Bulletin of Mathematical Sciences
Vol. 9, No. 3 (2019) 1950016 (12 pages)

DOI: 10.1142/S1664360719500164

Linear representations of random groups

Gady Kozma
Department of Mathematics
The Weizmann Institute of Science
Rehovot 76100, Israel
gady.kozma@weizmann.ac.il

Alexander Lubotzky
Einstein Institute of Mathematics
The Hebrew University
Jerusalem 91904, Israel
alex.lubotzky@mail.huji.ac.il

Received 11 October 2018
Accepted 8 April 2019
Published 15 May 2019

Communicated by Efim Zelmanov

We show that for a fixed $k \in \mathbb{N}$, Gromov random groups with any density $d > 0$ have no nontrivial degree k representations over any field, a.a.s. This is especially interesting in light of the results of Agol, Ollivier and Wise that when $d < \frac{1}{6}$ such groups have a faithful linear representation over $\mathbb{Q}$, a.a.s.

Keywords: Random groups; representations; Bezout Theorem.

Mathematics Subject Classification: 20C99, 20F69

1. Introduction

Gromov random groups. Let $m \geq 2$ and let F be the free group on m generators $x = \{x_1, \ldots, x_m\}$. For $l \in \mathbb{N}$, let S_l be the sphere of radius l in the Cayley graph of F with respect to x, i.e. the set of reduced words in $x_i^{\pm 1}$ of length l. Fix some $d \geq 0$ and let R be a random subset of S_l constructed by taking $\lfloor |S_l|^d \rfloor = \lfloor \left(2m(2m-1)^{l-1}\right)^d \rfloor$ elements of S_l uniformly, independently and with repetitions. The group $\Gamma = \langle x \mid R \rangle$, i.e. the group presented by the generators x and the relators R, is called a "Gromov random group of density d with m generators and relators of length l". For a group

property P, we say that Gromov random groups satisfy P asymptotically almost surely (a.a.s.) if the probability of P goes to 1 as $l \to \infty$. In a formula,

$$\lim_{l\to\infty} \mathbb{P}(\Gamma \text{ satisfies } P) = 1.$$

See [4] for an invitation to the topic. The goal of this paper is to prove:

Theorem 1. *Let $k \geq 1$, $m \geq 2$ and $d > 0$. Then Gromov random groups Γ at density d with m generators satisfy a.a.s. that for any field F and any $\rho : \Gamma \to \mathrm{GL}_k(F)$, $|\rho(\Gamma)| \leq 2$.*

In fact, we prove that logarithmically many relators are enough for this property, see the formulation of Theorem 8 below.

When l is odd, it is easy to see that $\mathbb{Z}/2\mathbb{Z}$ is a.a.s. not a quotient of Γ hence in fact we may strengthen Theorem 1 to state that $\rho(\Gamma) = \{1\}$. Similarly, when l is even, $\mathbb{Z}/2\mathbb{Z}$ is (deterministically) a quotient of Γ so the possibility of an image of size 2 cannot be removed.

We recall the well-known result of Gromov [4, §V] that for a fixed $m \geq 2$, Γ is a.a.s. an infinite hyperbolic group for $d < \frac{1}{2}$, while $|\Gamma| \leq 2$ for $d > \frac{1}{2}$. So our theorem is of interest only for $d \leq \frac{1}{2}$. But it is especially interesting for $d < \frac{1}{6}$. In this case, Agol [1] and Ollivier and Wise [5] proved the following remarkable result:

Theorem 2. *For a fixed $m \geq 2$ and $d < \frac{1}{6}$, the random group Γ is a.a.s. linear over $\mathbb{Z}$, i.e. has a faithful representation into $\mathrm{GL}_k(\mathbb{Z})$, for some $k \in \mathbb{N}$.*

Thus the main difference between these two results is whether k, the degree of linearity, is allowed to depend on l, the length of the relators, or not. If it is allowed, we are in the case of Theorem 2 and a representation exists. If it is fixed, we are in the case of Theorem 1 and no representation exists.

A remark on the field: while Theorem 2 constructs a representation into $\mathbb{Q}$, in fact it implies arbitrarily large representations for any field. We cannot show this by taking the representation into $\mathrm{GL}_k(\mathbb{Z})$ modulo p as that might be trivial. But we can, instead, use the fact that any subgroup of $\mathrm{GL}_k(\mathbb{Z})$ is residually finite (simply because $\cap_m \ker(\mathrm{GL}_k(\mathbb{Z}) \to \mathrm{GL}_k(\mathbb{Z}/m\mathbb{Z})) = \{1\}$) so has arbitrarily large finite quotients. These finite quotients may be embedded into a symmetric group, hence for some k' it will embed (as permutation matrices) into $\mathrm{GL}_{k'}(F)$ for any F.

