Fernando Chamizo; Dulcinea Raboso
Distributional properties of powers of matrices

Czechoslovak Mathematical Journal, Vol. 64 (2014), No. 3, 801–817

Persistent URL: http://dml.cz/dmlcz/144059

Czechoslovak Mathematical Journal, 64 (139) (2014), 801–817

DISTRIBUTIONAL PROPERTIES OF POWERS OF MATRICES

Fernando Chamizo, Dulcinea Raboso, Madrid

(Received July 5, 2013)

Abstract. We apply the larger sieve to bound the number of 2×2 matrices not having large order when reduced modulo the primes in an interval. Our motivation is the relation with linear recursive congruential generators. Basically our results establish that the probability of finding a matrix with large order modulo many primes drops drastically when a certain threshold involving the number of primes and the order is exceeded. We also study, for a given prime and a matrix, the existence of nearby non-similar matrices having large order. In this direction we find matrices of large order when the trace is restricted to take values in a short interval.

Keywords: larger sieve; pseudorandom number; finite field; special linear group of degree 2; general linear group of degree 2

MSC 2010: 11N36, 11C20, 11Z05, 11L05

1. Introduction and main results

The function $k \mapsto g^k \pmod p$ with g a generator of $\mathbb{F}_p^*$ is employed in practice for pseudorandom number generation. In general, the implementation of linear recursive congruential generators [16] suggests that one should look for matrices in $\mathrm{GL}_n(\mathbb{F}_p)$ having maximal order and there is some literature about the choice of these matrices and the statistical properties of the corresponding generators [5], [22].

In computer science applications usually there is a single built-in pseudorandom number generator function whose output is reduced modulo m to obtain a pseudorandom number in the range $[1, m)$. These ranges appear very often in run-time and it is impossible to choose in advance a common high order element for all of the corresponding moduli. From the mathematical point of view one expects that using $k \mapsto n^k \pmod p$ as a pseudorandom number generator, for p in a reasonably large

The first author is partially supported by the grant MTM2011-22851 from the Ministerio de Ciencia e Innovación (Spain).

range, gives good results for almost any choice of n. In other words, if $\exp_p(n)$ is defined to be the order of n in $\mathbb{F}_p^*$ if $p \nmid n$ and 0 if $p \mid n$, then it is very unlikely to find n such that $\exp_p(n)$ is small for many consecutive primes. This fact was proved by P. X. Gallagher as an application of his larger sieve.

Theorem 1.1 ([7], Theorem 2)**.** *Given $\varepsilon > 0$ the number of integers $n \leqslant N$ for which $\exp_p(n) \leqslant N^\theta$ for all primes $p \leqslant N^{\theta+\varepsilon}$ is $O(N^\theta)$, uniformly for $0 \leqslant \theta \leqslant 1$.*

In connection with this result, P. J. Stephens proved previously that Artin's conjecture holds on average and gave a nontrivial bound for the number of possible exceptions [25]. In practice there is no difference between maximal and large order elements when generating pseudorandom numbers.

Although linear recursive congruential generators have been employed since the 80's, it seems that Artin's conjecture in $\mathrm{GL}_n(\mathbb{F}_p)$ has not received much attention until recently. The case $n = 2$ seems to be a distinguished one. It was shown in [15] and [14] (see also [13]) to be related with quantum ergodicity on flat tori (an instance of arithmetic quantum chaos). In [23] it is also studied in connection with the order of the reduction of units in quadratic fields. On the other hand, our knowledge about the distribution of maximal order matrices in $\mathrm{GL}_2(\mathbb{F}_p)$ benefits from the recent uniform proof [3] of Burgess' inequality in $\mathbb{F}_{p^2}$ and a conjectural deterministic polynomial-time search procedure [24] for primitive roots in $\mathbb{F}_{p^2}$ (meaning that the output is a subset containing at least one primitive root).

Given $N \in \mathbb{Z}^+$ and an interval $I = [1, M]$, consider the probability $P_N(x)$ of a positive integer $n \leqslant N$ having exponent at most x for all primes in I. Of course, if $x \geqslant |I|$, we trivially have $P_N(x) = 1$. On the other hand, Theorem 1.1 implies that if x is slightly smaller than $|I|$ then this probability drops drastically. Namely, Theorem 1.1 can be rephrased as

$$P_N(x) \ll \frac{|I|}{N^{1+\varepsilon}} \quad \text{whenever} \quad \frac{x}{|I|} < N^{-\varepsilon}.$$

In some way, $N^{-\varepsilon}$ establishes a threshold to get a saving $O(|I|N^{-1-\varepsilon})$ with respect to the trivial bound $P_N(x) \leqslant 1$.

In this paper we study this phenomenon for nonsingular integral matrices, showing that there is a value of x very close to the size of the interval such that there are few matrices with order less than x. Furthermore, in the last section we study some properties of high-order elements.

We extend the previous notation writing $\exp_p(A)$ to denote the order of the matrix A in $\mathrm{GL}_2(\mathbb{F}_p)$ when reduced modulo p if $p \nmid \det(A)$ and $\exp_p(A) = 0$ if $p \mid \det(A)$.

