-\rho^n$. However by Lemma 4.2 (iii), we have,

$$d(S_n x, V^+_{S_n}) \cdot d(x, V^-_{S_n}) \leqslant e^{-4\varepsilon n},$$

with probability $> 1-\rho^n$. Hence

$$d(S_n x, V^+_{S_n}) \leqslant e^{-3\varepsilon n}$$

. And using Lemma 4.1, we conclude that

$$d(S_n x, y) \geqslant d(y, V^+_{S_n}) - d(S_n x, V^+_{S_n}) \geqslant e^{-3\varepsilon n}$$

holds for all n larger than some $n_2(\varepsilon) > 0$ and with probability at least $1-\rho_1^n$ for some $\rho_1 = \rho_1(\varepsilon) < 1$. This ends the proof of Proposition 1.2.

We now end by stating a strengthening of Proposition 1.2, whose proof is entirely similar to the above and is left to the reader.

If $\mathrm{Grass}(r,d)$ denotes the Grassmannian of r-planes in k^d (for $1 \leqslant r \leqslant \dim V - 1$), then the subset $\mathcal{G}(r,\mu)$ of all $V \in \mathrm{Grass}(r,d)$ with non trivial intersection with $Im(\pi)$ for some $\pi \in \Pi$, is a closed (algebraic) subset of $\mathrm{Grass}(r,d)$. It is empty if $r = 1$ (i.e. for the projective space $\mathbb{P}(k^d)$) due to the irreducibility of G_μ on k^d. It is all of $\mathrm{Grass}(r,d)$ if $r > d-p$, by definition of the proximal dimension p.

We have:

Proposition 5.1. *Keep the assumptions of Proposition 1.2. Let $1 \leqslant r \leqslant d-p$. Given $\varepsilon > 0$, there is $c = c(\mu,\varepsilon) > 0$ such that, for every open set $\mathcal{U}$ of* $\mathrm{Grass}(r,d)$ *containing $\mathcal{G}(r,\mu)$, there is $n_0 = n_0(\varepsilon,\mu,\mathcal{U}) > 0$ such that for every $n \geqslant n_0$*

$$\mu^n(\{g \in \mathrm{GL}_d(k); d(g(x), V) < e^{-\varepsilon n}\}) < e^{-cn},$$

for every subspace $V \in \mathrm{Grass}(r,d) \setminus \mathcal{U}$ and every $x \in \mathbb{P}(k^d)$.

6. Proof of Theorem 1.1 and Corollary 1.3

We keep the notation of Theorem 1.1. Note first that up to passing to a finite extension of k, we may assume that $\mathbb{G}$ is k-split. Theorem 1.1 then follows immediately from the combination of Proposition 1.2 and the following observation:

Lemma 6.1. *There are finitely many non trivial irreducible k-modules of $\mathbb{G}$, say $\pi_1, \ldots, \pi_t$ such that every subgroup of $\mathbb{G}(k)$ which is not Zariski dense must fix a line in one of these modules.*

Proof. Consider the adjoint representation $(\rho_1, Lie(\mathbb{G}))$ of $\mathbb{G}$. If $\Gamma \leqslant \mathbb{G}(k)$ is not Zariski dense, but is infinite, then it preserves the Lie algebra of its Zariski-closure, which is a non trivial proper subspace of $Lie(\mathbb{G}(k))$. If it is finite, then Jordan's theorem tells us that Γ has a normal abelian subgroup of index bounded in terms of $\dim \mathbb{G}$ only. Let ρ_2 be any irreducible k-module of $\mathbb{G}$ whose dimension is larger than this bound. Then no finite subgroup of $\mathbb{G}$ will act irreducibly on it. So every non Zariski dense subgroup of $\mathbb{G}$ must preserve a proper subspace of either ρ_1 or ρ_2. The lemma follows if we take for the π_i's the collection of irreducible submodules appearing in the decomposition of the wedge powers of ρ_1 and ρ_2. □

Now Corollary 1.3 is an easy consequence of Theorem 1.1 via the following well-known lemma, for which we refer to [20, lemma 4.1.].