2. Algebraic Geometry Preliminaries

The proof uses some results from algebraic geometry. We will now survey briefly the notions and results we need, assuming only that the reader is familiar with undergraduate algebra.

Let F be an algebraically closed field of any characteristic, and let $n \geq 0$. A subset W of the affine space $\mathbb{A}^n := F^n$ is called an *(affine) variety* if

$$W = \bigcap_{i=1}^{k} Z(p_i) \quad Z(p) = \{x \in \mathbb{A}^n : p(x) = 0\},$$

where $p_1, \ldots, p_k$ are polynomials in n variables. We will use the notations F (for the underlying algebraically closed field) and $Z(p)$ throughout the paper.

A variety is called *irreducible* if it cannot be written as a union of two proper subvarieties. Any variety can be written as a finite union of irreducible varieties. Assuming that the representation is not redundant (i.e. if $W = \bigcup X_i$ then $X_i \nsubseteq X_j$ for any $i \neq j$), it is unique. The X_i of this unique representation are called the *irreducible components* of W. See [6, Theorems 1.4 and 1.5]. Let us remark that in some of the literature, including [3, 6], a variety is defined to be automatically irreducible. But for us it will be convenient to define it as above.

For any affine variety one can define its *dimension*, denoted by dim. Heuristically it corresponds with the natural notion of dimension, but the formal definition requires some preliminaries which we prefer to skip. The reader may consult [6, Chap. 1, §6]. We will need the following properties of it:

(i) $\dim(W) \in \{-1, 0, 1, 2, \ldots\}$.
(ii) $\dim(\mathbb{A}^n) = n$.
(iii) $\dim(W) = -1$ only for $W = \emptyset$ and $\dim(W) = 0$ implies that W is finite.
(iv) If W is an irreducible algebraic variety with $\dim(W) = k$ and if p is a polynomial not identically zero on W, then any irreducible component of $W \cap Z(p)$ has dimension $k - 1$.

See [6, Corollary 1.13] for this last, and most remarkable property. (Note that theorem numbering in the third edition of Shafarevich is different from those of the previous editions).

The following result will be referred to as "Bézout's theorem" (the literature is abound with results called "Bézout's theorem", some of them very close in formulation to it, so we are certainly following tradition here).

Theorem 3 (Bézout). *Let W be an affine variety defined by polynomials $f_1, f_2, \ldots, f_m$ in n variables, i.e. $W = \cap Z(f_i) \subset \mathbb{A}^n$. Suppose $\deg f_i \leq d$ for all i. Then the number of irreducible components of W is bounded by $d^{\min(m,n)}$.*

This result is well-known, even classic. And yet we could not find a reference to it in this form. Hence we supply a proof.

Proof. The literature is far more complete for *projective* varieties. Hence our first step will be to define the projective space $\mathbb{P}^n$ and show how the projective Bézout theorem implies the affine one (we hope no confusion will arise from the use of $\mathbb{P}$ for "probability" in other parts of the paper. The use of $\mathbb{P}$ for the projective space will be restricted to the proof of Theorem 3).

The projective space $\mathbb{P}^n$ over a field F is the space $F^{n+1}\backslash\{0\}$ (we consider the coordinates $0, \ldots, n$), modulo the relation $v \sim av$ for every $v \in F^{n+1}\backslash\{0\}$ and $a \in F\backslash\{0\}$. A projective variety is the intersection of zeroes of *homogeneous* polynomials. Irreducible projective varieties are defined like affine ones, and the decomposition result that allows to define irreducible components is as in the affine

case ([6, p. 46] claims that "the proof carries over word-for-word"). We will need two maps between subvarieties of $\mathbb{A}^n$ and $\mathbb{P}^n$. The first, restriction, takes the projective variety $\cap Z(f_i)$, f_i homogeneous polynomials in $x_0, \ldots, x_n$ and maps it to the affine variety $\cap Z(g_i)$ where $g_i(x_1, \ldots, x_n) = f_i(1, x_1, \ldots, x_n)$. The second, homogenization, maps an affine variety $\cap Z(g_i)$ into the projective variety $\cap Z(f_i)$, where f_i are produced from g_i by taking every monomial $ax_1^{b_1} \cdots x_n^{b_n}$ of g_i and mapping it to $ax_0^{b_0} \cdots x_n^{b_n}$ where $b_0 = \deg g_i - (b_1 + \cdots + b_n)$, and summing those to get f_i. We will denote "W is the restriction of V" by $W = V \cap \mathbb{A}^n$, and "V is the homogenization of W" by $V = \overline{W}$. Clearly $\overline{W} \cap \mathbb{A}^n = W$ for any affine W. (We remark that $\overline{W}$ depends only on W as a set, and not on the choice of the polynomials f_i such that $W = \cap Z(f_i)$, but we will not need this fact).