We allow thin intervals of primes if they have positive density and are wide enough. Namely, we consider intervals $I = [a, b]$, $0 < a < b - 3$, such that

$$\sum_{p \in I} \log p \gg |I| \quad \text{and} \quad \log |I| \gg \log b \tag{1.1}$$

when p runs over the primes. The prime number theorem implies that this is the case for $I = [1, x]$ in a stronger asymptotic form that extends to $I = [x - x^\alpha, x]$ for $\alpha > 7/12$ using the unconditionally known density hypothesis [10]. In [2] (see also [9]) sieve methods are pushed to prove (1.1) for $\alpha \geqslant 0.525$. With the present knowledge $\log |I| \gg \log b$ holds in every case in which the positive density condition is known [18].

The natural analog of the interval $[0, N]$ in $\mathrm{SL}_2(\mathbb{Z})$ is the set (of cardinality comparable to N^2, see Lemma 2.8)

$$\mathcal{I}_N = \{A \in \mathrm{SL}_2(\mathbb{Z})\colon\ 0 \leqslant a_{ij} \leqslant N\}.$$

We define the probability

$$\mathcal{P}_N(x) = \frac{|\mathcal{M}_N(x)|}{|\mathcal{I}_N|} \quad \text{where } \mathcal{M}_N(x) = \{A \in \mathcal{I}_N\colon\ \exp_p(A) \leqslant x \text{ for } p \in I\}$$

and we want to find a threshold function $\mathfrak{T} = \mathfrak{T}(N, |I|)$ and a saving function $\mathfrak{S} = \mathfrak{S}(N, |I|)$ such that

$$\mathcal{P}_N(x) \leqslant \mathfrak{S} \quad \text{whenever} \quad \frac{x}{|I|} \leqslant \mathfrak{T}. \tag{1.2}$$

In the same way, we also consider arbitrary nonsingular integral matrices. We introduce (see Lemma 2.8)

$$\mathcal{I}_N^* = \{A \in \mathrm{M}_{2\times 2}(\mathbb{Z})\colon\ \det(A) \neq 0,\ 0 \leqslant a_{ij} \leqslant N\},$$

and define

$$\mathcal{P}_N^*(x) = \frac{|\mathcal{M}_N^*(x)|}{|\mathcal{I}_N^*|} \quad \text{where } \mathcal{M}_N^*(x) = \{A \in \mathcal{I}_N^*\colon\ 0 < \exp_p(A) \leqslant x \text{ for } p \in I\}.$$

Again we look for a threshold function $\mathfrak{T}^* = \mathfrak{T}^*(N, |I|)$ and a saving function $\mathfrak{S}^* = \mathfrak{S}^*(N, |I|)$ such that

$$\mathcal{P}_N^*(x) \leqslant \mathfrak{S}^* \quad \text{whenever } \frac{x}{|I|} \leqslant \mathfrak{T}^*. \tag{1.3}$$

Our results prove that a logarithmic threshold is enough to get a substantial saving.

Theorem 1.2. *Let I be an interval satisfying* (1.1) *and $N \geqslant 3$. Then there exists an absolute constant $C > 0$ such that* (1.2) *holds with*

$$\mathfrak{S} = \frac{|I|}{N} \log N (\log\log N)^2 \quad \textit{and} \quad \mathfrak{T} = L(\log N \log |I|)$$

where $L(t) = Ct^{-1} \log t$.

Theorem 1.3. *With the notation of Theorem* 1.2, (1.3) *holds with*

$$\mathfrak{S}^* = \frac{|I|^2 (\log\log N)^2}{N^3 \log\log |I|} \quad \textit{and} \quad \mathfrak{T}^* = L(\log N \log |I| \log\log |I|).$$

The meaning of these results is easier to appreciate when $|I|$ is expressed as a power of N.

Corollary 1.4. *If $|I| = N^\delta \geqslant 3$, then the number of matrices in $\mathcal{I}_N$ for which*

$$\exp_p(A) \leqslant CN^\delta \frac{\log\log N}{\delta (\log N)^2}$$

for every $p \in I$ is less than $N^{\delta+1} \log N (\log\log N)^2$.

Corollary 1.5. *If $|I| = N^\delta \geqslant 3$, then the number of matrices in $\mathcal{I}_N^*$ for which*

$$0 < \exp_p(A) \leqslant CN^\delta \frac{\log\log N}{\delta (\log N)^2 \log\log N^\delta}$$

for every $p \in I$ is less than $N^{2\delta+1} (\log\log N)^2 (\log\log N^\delta)^{-1}$.

These results suggest that it is very unlikely to find a matrix with low exponent for many primes. Keeping the analogy with the integral case, we have many possibilities for good pseudorandom matrix generators.

2. Auxiliary results

Given $m \in \mathbb{Z} - \{0\}$ and an odd prime $p \nmid m$, we define $f(n)$ to be the number of distinct possible Jordan canonical forms (over $\mathbb{F}_{p^2}$) of the diagonalizable matrices belonging to the set

$$\{A \in \mathrm{GL}_2(\mathbb{F}_p) \colon \det A = m,\ \exp_p(A) = n\}.$$

We write henceforth $e = \exp_p(m)$. As the determinant is multiplicative, we have trivially that $f(n) = 0$ if $e \nmid n$. The following lemma takes care of the rest of the cases.