Lemma 6.2. *If K is a finitely generated field and $t \in K^{\times}$ an element of infinite order, then there is an embedding of K into a local field k with absolute value $|\cdot|$, for which $|t| \neq 1$.*

Proof of Corollary 1.3. Since Γ is finitely generated, we may assume that K is a finitely generated extension of $\mathbb{Q}$. There are thus only finitely many roots of unity which are roots of polynomials of bounded degree with coefficients in K. Then Γ must contain an element with an eigenvalue t (in some faithful linear representation of $\mathbb{G}$) which is not a root of unity, since otherwise every $\gamma \in \Gamma$ would be killed by the product of a bounded number of cyclotomic polynomials, contradicting the Zariski-density of Γ. Now apply Lemma 6.2 to obtain the desired local field k. The image in of Γ in $\mathbb{G}(k)$ is still Zariski-dense and is not relatively compact. We may then apply Theorem 1.1 to get the desired conclusion. □

References

[1] R. Aoun. Transience of algebraic varieties in linear groups and application to generic Zariski density. *to appear Annales de l'Institut Henri Poincaré, arXiv:1103.0944.*

[2] R. Aoun. Random subgroups of linear groups are free. *Duke Math. J.*, 160(1):117–173, 2011.

[3] Y. Benoist and N. de Saxcé. A spectral gap theorem in simple lie groups. preprint available on the authors website.

[4] Y. Benoist and J.-F. Quint. Central limit theorem for linear groups. preprint available on the authors website, to appear in Annals of Proba.

[5] Y. Benoist and J.-F. Quint. *Random walks on reductive groups.* manuscript available on the authors website.

[6] E. Bombieri and W. Gubler. *Heights in Diophantine geometry*, volume 4 of *New Mathematical Monographs.* Cambridge University Press, Cambridge, 2006.

[7] P. Bougerol and J. Lacroix. *Products of random matrices with applications to Schrödinger operators*, volume 8 of *Progress in Probability and Statistics.* Birkhäuser Boston, Inc., Boston, MA, 1985.

[8] J. Bourgain, A. Furman, E. Lindenstrauss, and S. Mozes. Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus. *J. Amer. Math. Soc.*, 24(1):231–280, 2011.

[9] J. Bourgain and A. Gamburd. Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. *Ann. of Math. (2)*, 167(2):625–642, 2008.

[10] J. Bourgain and A. Gamburd. Expansion and random walks in $\mathrm{SL}_d(\mathbb{Z}/p^n\mathbb{Z})$. II. *J. Eur. Math. Soc. (JEMS)*, 11(5):1057–1103, 2009. With an appendix by Bourgain.

[11] E. Breuillard. Approximate groups and super-strong approximation. survey article to appear in proceedings of the 2013 Groups St.Andrews meeting.

[12] E. Breuillard. Geometry of groups of polynomial growth and shape of large balls. 2007. Preprint.

[13] I. Y. Gol′dsheĭd and G. A. Margulis. Lyapunov exponents of a product of random matrices. *Uspekhi Mat. Nauk*, 44(5(269)):13–60, 1989.

[14] A. S. Golsefidy and P. P. Varjú. Expansion in perfect groups. *Geom. Funct. Anal.*, 22(6):1832–1891, 2012.

[15] F. Guimier. Simplicité du spectre de Liapounoff d'un produit de matrices aléatoires sur un corps ultramétrique. *C. R. Acad. Sci. Paris Sér. I Math.*, 309(15):885–888, 1989.

[16] Y. Guivarc'h. Produits de matrices aléatoires et applications aux propriétés géométriques des sous-groupes du groupe linéaire. *Ergodic Theory Dynam. Systems*, 10(3):483–512, 1990.

[17] Y. Guivarc'h and A. Raugi. Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence. *Z. Wahrsch. Verw. Gebiete*, 69(2):187–242, 1985.

[18] É. Le Page. Théorèmes limites pour les produits de matrices aléatoires. In *Probability measures on groups (Oberwolfach, 1981)*, volume 928 of *Lecture Notes in Math.*, pages 258–303. Springer, Berlin-New York, 1982.

[19] J.-F. Quint. Cônes limites des sous-groupes discrets des groupes réductifs sur un corps local. *Transform. Groups*, 7(3):247–266, 2002.

[20] J. Tits. Free subgroups in linear groups. *J. Algebra*, 20:250–270, 1972.

Laboratoire de Mathématiques, Bâtiment 425, Université Paris Sud 11, 91405 Orsay, FRANCE

E-mail address: emmanuel.breuillard@math.u-psud.fr