Claim. *The restriction of an irreducible projective variety is irreducible.*

Proof. Let V be the irreducible projective variety, and let $W = V \cap \mathbb{A}^n$. Assume by contradiction that $W = W_1 \cup W_2$ in a nontrivial way. We now claim that $(\overline{W_1} \cap V) \cup (\overline{W_2} \cap V) \cup (\{x_0 = 0\} \cap V)$ is a nontrivial decomposition of V. Indeed, this is clearly a decomposition of V, and it is nontrivial because any $(x_1, \ldots, x_n) \in W_1 \backslash W_2$ would satisfy that $(1, x_1, \ldots, x_n) \in \overline{W_1} \backslash \overline{W_2}$, and similarly for $W_2 \backslash W_1$. □

With the claim, the affine Bézout theorem follows from the projective one as follows: Let $W = \cap_{i=1}^m Z(f_i)$ with $\deg f_i \leq d$. Then $\overline{W}$ has the same structure, and hence by the projective Bézout theorem its decomposition to irreducible components $\overline{W} = X_1 \cup \cdots \cup X_K$ satisfies $K \leq d^{\min(n,m)}$. By the claim, $X_i \cap \mathbb{A}^n$ are irreducible, and of course

$$W = \overline{W} \cap \mathbb{A}^n = (X_1 \cup \cdots \cup X_K) \cap \mathbb{A}^n = (X_1 \cap \mathbb{A}^n) \cup \cdots \cup (X_K \cap \mathbb{A}^n),$$

which is a decomposition of W to irreducible subvarieties (it might be redundant, but that would only mean the number of components of W is smaller than K). Thus, we need only show the projective Bézout theorem.

For the projective Bézout theorem we will need the concepts of dimension and degree of a projective variety. The dimension of a projective variety is as for an affine variety, and has the same four properties listed above ($\dim(\mathbb{P}^n) = n$), with the same references in [6]. As for the degree, heuristically if $W \subseteq \mathbb{P}^n$ is some irreducible variety then $\deg W$ is the number of intersections of W with a generic linear variety of dimension $n - \dim W$. Again, the formal definition is different and we will skip it, the reader may consult [3, p. 50]. We only need the following properties to use Theorem 4 below:

(i) $\deg(W)$ is always a positive integer, except $\deg(\emptyset) = 0$,
(ii) $\deg(\mathbb{P}^n) = 1$.

For both properties, see [3, Chap. 1, Proposition 7.6]. The projective Bézout theorem follows as a corollary from the following result:

Theorem 4. *Let W be an irreducible projective variety. Let f be a homogeneous polynomial. Let $X_1, \dots, X_s$ be the irreducible components of $W \cap Z(f)$. Then*

$$\sum_{j=1}^{s} \deg(X_j) \leq \deg W \cdot \deg f,$$

where $\deg W$ is the degree of a projective variety just mentioned, while $\deg f$ is the usual degree of a polynomial. See [3, Theorem 7.7 and Proposition 7.6d]. The formulation in [3] has some additional quantities, intersection multiplicities, denoted by $i(\cdot)$ — all we need from them is that they are at least 1, which follows because they are defined as lengths of some modules ([3], top of p. 53 and the definition at p. 51), and the length of a module is the maximal size of a decreasing sequence of submodules. The formulation in [3] requires that $\dim W \geq 1$ and that f is not identically zero on W, but the case $\dim W = 0$ (i.e. W is a single point) is obvious, and so is the case $W \subseteq Z(f)$.

Recall the formulation of the projective Bézout theorem: we are given homogeneous polynomials $f_1, \dots, f_m$ in $n+1$ variables with $\deg f_i \leq d$. Let

$$\bigcap_{i=1}^{m} Z(f_i) = X_1 \cup \cdots \cup X_K \tag{1}$$

be the decomposition of $\cap Z(f_i)$ into irreducible components. We claim that

$$\sum_{j=1}^{K} \deg(X_j) d^{\dim X_j} \leq d^n. \tag{2}$$

We show (2) by induction on m. Indeed, $m = 0$ is obvious. Assume (2) has been proved for m and write (using the X_i of (1))

$$\bigcap_{i=1}^{m+1} Z(f_i) = \bigcup_{j=1}^{K} X_j \cap Z(f_{m+1}).$$

Fix j and let $Y_{j,k}$ be the irreducible components of $X_j \cap Z(f_{m+1})$. By Theorem 4,