Lemma 2.1. *For $e \mid n$ write $k = n/e$. Then*

$$f(n) = \begin{cases} \varphi(k)e & \text{if } n \mid p-1, \\ \dfrac{\varphi(n)}{\varphi(e)} & \text{if } n \nmid p-1 \text{ and } k \mid p+1 \text{ and } e \text{ or } (p+1)/k \text{ is odd}, \\ 0 & \text{otherwise.} \end{cases}$$

P r o o f. First suppose that the Jordan canonical form of the matrix is of the form $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & \beta \end{smallmatrix}\right)$ with $\alpha, \beta \in \mathbb{F}_p^*$. Then we can write $\beta = m\alpha^{-1}$. Clearly $n = \operatorname{lcm}(\exp_p(\alpha), \exp_p(m\alpha^{-1}))$, and furthermore

$$\exp_p(m\alpha^{-1}) \mid \operatorname{lcm}(e, \exp_p(\alpha^{-1})),$$

which is a general fact of abelian groups, whence $n = \operatorname{lcm}(e, \exp_p(\alpha))$.

For a given $a \mid p-1$ there are $\varphi(a)$ elements in $\mathbb{F}_p^*$ having order a, and there are $F(n,e)$ of them which give matrices of order n, where

$$F(n,e) = \sum_{a\colon \operatorname{lcm}(a,e)=n} \varphi(a).$$

It is easy to see that $F(q^{r+s}, q^r) = \varphi(q^s)q^r$ for q prime. As φ is multiplicative, denoting by e_q and n_q the maximal q-powers dividing e and n, respectively, we have

$$F(n,e) = \prod_q F(n_q, e_q) = \prod_q \varphi\Big(\frac{n_q}{e_q}\Big)e_r = \varphi\Big(\frac{n}{e}\Big)e,$$

which gives the first part of the result.

Now, suppose that the Jordan canonical form of the matrix is $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & \beta \end{smallmatrix}\right)$ with distinct $\alpha, \beta \in \mathbb{F}_{p^2} \setminus \mathbb{F}_p$. The Frobenius endomorphism generates the associated Galois group which is isomorphic to S_2, then it permutes the roots of the characteristic polynomial and we have $\beta = \alpha^p$. As m is fixed, we can choose a generator $g \in \mathbb{F}_{p^2}^*$ such that $m = g^{(p^2-1)/e}$. Therefore, we seek elements in $\mathbb{F}_{p^2}$ of the form $g^{(p^2-1)r/n}$ with $0 \leqslant r \leqslant n$, $\gcd(r,n) = 1$, which do not belong to $\mathbb{F}_p$ and also satisfy $(g^{(p^2-1)r/n})^{p+1} = g^{(p^2-1)/e}$. The first condition is equivalent to $n \nmid p-1$, and the second leads us to compute

$$\#\Big\{0 \leqslant r \leqslant ke\colon \gcd(r, ke) = 1,\ \frac{p+1}{k}r \equiv 1 \pmod{e}\Big\}. \tag{2.1}$$

Of course, necessarily $\gcd((p+1)/k, e) = 1$, and noting that $\gcd(p+1, p-1) = 2$, this is equivalent to saying that either $(p+1)/k$ or e is odd. For $\gcd(a,n) = 1$, let

$$S(a) = \{0 \leqslant r \leqslant ke\colon \gcd(r,k) = 1 \text{ and } r \equiv a \pmod{e}\}.$$

Clearly $|S(a)|$ does not depend on the choice of a. Let $\{a = a_1, \dots, a_{\varphi(e)}\}$ be a complete set of representatives of $(\mathbb{Z}/e\mathbb{Z})^*$ with $\gcd(a_i, k) = 1$. Then (2.1) coincides with

$$\frac{1}{\varphi(e)} \sum_{i=1}^{\varphi(e)} |S(a_i)| = \frac{1}{\varphi(e)} \#\{0 \leqslant r \leqslant ke\colon \gcd(r,k) = 1,\ (r,e) = 1\} = \frac{\varphi(n)}{\varphi(e)},$$

which completes the proof. □

Define $g(n)$ like $f(n)$ but now considering non-diagonalizable matrices in the same set. Again $g(n) = 0$ if $e \nmid n$. Moreover, the non-vanishing of g requires m to be a quadratic residue or equivalently $(p-1)/e$ to be even because of the double root of the characteristic polynomial. We have

Lemma 2.2. *Let $(p-1)/e$ be even, then*

$$g(n) = \begin{cases} \dfrac{1-(-1)^e}{2} & \text{if } n = ep, \\ \dfrac{3+(-1)^e}{2} & \text{if } n = 2ep, \\ 0 & \text{otherwise.} \end{cases}$$

P r o o f. Since the matrix is not diagonalizable, it must be similar to one of the form $\left(\begin{smallmatrix} \alpha & 1 \\ 0 & \alpha \end{smallmatrix}\right)$ with $\alpha \in \mathbb{F}_p^*$.

Let g be a generator of $\mathbb{F}_p^*$ such that $m = g^{(p-1)/e}$. As $\alpha^2 = m$, clearly $2 \mid (p-1)/e$, and so $\alpha = g^{(p-1)/2e}$ or $g^{(p-1)/2e+(p-1)/2}$. In the first case, the order of α is $2e$ and therefore the order of the matrix is $2ep$, while in the second case the order of α can be $2e$ or e, depending on whether e is even or not, in which case the order of the matrix is $2ep$ or ep, respectively. □

Lemma 2.3. *Let p be an odd prime, $p \nmid m$ and $x > 0$. Consider*

$$S_{m,p}(x) = \{\operatorname{tr}(A)\colon\ A \in \mathrm{GL}_2(\mathbb{F}_p) \text{ with } \det A = m,\ \exp_p(A) \leqslant x\},$$

then

$$|S_{m,p}(x)| = \frac{1}{2} \sum_{\substack{k \leqslant x/e \\ k \mid (p-1)/e}} \varphi(k)e + \frac{1}{2} \sum_{\substack{k \leqslant x/e \\ k \mid p+1}} \frac{\varphi(ke)}{\varphi(e)} + O(1).$$

P r o o f. First, note that the part coming from matrices with a double eigenvalue contributes $O(1)$.