$$\sum_{k} \deg(Y_{j,k}) \leq d \deg(X_j).$$

If $X_j \nsubseteq Z(f_{m+1})$ then by property 4 of the dimension, $\dim Y_{j,k} = \dim X_j - 1$ so

$$\sum_{k} \deg(Y_{j,k}) d^{\dim Y_{j,k}} \leq \deg(X_j) d^{\dim X_j}. \tag{3}$$

But if $X_j \subseteq Z(f_{m+1})$ then (3) holds trivially (with no need to invoke Theorem 4). So (3) holds always. We sum (3) over j to get

$$\sum_{j,k} \deg(Y_{j,k}) d^{\dim Y_{j,k}} \leq \sum_{j} \deg(X_j) d^{\dim X_j} \leq d^n,$$

where the second inequality is the induction assumption. Now, $Y_{j,k}$ is a decomposition of $\cap_{i=1}^{m+1} Z(f_i)$ to irreducible components — it may be redundant, but that

only reduces the sum in (2) further. Hence (2) holds for $m+1$ and the induction is complete.

The projective Bézout theorem now follows easily. We drop the degrees (as we may, as they are always at least 1) and get

$$\sum_{j=1}^{K} d^{\dim X_j} \leq d^n.$$

If $m \geq n$ the theorem follows immediately. If $m < n$ it follows because then each X_j has dimension at least $n-m$. This shows the projective Bézout theorem and, as explained above, also the affine one, namely Theorem 3. □

The next result we need is an effective version of the nullstellensatz. Hilbert's nullstellensatz is the following: Suppose p_i are polynomials in n variables with $\cap Z(p_i) = \emptyset$. Then there exists polynomials q_i such that $\sum p_i q_i \equiv 1$. There is also a version of the nullstellensatz when $W := \cap Z(p_i) \neq \emptyset$. It states that if r is a polynomial which is zero on every point of W, then there exists a $\nu \geq 1$ and q_i such that $\sum p_i q_i = r^\nu$. These theorems hold in an algebraically closed field, but if the coefficients of the p_i and of r belong to a smaller field then the coefficients of the q_i can be chosen in the same smaller field since the condition $\sum p_i q_i = r^\nu$ is a linear condition on the coefficients of the q_i. Let us consider the case that the smaller field is $\mathbb{Q}$ (below the notation Z will refer to zeroes over $\mathbb{C}$). Multiplying by the common denominator we get a result that holds in $\mathbb{Z}$, i.e. if the p_i and the r have integer coefficients then one may find polynomials q_i with integer coefficients, and integers ν and b such that $\sum p_i q_i = b r^\nu$. We will need an effective version of this result but, in fact, the only quantity we need to control is b. Hence the effective version is as follows:

Theorem 5. *Let $p_1, \ldots, p_t, r \in \mathbb{Z}[x_1, \ldots, x_n]$ and assume r vanishes on $\cap_{i=1}^t Z(p_i)$. Assume also that $\deg p_i \leq d \ \forall i$, $\deg r \leq d$ and all coefficients of all p_i are bounded in absolute value by h. Then there exists $q_i \in \mathbb{Z}[x_1, \ldots, x_n]$, $i = 1, \ldots, t$ and $b, \nu \in \mathbb{N}$ such that*

$$\sum_{i=1}^{t} p_i q_i = b r^\nu$$

with the bound

$$\log b \leq C^n n^{2n} (d+1)^{n(n+2)} (\log h + C n^2 \log d).$$

Further, b is independent of r.

Here and below C and c will stand for absolute constants whose value might change from line to line. We will only use the following, rough bound, which holds for a fixed n and d sufficiently large (i.e. $d > d_0(n)$),

$$\log b \leq (2d)^{n(n+2)+1} \log h. \tag{4}$$

Proof. We will find $q_i \in \mathbb{Q}[x_1, \ldots, x_n]$ such that $\sum p_i q_i = r^\nu$ and then b will be bounded by the lcm of the denominators of the q_i. By the corollary to Theorem 1 of [2], we may take $q_i \in \mathbb{Q}[x_1, \ldots, x_n]$ with