For the rest of the cases we apply Lemma 2.1, noting that swapping the eigenvalues, a pair of diagonal Jordan canonical forms corresponds to a class of matrices under similarity and hence to a value of the trace. Note that $\operatorname{tr}(A) = \operatorname{tr}(B)$, $\det A = \det B = m$ defines uniquely the eigenvalues and hence the Jordan canonical form (up to a permutation) when they are distinct. □

Lemma 2.4. *For p an odd prime, $p \nmid m$ and $x > 0$ we have*

$$|S_{m,p}(x)| \ll \varepsilon_m(p) \sum_{\substack{n \leqslant x \\ n \mid p-1}} \varphi(n) + \sum_{\substack{n \leqslant x \\ n \mid p+1}} \varphi(n),$$

where

$$\varepsilon_m(x) = \begin{cases} 1 & \text{if } m = \pm 1, \\ \log\log x & \text{otherwise.} \end{cases}$$

P r o o f. This is a consequence of the previous lemma. The case $m = \pm 1$ is trivial. For $m > 1$, using the definition of φ and [8], Theorem 328, we have

$$\varphi\Big(\frac{n}{e}\Big)e \leqslant \varphi(n)\frac{e}{\varphi(e)} \ll \varphi(n)\log\log e.$$

For the second sum, note that $\gcd(k, e) \mid 2$ because $k \mid p + 1$, and therefore

$$\frac{\varphi(ke)}{\varphi(e)} \leqslant 2\varphi(k)$$

follows easily. □

We need the larger sieve inequality [7]:

Theorem 2.5 ([7], Theorem 1)**.** *If all but $g(p)$ residue classes $\pmod p$ are removed for each prime p in a finite set $\mathcal{S}$, then the number of integers which remain in any interval of length N is at most*

$$\Big(\sum_{p \in \mathcal{S}} \log p - \log N\Big) \Big/ \Big(\sum_{p \in \mathcal{S}} \frac{\log p}{g(p)} - \log N\Big)$$

provided the denominator is positive.

We remark that in [7] this result is stated in a slightly more general form allowing $\mathcal{S}$ to contain prime powers (see Proposition 9.13 of [6] for a flexible version). Of course, the upper bound increases when we sieve with less elements.

Lemma 2.6. *Consider*

$$\mathcal{T}_m(x) = \{1 \leqslant t \leqslant 2N\colon\ t \in S_{m,p}(x) \text{ for every } p \in I\},$$

with $S_{m,p}$ as in Lemma 2.3. For $x = |I|M^{-1}\log M$ with $M > C'\varepsilon_m(|I|)\log|I|\log N$ where C' is a constant, we have

$$|\mathcal{T}_m(x)| \ll \varepsilon_m(|I|)M^{-1}|I|\log|I|.$$

P r o o f. The proof is similar to that of Theorem 1.1 (see [7]). From the Cauchy-Schwarz inequality we obtain

$$\Big(\sum_{p\in I}\frac{\log p}{|S_{m,p}(x)|}\Big)\Big(\sum_{p\in I}|S_{m,p}(x)|\log p\Big) \geqslant \Big(\sum_{p\in I}\log p\Big)^2 \gg |I|^2.$$

On the other hand, the Brun-Titchmarsh theorem [21] gives the bound

$$\pi(y_0+y_1;q,c)-\pi(y_0;q,c) < \frac{2y_1}{\varphi(q)\log(y_1/q)}, \quad 1 \leqslant q < y_1$$

which, together with Lemma 2.4, gives

$$\begin{aligned}\sum_{p\in I}|S_{m,p}(x)|\log p &\ll \sum_{p\in I}\Big(\varepsilon_m(p)\sum_{\substack{n\leqslant x\\ n|p-1}}\varphi(n)+\sum_{\substack{n\leqslant x\\ n|p+1}}\varphi(n)\Big)\log p\\ &\ll |I|\log|I|\varepsilon_m(|I|)\sum_{n\leqslant x}\Big(\log\frac{|I|}{n}\Big)^{-1}\\ &\ll \varepsilon_m(|I|)M^{-1}|I|^2\log|I|\end{aligned}$$

where in the second inequality we have changed the order of summation to

$$\sum_{n\leqslant x}\varphi(n)\Bigg(\sum_{\substack{p\in I\\ p\equiv 1\ (n)}}\varepsilon_m(p)\log p+\sum_{\substack{p\in I\\ p\equiv -1\ (n)}}\log p\Bigg)$$

before applying the Brun-Titchmarsh inequality.

Thus, it follows that

$$\sum_{p\in I}\frac{\log p}{|S_{m,p}(x)|} \gg \frac{M}{\varepsilon_m(|I|)\log|I|}.$$

Now, by Theorem 2.5 we obtain the result

$$|\mathcal{T}_m(x)| \ll \Big(\sum_{p\in I}\log p\Big)\Big/\Big(\sum_{p\in I}\frac{\log p}{|S_{m,p}(x)|}\Big) \ll \varepsilon_m(|I|)M^{-1}|I|\log|I|.$$

The size of M ensures that the denominator in the statement of Theorem 2.5 is positive and, indeed, for a suitably chosen C' it is greater than $c\log N$ with $c>0$.

□

Lemma 2.7. *Let*

$$\mathcal{A}_m = \sup_t \#\{A\in\mathcal{I}_N^*\colon \det A = m,\ \operatorname{tr}(A)=t\},$$

then

$$\mathcal{A}_m \ll N(\log N)^2(\log\log N)^2, \quad \textit{for every } m.$$

P r o o f. The problem is reduced to counting the number of solutions of

$$\begin{cases} 0\leqslant a_{11},a_{12},a_{21},a_{22}\leqslant N,\\ a_{11}+a_{22}=t,\\ a_{11}a_{22}-a_{12}a_{21}=m. \end{cases}$$

Writing $h(n)=n(t-n)-m$, we have $h(a_{11})=a_{12}a_{21}$, and then the number of solutions is bounded by

$$\sum_{n\leqslant N}\sum_{k\mid h(n)}1 = \sum_{k\leqslant N}\sum_{\substack{n\leqslant N\\ h(n)\equiv 0\ (k)}}1 \ll \sum_{k\leqslant N}\varrho(k)\frac{N}{k}$$

where $\varrho(k)$ represents the number of solutions of $h(n)\equiv 0 \pmod k$. Since ϱ is multiplicative, we have

$$\sum_{k\leqslant N}\frac{\varrho(k)}{k} \leqslant \prod_{p\leqslant N}\Big(1+\frac{\varrho(p)}{p}+\frac{\varrho(p^2)}{p^2}+\dots\Big).$$

We separate the product in two parts. In the former one we consider the primes which satisfy $\varrho(p)=0,2$ and thus by [19], Lemma 6.1, verify $\varrho(p^k)\leqslant 2$, and in the latter part those with $\varrho(p)=1$, equivalently $p\mid\Delta=t^2-4m$, in which case ([19], Lemma 6.1) ensures $\varrho(p^k)\leqslant p^{[k/2]}$ for $k\geqslant 1$. Thus, the product is bounded by

$$\begin{aligned}&\prod_{\substack{p\leqslant N\\ \varrho(p)=0,2}}\Big(1+\frac{2}{p}+\frac{2}{p^2}+\dots\Big)\prod_{\substack{p\leqslant N\\ p\mid\Delta}}\Big(1+\frac{1}{p}+\frac{p}{p^2}+\dots\Big)\\ &\qquad\ll \prod_{p\leqslant N}\Big(1+\frac{2}{p}\Big)\prod_{\substack{p\leqslant N\\ p\mid\Delta}}\Big(1+\frac{2}{p}\Big) \ll (\log N)^2\prod_{\substack{p\leqslant N\\ p\mid\Delta}}\Big(1+\frac{2}{p}\Big).\end{aligned}$$

Now, using [8], Theorem 323, the last product is bounded by $(\log\log N)^2$, hence

$$\mathcal{A}_m \ll N(\log N)^2(\log\log N)^2$$

and the result follows. □

Lemma 2.8. *We have*

$$|\mathcal{I}_N| = \frac{12}{\pi^2}N^2 + O(N(\log N)^2) \quad \textit{and} \quad |\mathcal{I}_N^*| = (N+1)^4 + O(N^2(\log N)^3).$$

P r o o f. Let N_{cd} be the number of matrices in $\mathcal{I}_N$ having (c,d) as the lower row. Of course c and d must be coprime and it is easy to see

$$|\mathcal{I}_N| = \sum_{0<d<c\leqslant N} N_{cd} + \sum_{0<c<d\leqslant N} N_{cd} + O(N).$$

If $0 < d < c \leqslant N$ then the possible upper rows of a matrix counted in N_{cd} are $(x_0 + ct, y_0 + dt)$ with $0 \leqslant t \leqslant (N - x_0)/c$ where $x_0 d - y_0 c = 1$ and $x_0 = \bar{d}$, defined as the solution of $dx \equiv 1 \pmod{c}$ with $0 \leqslant x < c$.

The case $0 < c < d \leqslant N$ is very similar but now $0 \leqslant t \leqslant (N - y_0)/c$ and $y_0 = d - \bar{c}$ where $\bar{c}$ is the solution of $cx \equiv -1 \pmod{d}$ with $0 \leqslant x < d$.

Then we have

$$|\mathcal{I}_N| = \sum_{\substack{d<c\leqslant N\\ \gcd(c,d)=1}} \Big(\Big[\frac{N-\bar{d}}{c}\Big] + 1\Big) + \sum_{\substack{c<d\leqslant N\\ \gcd(c,d)=1}} \Big(\Big[\frac{N-c+\bar{d}}{c}\Big] + 1\Big) + O(N)$$

where $[\cdot]$ denotes the integral part. Exchanging c and d in the last sum and introducing the function $\psi(x) = x - [x] - 1/2$, we can write the previous formula as

$$|\mathcal{I}_N| = \sum_{\substack{d<c\leqslant N\\ \gcd(c,d)=1}} \frac{2N}{c} - \sum_{\substack{d<c\leqslant N\\ \gcd(c,d)=1}} \Big(\psi\Big(\frac{N-\bar{d}}{c}\Big) + \psi\Big(\frac{N+\bar{d}}{c}\Big)\Big) + O(N). \tag{2.2}$$

The first sum gives the main term plus an admissible error terms by partial summation of $\sum_{n\leqslant x} \varphi(n) = 3x^2/\pi^2 + O(x\log x)$ that is well known ([8], Theorem 330). It remains to prove that the second sum is $O(N(\log N)^2)$.