$$\deg q_i \leq (n+1)(n+2)(d+1)^{n+2} =: Q. \tag{5}$$

Once the degree is bounded, the coefficients of the q_i are given by the solution of a system of linear equations (depending on the p_i, on r and on ν). Let $f(d,n)$ be the dimension of the space of polynomials with n variables and degree $\leq d$ (so $f(d,n) \leq (d+1)^n$). The system might be underdetermined, we have $tf(Q,n)$ variables and at most $f(Q+d,n)$ equations, one for each coefficient of one monomial in the equality $\sum p_i q_i = r^\nu$, up to the degree of the left-hand side. Let R be the rank of this system of equations, so $R \leq f(Q+d,n)$. Pick arbitrarily R variables and R equations such that the corresponding submatrix M is invertible and solve the restricted equations. Set the rest of the variables to zero, and the remaining equations (if any) will be fulfilled automatically. It follows that some choice of the q_i can be achieved by inverting M and applying the result to the vector of the coefficients of r^ν, themselves integers. Since $M^{-1} = M'/\det M$, where M' has integer entries, we may choose $b = \det M$ (and in particular b is independent of r). But the entries of M are simply the coefficients of the p_i, all of them bounded by h. Applying Hadamard's inequality (the determinant is bounded by the product of the l^2 norms of the rows) gives

$$b \leq (h\sqrt{R})^R$$

or

$$\begin{aligned}
\log b &\leq R\left(\log h + \frac{1}{2}\log R\right) \leq f(Q+d,n)\left(\log h + \frac{1}{2}\log f(Q+d,n)\right) \\
&\leq (Q+d+1)^n\left(\log h + \frac{1}{2}n\log(Q+d+1)\right) \\
&\overset{(5)}{\leq} ((n+1)(n+2)(d+1)^{n+2} + d + 1)^n \\
&\quad \cdot \left(\log h + \frac{1}{2}n\log((n+1)(n+2)(d+1)^{n+2} + d + 1)\right) \\
&\leq C^n n^{2n}(d+1)^{n(n+2)}(\log h + Cn^2 \log d)
\end{aligned}$$

as claimed. □

3. Proof of the Main Result

Lemma 6. *Let G be any d-regular connected multigraph with $d \geq 4$ and more than 2 vertices, and let x be some vertex of G. Let $t > 1$. Then*

$$\mathbb{P}^x(N(t) = x) \leq \frac{d-2}{d-1},$$

where $N(t)$ is a nonbacktracking random walk on G at the t^{th} step and $\mathbb{P}^x$ denotes the probability when $N(0) = x$.

Let us define precisely what we mean by "multigraph" and "nonbacktracking random walk". A multigraph is a graph which might contain multiple edges and self-loops. It is d-regular if every vertex has exactly d edges connected to it, with a self-loop counted as two edges. A nonbacktracking random walk is a walk that is not allowed to traverse an edge and on the next step traverse it in the opposite direction (there are no restrictions on the first step). A self-loop can be traversed in either direction, and the nonbacktracking condition is that it cannot be traversed and then traversed backwards. When the multigraph is d-regular, this process has exactly $d-1$ possibilities at each step (except the first one), and it chooses each with probability $1/(d-1)$, independently of the past.

Proof. Fix the vertex x for the rest of the proof. Every edge of our multigraph we consider as two directed edges (a self-loop too corresponds to two directed edges), and for a directed edge e we denote by $\overline{e}$ the inverted edge. Hence, the nonbacktracking condition is that the walk is not allowed to traverse e immediately after traversing $\overline{e}$. (Note that we have a multigraph, so there can be $e \neq f$ that both go from vertex x to vertex y. Still, we may traverse $\overline{e}$ and then f, or $\overline{f}$ and then e. It is only the couples $\overline{e}$, e and $\overline{f}$, f that are prohibited. Each self-loop corresponds to two directed edges which are $\bar{\cdot}$ of one another). Let $q_t(e)$ be the probability that the edge e was traversed at time t, i.e. if $e : v \to w$ (i.e. e is from v to w, we will also use the notation $e :\to v$ and $e : v \to$ if we do not care about the other vertex) then it is the probability that $N(t-1) = v$ and then the process continues through e (which means, in particular, that $N(t) = w$). Let $Q(t) = \max_e q_t(e)$. Then $Q(t)$ is non-increasing because

$$q_{t+1}(e) = \frac{1}{d-1} \sum_{f :\to v,\, f \neq \overline{e}} q_t(f) \leq Q(t) \quad \forall\, e : v \to .$$

Examine now the event that $e : y \to z$ was traversed in the second step. It requires that $N(1) = y$. Assume first (call this "case I") that each neighbor y of x is connected to x by $\leq d-2$ edges (including x itself, if there are self-loops). Then $\mathbb{P}(N(1) = y) \leq (d-2)/d$ for every y and hence $Q(2) \leq (d-2)/(d(d-1))$.