Given a positive integer M, there exist real numbers $a_m^{\pm} \ll m^{-1}$ and $a_0^{\pm} \ll M^{-1}$ such that (see for instance Vaaler's lemma in [20], §1.2)

$$\sum_{|m|\leqslant M} a_m^- e(mx) \leqslant \psi(x) \leqslant \sum_{|m|\leqslant M} a_m^+ e(mx) \quad \text{with } e(t) = \mathrm{e}^{2\pi \mathrm{i} t}.$$

Using the evaluation of the Ramanujan sums and $\varphi(ab) \leqslant a\varphi(b)$, we arrive at

$$\Bigg| \sum_{\substack{d=1 \\ \gcd(c,d)=1}}^{c} e\Big(m\frac{\bar{d}}{c}\Big) \Bigg| \leqslant \gcd(c,m).$$

Hence the second sum in (2.2) is bounded by

$$\frac{N^2}{M} + \sum_{c\leqslant N} \sum_{m\leqslant M} \frac{\gcd(c,m)}{m} \leqslant \frac{N^2}{M} + \sum_{d\leqslant M} d \sum_{d|c\leqslant N} \sum_{d|m\leqslant M} \frac{1}{m}.$$

Choosing $M = N(\log N)^{-2}$ one gets the result.

The second formula in the statement reduces to proving that the number of singular matrices with entries $0 \leqslant a, b, c, d \leqslant N$ is $O(N^2(\log N)^3)$. It is easy to see that there are only $O(N^2)$ of them with $abcd = 0$, hence we assume $a, b, c, d > 0$. These singular matrices are clearly overcounted by

$$\sum_{a\leqslant N} \sum_{d\leqslant N} \sum_{b|ad} 1 \leqslant \sum_{m\leqslant N^2} \sum_{d|m} \sum_{b|m} 1 = \sum_{m\leqslant N^2} \tau^2(m)$$

where $\tau(m)$ is the divisor function. Using elementary arguments ([4], page 140) we deduce that the last sum has the expected order of magnitude. □

3. Proof of the main results

P r o o f of Theorem 1.2. Recall that we defined

$$\mathcal{M}_N(x) = \{A \in \mathcal{I}_N \colon \exp_p(A) \leqslant x \text{ for every } p \in I\}.$$

Clearly, with the same notation as in Lemma 2.6, we have

$$|\mathcal{M}_N(x)| \leqslant \sum_{t\in\mathcal{T}_1(x)} |\{A \in \mathcal{I}_N \colon \operatorname{tr}(A) = t\}|.$$

Hence $|\mathcal{M}_N(x)| \leqslant |\mathcal{T}_1(x)|\mathcal{A}_1$, with $\mathcal{A}_1$ as in Lemma 2.7, and the bounds in Lemmas 2.6, 2.7 and 2.8 give, choosing M comparable to $\log N \log |I|$,

$$\mathcal{P}_N(x) \ll C\mathfrak{S} \quad \text{for } x = |I|\mathfrak{T}.$$

Choosing C small enough we obtain the result. □

P r o o f of Theorem 1.3. In this case we are interested in the set

$$\mathcal{M}_N^*(x) = \{A \in \mathcal{I}_N^* \colon\ 0 < \exp_p(A) \leqslant x \text{ for every } p \in I\}.$$

As the determinant is multiplicative, if $A \in \mathrm{GL}_2(\mathbb{F}_p)$ has order n, then the order of $\det A$, seen as an element of $\mathbb{F}_p$, divides n. Hence

$$|\mathcal{M}_N^*(x)| \leqslant \sum_{m \in \mathcal{Z}} |\{A \in \mathcal{I}_N^* \colon\ \det A = m,\ 0 < \exp_p(A) \leqslant x \text{ for every } p \in I\}|$$

where $\mathcal{Z} = \{m \colon\ 1 \leqslant |m| \leqslant N^2,\ 0 < \exp_p(m) \leqslant x \text{ for every } p \in I\}$.

The number of elements of order than less or equal to x in $\mathbb{F}_p^*$ is $\sum \varphi(n)$ where the sum runs over $n \leqslant x$ with $n \mid p-1$, which is majorized by $2|S_{1,p}(x)|$ (see Lemma 2.3). Then proceeding as in Lemma 2.6, we get a bound for $|\mathcal{Z}|$ similar to that for $|\mathcal{T}_1(x)|$,

$$|\mathcal{Z}| \ll M^{-1}|I| \log |I| \ll \frac{C|I|}{\log N \log\log |I|}$$

with M comparable with $C^{-1} \log N \log |I| \log\log |I|$ that corresponds to $x = |I|\mathfrak{T}^*$.

Writing

$$|\{A \in \mathcal{I}_N^* \colon\ \det A = m,\ 0 < \exp_p(A) \leqslant x \text{ for every } p \in I\}| \leqslant \sum_{t \in \mathcal{T}_m(x)} \mathcal{A}_m$$

and using Lemmas 2.6, 2.7 and 2.8 (see the proof of Theorem 1.2), we conclude

$$\mathcal{P}_N^*(x) \ll C\mathfrak{S}^*$$

and again it is enough to choose C small enough. □

4. Some other questions about the distribution

We are interested in knowing whether we can always get large order matrices with small perturbations. To do this, we can restrict ourselves to the study of traces, and then translate the results to matrices through the following lemma.

Lemma 4.1. *Let A be an element of $\mathrm{SL}_2(\mathbb{Z})$. Fixing a generator g of $\mathbb{F}_p^*$ and a prime factor q of $p-1$, we have that $\exp_p(A) = q$ if and only if*

$$\mathrm{tr}(A) = g^{k(p-1)/q} + g^{-k(p-1)/q}$$

for some $1 \leqslant k < q$.

P r o o f. The condition imposed on the order of the matrix forces its Jordan canonical form to be diagonal with entries $\alpha, \alpha^{-1} \in \mathbb{F}_p$. The result is proved by writing $\alpha = g^{k(p-1)/q}$ since $\exp_p(A) = q$. □

Identifying matrices with the same trace and taking the distance (between classes) to be the distance between traces, we can obtain results about a kind of discrepancy of matrices. The proofs are provided later in this section.

Theorem 4.2. *Let J be an interval of length greater than $6p^{3/2}(q-1)^{-1}\log p$, where q is a prime divisor of $p-1$. Then there exists a matrix $A \in \mathrm{SL}_2(\mathbb{F}_p)$ such that $\mathrm{tr}(A) \in J$ and $\exp_p(A) = q$.*

R. C. Baker and G. Harman proved in [1] that for infinitely many primes p the largest factor of $p-1$ is greater than $p^{0.677}$. In fact, this is actually proved for a positive proportion of the primes (see [9], §8.1, specially 8.1.7, and the nearby formulas). Using this result, we obtain the following corollary:

Corollary 4.3. *There exist positive constants C_1 and C_2 such that for at least $C_1 N/\log N$ prime numbers $p \in [N, 2N]$, there are matrices with $\exp_p(A) > N^{0.677}$ in any interval of length greater than $C_2 N^{0.823}\log N$.*

Now, we change our point of view. We fix a matrix of large order and proceed to study the distribution of its powers. Observe that the maximum order of a diagonalizable matrix in $\mathrm{SL}_2(\mathbb{F}_p)$ is $p-1$, so we expect a matrix of this order to be a good random vector generator. The next result shows that the powers of these matrices are well distributed.

Theorem 4.4. *Let $A \in \mathrm{SL}_2(\mathbb{Z})$ be such that $\exp_p(A) = p-1$. Then*

$$\#\{A^k,\ 1 \leqslant k \leqslant N\colon\ \mathrm{tr}(A^k) \in J\} = \frac{N|J|}{p} + O(p^{1/2}(\log p)^2),$$

where J is an interval contained in $[1,p]$ and $N < p$.

For every pair of integers m and n, with $p \nmid m$, we define the trigonometric sum

$$S(N) = \sum_{k=1}^{N} e\Big(\frac{m}{p}(g^k + g^{-k})\Big)e\Big(\frac{nk}{p-1}\Big) \quad \text{where } e(t) = \mathrm{e}^{2\pi i t}.$$

To prove the previous theorems we address before the following lemmas.

Lemma 4.5. *We have*

$$a)\ \ |S(p-1)| \leqslant 2p^{1/2} \quad \textit{and} \quad b)\ \ |S(N)| \leqslant 7p^{1/2}\log p$$

where $p \nmid m$ *and* $N < p$.

P r o o f. In the former case, apply the change of variable $x = g^k$ to obtain

$$S(p-1) = \sum_{x=1}^{p-1} e\Big(\frac{m}{p}(x+\bar{x})\Big)\chi(x)$$

where χ is a certain Dirichlet character, and $\bar{x}$ denotes the inverse of x modulo p. Now the result follows from [17], Theorem 10. To prove the bound b), we can employ the completing technique (see Lemma 12.1 in [11], §12.2). Let $\widetilde{S}$ be given by

$$\widetilde{S}\Big(\frac{a}{p}\Big) = \sum_{0<k\leqslant p-1} e\Big(\frac{(m-a)g^k + mg^{-k}}{p}\Big)e\Big(\frac{nk}{p-1}\Big),$$

so that

$$S(N) = \frac{1}{p}\sum_{a \pmod p} \lambda\Big(\frac{a}{p}\Big)\widetilde{S}\Big(\frac{a}{p}\Big)$$

where

$$\lambda\Big(\frac{a}{p}\Big) = \sum_{0<y\leqslant N} e\Big(\frac{ay}{p}\Big).$$

On the one hand $\lambda(0) = N$, while on the other hand, for $0 < |a| \leqslant p/2$ we have $|\lambda(a/p)| \leqslant p|a|^{-1}$ (note that $\lambda(a/p)$ is the sum of a geometric progression). Therefore, using these observations and the bound a), we conclude

$$\begin{aligned}|S(N)| &\leqslant \frac{N}{p}|\widetilde{S}(0)| + \sum_{0<|a|\leqslant p/2} |a|^{-1}\Big|\widetilde{S}\Big(\frac{a}{p}\Big)\Big| \\ &\leqslant 2p^{1/2} + 4p^{1/2}(1+\log(p/2)) \leqslant 7p^{1/2}\log p.\end{aligned}$$