If x has a neighbor y to which it is connected by more than $d-2$ edges ("case II"), then the requirements of regularity and more than 2 vertices say that it must be connected to y by exactly $d-1$ edges, and further that it has a second neighbor to which it is connected by 1 edge. This means that $\mathbb{P}(N(2) = x) = (d-2)/d$ and, in particular, any vertex $z \neq x$, we have $\mathbb{P}(N(2) = z) \leq 2/d$. Hence

$$Q(3) \leq \frac{\max\{2, (d-2)\}}{d(d-1)}.$$

Since $d \geq 4$ we get the same bound as in case I. Hence $Q(t) \leq (d-2)/(d(d-1))$ for all $t \geq 3$. But this means that

$$\mathbb{P}(N(t) = x) = \sum_{e:\to x} q_t(e) \leq dQ(t) \leq \frac{d-2}{d-1}.$$

This covers all cases of the lemma except $t = 2$ in case II, but we just calculated that in this case $\mathbb{P}(N(2) = x) = (d-2)/d$. The lemma is thus proved. □

The next lemma is quite close to the formulation of Theorem 8, the only difference is that it handles only one field.

Lemma 7. *Let F be an algebraically closed field, let $k, m \in \mathbb{N}$, $m \geq 2$ and let l be sufficiently large (i.e. $l > l_0(k,m)$). Let R be given by taking $u \geq 15\, m^3 k^4 \log l$ random reduced words of length l in the letters $\{x_1, \ldots, x_m, x_1^{-1}, \ldots, x_m^{-1}\}$ independently, uniformly, with repetitions. Let $\Gamma = \langle x_1, \ldots, x_m \,|\, R\rangle$. Then*

$$\mathbb{P}(\exists\, \rho : \Gamma \to \mathrm{GL}_k(F) \text{ such that } |\rho(\Gamma)| > 2) \leq \exp(-cu/m^2k^2),$$

where ρ is a group homomorphism.

(log here is the natural logarithm).

Proof. We consider $\mathrm{GL}_k(F)$ as a subvariety of F^{2k^2} by considering the first set of k^2 variables as the entries of the matrix and the second set of k^2 variables as the entries of the inverse matrix, and adding polynomial equations (k^2 of them, all of degree 2) that ensure that indeed, the product of the two matrices is 1. Similarly we consider $\mathrm{GL}_k(F) \times \cdots \times \mathrm{GL}_k(F)$ (m times) as a subvariety of F^{2mk^2}. Denote this variety by X. For $A \in \mathrm{GL}_k(F) \times \cdots \times \mathrm{GL}_k(F)$ we denote $(A, A^{-1}) := (A_1, A_1^{-1}, \ldots, A_m, A_m^{-1}) \in X$.

Let $E_j \subset X$ be the collection of (A, A^{-1}) such that the matrices $A_1, \ldots, A_m$ satisfy the first j words in R. Since these (random) words can be thought of as (random) polynomial equations in $2\,mk^2$ variables as above, E_j is a (random) variety in F^{2mk^2}. Let (A, A^{-1}) be a point in E_j with $|\langle A\rangle| > 2$. Examine the event that $(A, A^{-1}) \in E_{j+1}$, conditioned on E_j. The new reduced word ω that was added to form E_{j+1} is independent of the past, and hence $\omega(A)$ is distributed like a nonbacktracking random walk of length l on the Cayley graph generated by A. (If, for some i, $A_i = 1$ then the graph will contain one corresponding self-loop on each vertex. The two directions of this self-loop will correspond to multiplying by A_i and A_i^{-1}. This matches with the definitions we gave around Lemma 6). This Cayley graph is a $2m$-regular multigraph, and by assumption it has more than 2 vertices. Hence, we may use Lemma 6 to get

$$\mathbb{P}((A, A^{-1}) \in E_{j+1} \,|\, (A, A^{-1}) \in E_j) \leq \frac{2m-2}{2m-1}.$$

In other words, with probability $\geq \frac{1}{2m-1}$, adding a relation breaks the irreducible component containing (A, A^{-1}) into further irreducible components, which then must have smaller dimension, by property (iv) of the dimension (see §2).

Repeating this λ times we get that after adding λ words we break any fixed irreducible component with probability at least $1-\exp(-\lambda/2m)$. By Bézout's theorem (Theorem 3), E_j has no more than l^{2mk^2} irreducible components (the initial polynomial equations defining X have degree 2). Hence a simple union bound shows that, for $\lambda \geq 5m^2k^2 \log l$,

$$\mathbb{P}(\text{all components are broken}) \geq 1 - l^{2mk^2} \exp\left(-\frac{\lambda}{2m}\right) \geq 1 - \exp\left(-\frac{\lambda}{10m}\right).$$

Use that for $\lambda = u/(2\,mk^2+1) \geq 5\,m^2k^2 \log l$ words, and get that with probability at least $1-\exp(-cu/m^2k^2)$ one breaks all components. Therefore the maximal degree decreases by 1. Repeating this a further $2\,mk^2+1$ times, the maximal degree of any component which contains any (A, A^{-1}) with $|\langle A\rangle| > 2$ is -1, so they are in fact empty. We get the claim of the lemma with probability $(2\,mk^2+1)\exp(-cu/m^2k^2)$ but of course the outer $2\,mk^2+1$ can be ignored (perhaps changing the constant inside the exponent). □

Theorem 8. *Let $k, l, m \in \mathbb{N}$, $m \geq 2$. Let R be given by taking at least $Cm^4k^6 \log l$ random reduced words of length l in the letters $\{x_1, \ldots, x_m, x_1^{-1}, \ldots, x_m^{-1}\}$ independently, uniformly, with repetitions. Let $\Gamma = \langle x_1, \ldots, x_m | R\rangle$. Then*

$$\lim_{l\to\infty} \mathbb{P}(\exists\, F, \rho : \Gamma \to \mathrm{GL}_k(F) \text{ such that } |\rho(\Gamma)| > 2) = 0,$$

where F runs over the all fields, and where ρ is a group homomorphism.

Proof. First apply Lemma 7 with $F = \mathbb{C}$ and with $15\,m^3k^4\lceil \log l\rceil$ relators. We get that with high probability, any $\rho : \Gamma \to \mathrm{GL}_k(\mathbb{C})$ has $|\rho(\Gamma)| \leq 2$. Of course, the image of the generators $\{x_1, \ldots, x_m\}$ is easy to characterize: we have $\rho(x_i)^2 = 1$ $\forall\, i$ and for some $S \subset \{1, \ldots, m\}$ we have $\rho(x_i) = \rho(x_j)$ $\forall\, i, j \in S$ and $\rho(x_i) = 1$ $\forall\, i \notin S$. Compactly, $|\{\rho(x_i)\}\backslash\{1\}| \leq 1$.

Recall from the proof of Lemma 7 the variety X in $\mathbb{C}^{2mk^2}$ and denote by $p_1, \ldots,$ p_{mk^2} the polynomials defining it; and the notation (A, A^{-1}). The condition that the matrices A satisfy a given word is a collection of k^2 polynomials in $2\,mk^2$ variables. Denote the polynomials that the polynomials that correspond to the ith word by $p_{(m+(i-1))k^2+1}, \ldots, p_{(m+i)k^2}$ and let $M = mk^2 + 15\,m^3k^6\lceil \log l\rceil$. Note that each of these polynomials has integer coefficients. We get that

$$\bigcap_{i=1}^{M} Z(p_i) \subseteq \{(A, A^{-1}) : A_i^2 = 1\ \forall\, i, |A\backslash\{1\}| \leq 1\}.$$

Denote the variety on the right by Y.

We now claim that Y can by written as $\cap Z(r_j)$ for some $r_1, \ldots, r_K$ which depend only on k and m, and in particular do not depend on the field. Here is how: the condition $A_i^2 = 1$ corresponds to k^2 polynomials for each i. The condition

$$\bigcup_{S\subset\{1,\ldots,m\}} \{A_i = A_j \forall\, i, j \in S, A_i = 1 \forall\, i \notin S\}.$$

gives us at most $(mk^2)^{2^m}$ polynomials because for every S the corresponding variety is described by at most mk^2 polynomials, but taking union requires to take every possible choice of a polynomial for each S, and multiply them out. This describes our $r_1, \ldots, r_K$ (and gives $K \leq mk^2 + (mk^2)^{2^m}$, but we will have no use for this fact).

Now apply the effective nullstellensatz (Theorem 5) K times as follows. In all applications the polynomials p_i from the nullstellensatz are our p_i, but the polynomial r we take corresponding to the r_j above. We get corresponding $q_{i,j}$, ν_j and b (we use here the "further" clause of Theorem 5 which claims that b is independent of r, so our b is independent of j), with b bounded by (4). The number of variables is $2\,mk^2$ while the maximal value of the coefficients, h, can be bounded roughly by $(2\,mk^2)^l$. We get,

$$b \leq \exp((2l)^{4m^2k^4+4mk^2+1} \cdot l \log(2\,mk^2)) \leq \exp((2l)^{7m^2k^4}),$$

which holds for l sufficiently large.

Consider now a field F of characteristic that does not divide b. Then

$$\sum p_i q_{i,j} = b r_j^{\nu_j}$$

holds also in F, and because $\operatorname{char} F$ does not divide b we get that $b \neq 0$ in the field and we may divide by it. This means that whenever $p_i = 0$ for all i so are r_j for all j, but that means that any A which satisfy our first $15\,m^3k^4\lceil \log l \rceil$ words must also satisfy that $A_i^2 = 1\ \forall i$ and that $|A\backslash\{1\}| \leq 1$. So in $\mathrm{GL}_k(F)$ too we get $|\langle A \rangle| \leq 2$.

Finally, let $\mathscr{P}$ be the set of primes that divide b and note that $|\mathscr{P}| \leq C \log b \leq C(2l)^{7m^2k^4}$. For every $\tau \in \mathscr{P}$ apply Lemma 7 again, but this time with the field F_τ being the algebraic closure of $\mathbb{Z}/\tau\mathbb{Z}$, with $u = \lambda m^4 k^6 \log l$ for some λ to be fixed soon, and with the set of relators $R_{\text{Lemma 7}}$ not containing the first $15\,m^3k^4\lceil \log l \rceil$ relations already used. We get that

$$\mathbb{P}(\exists\, \rho : \Gamma \to \mathrm{GL}_k(F_\tau) \text{ s.t. } |\rho(\Gamma)| > 2 \,|\, \mathscr{P}) \leq \exp(-cu/m^2k^2) = \exp(-c\lambda m^2 k^4 \log l),$$

(the conditioning over $\mathscr{P}$, itself a random variable, just like b is, is not important because $\mathscr{P}$ depends on the first $15\,m^3k^4\lceil \log l \rceil$ relations, while the estimate above used the other relations). Summing over all $\tau \in \mathscr{P}$ gives

$$\begin{aligned}\mathbb{P}(\exists\, \tau \in \mathscr{P} \exists\, \rho : \Gamma \to \mathrm{GL}_k(F_\tau) \text{ such that } |\rho(\Gamma)| > 2 \,|\, \mathscr{P}) \\ \leq |\mathscr{P}| \exp(-c\lambda m^2 k^4 \log l) \leq \exp((C - c\lambda) m^2 k^4 \log l).\end{aligned}$$

Integrating over the conditioning gives

$$\mathbb{P}(\exists\, \tau \in \mathscr{P} \exists\, \rho : \Gamma \to \mathrm{GL}_k(F_\tau) \text{ such that } |\rho(\Gamma)| > 2) \leq \exp((C - c\lambda) m^2 k^4 \log l).$$

Taking $\lambda = 2C/c$ we get that this probability goes to zero. Moving from F_τ to a general field of characteristic τ is done using the (usual, non-effective) nullstellensatz: find polynomials $q_{i,j} \in \mathbb{Z}/\tau\mathbb{Z}[x_1, \ldots, x_{2mk^2}]$ such that $\sum p_i q_{i,j} = r_j^{\nu_j}$ in $\mathbb{Z}/\tau\mathbb{Z}$ with the same p_i and $r_1, \ldots, r_K$ as above, and note that the existence of these $q_{i,j}$ ensures that in any field F of characteristic τ, if $A_1, \ldots, A_m$ are in $\mathrm{GL}_k(F)$ and satisfy all words in R then $|\langle A \rangle| \leq 2$, proving the theorem. □

Acknowledgments

We thank Nir Avni for the idea to use the effective nullstellensatz. We thank Ron Livne for help with Bézout's theorem. We thank Tsachik Gelander, Shahar Mozes and Michael Ben Or for many interesting discussions.

G. Kozma was supported by the Israel Science Foundation, by the Jesselson Foundation and by Paul and Tina Gardner. A. Lubotzky was supported by the Israel Science Foundation, by the National Science Foundation and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 692854).

References

[1] I. Agol, The virtual Haken conjecture, *Doc. Math.* **18** (2013) 1045–1087. With an appendix by Ian Agol, Daniel Groves, and Jason Manning.
[2] W. D. Brownawell, Bounds for the degrees in the Nullstellensatz. *Ann. of Math.* **126**(3) (1987) 577–591.
[3] R. Hartshorne, *Algebraic Geometry*, Graduate Texts in Mathematics, Vol. 52 (Springer-Verlag, New York-Heidelberg, 1977).
[4] Y. Ollivier, *A January 2005 Invitation to Random Groups*, Ensaios Matemáticos [Mathematical Surveys], Vol. 10 (Sociedade Brasileira de Matemática, Rio de Janeiro, 2005).
[5] Y. Ollivier and D. T. Wise, Cubulating random groups at density less than 1/6, *Trans. Amer. Math. Soc.* **363**(9) (2011) 4701–4733.
[6] I. R. Shafarevich, *Basic Algebraic Geometry. 1. Varieties in Projective Space*. 3rd edn. Translated from the 2007 third Russian edition. (Springer, Heidelberg, 2013).