□

Lemma 4.6. *We write $S(m,n;N)$ instead of $S(N)$ to emphasize the dependance on the parameters. Then*

$$\sum_{k=1}^{q} e\Big(\frac{m}{p}(g^{k(p-1)/q}+g^{-k(p-1)/q})\Big) = \frac{q}{p-1}\sum_{h=1}^{(p-1)/q} S(m,hq;p-1).$$

P r o o f. By the definition of $S(N)$, the sum on the right-hand side is

$$\sum_{s=1}^{p-1} e\Big(\frac{m}{p}(g^s+g^{-s})\Big)\sum_{h=1}^{(p-1)/q} e\Big(\frac{hqs}{p-1}\Big)$$

where the inner sum is $(p-1)/q$ if $(p-1)/q$ divides s and zero otherwise. After the change of variables $s \to k(p-1)/q$, we obtain the result. □

P r o o f of Theorem 4.2. We can assume $q-1 > \mathrm{e}p^{1/2}\log p$, where here and in the rest of the proof "e" is the base of the natural logarithm, because otherwise the result is trivial. In particular we can assume $p > 211$ (note that for $p \leqslant 211$, $p < (\mathrm{e}\log p)^2$ and hence $q-1 < \mathrm{e}p^{1/2}\log p$ because $q \leqslant p-1$).

Let g be a generator of $\mathbb{F}_p^*$. Consider $\alpha = g^{(p-1)/q}$ and let $t_k = (\alpha^k+\alpha^{-k})p^{-1}$ be the normalized traces modulo p in $[0,1]$. Then

$$\#\{k \leqslant q-1\colon\ t_k \in [a,b]\} \geqslant (b-a)(q-1) - D(q-1)(q-1),$$

where

$$D(N) = \sup_{0\leqslant a<b\leqslant 1}\Big|\frac{\#\{k\leqslant N\colon\ a\leqslant t_k\leqslant b\}}{N} - (b-a)\Big|$$

is the discrepancy of the sequence $\{t_k\}$. On the one hand, by Lemmas 4.6 and 4.5,

$$\Big|\sum_{k=1}^{q-1}\mathrm{e}(mt_k)\Big| \leqslant \frac{q}{p-1}\sum_{h=1}^{(p-1)/q}|S(m,hq;p-1)| \leqslant 2p^{1/2}.$$

On the other hand, the Erdős-Turán inequality ([20], Corollary 1.1) yields

$$\begin{aligned} D(q-1) &\leqslant \frac{1}{M+1} + \frac{3}{q-1}\sum_{m=1}^{M}\frac{1}{m}\Big|\sum_{k=1}^{q-1}\mathrm{e}(mt_k)\Big| \\ &\leqslant \frac{1}{M+1} + \frac{6p^{1/2}}{q-1}(1+\log(M+1)), \end{aligned}$$

and choosing M to be the integral part of $(q-1)/\mathrm{e}p^{1/2}\log p$, we get $(q-1)D(q-1) \leqslant 6p^{1/2}\log p$ because $1/(M+1) \leqslant \mathrm{e}p^{1/2}\log p/(q-1)$ and (recall that $p > 211$)

$$\log(M+1) \leqslant \log\Big(\frac{p-1}{\mathrm{e}p^{1/2}\log p}+1\Big) \leqslant \log\frac{p^{1/2}}{\mathrm{e}} = -1+\frac{1}{2}\log p.$$

We conclude that the interval $J = J(p,q)$ satisfies $|J| > 6p^{3/2}(q-1)^{-1}\log p$. □

P r o o f of Theorem 4.4. In this case, the matrix is similar to one of the form $\left(\begin{smallmatrix} g & 0 \\ 0 & g^{-1} \end{smallmatrix}\right)$ with g a generator of $\mathbb{F}_p^*$, so $tr(A^k) = g^k + g^{-k}$.

Let $t_k = (g^k + g^{-k})p^{-1}$ be the normalized traces modulo p in $[0, 1]$. By Lemma 4.5,

$$\Big|\sum_{k=1}^{N} \mathrm{e}\Big(\frac{m}{p}(g^k + g^{-k})\Big)\Big| \leqslant 7p^{1/2}\log p,$$

and we can apply the Erdős-Turán inequality again to obtain

$$D(N) \leqslant \frac{1}{M+1} + \frac{21p^{1/2}\log p}{N}(1 + \log(M+1)).$$

Taking

$$M = \Big\lceil \frac{N}{p^{1/2}(\log p)^2} \Big\rceil$$

where $\lceil x \rceil$ is the smallest integer not less than x, we get $D(N) \ll N^{-1}p^{1/2}(\log p)^2$ and the result follows by the definition of discrepancy [12]. □

Acknowledgment. We are grateful to M. Z. Garaev for insightful comments on Lemma 4.5. We thank the anonymous referee for careful reading of the manuscript and useful comments.

References

[1] *R. C. Baker, G. Harman*: Shifted primes without large prime factors. Acta Arith. *83* (1998), 331–361.
[2] *R. C. Baker, G. Harman, J. Pintz*: The difference between consecutive primes II. Proc. Lond. Math. Soc. (3) *83* (2001), 532–562.
[3] *M.-C. Chang*: Burgess inequality in $\mathbb{F}_{p^2}$. Geom. Funct. Anal. *19* (2009), 1001–1016.
[4] *H. Davenport*: Multiplicative Number Theory (2nd rev. ed.). Graduate Texts in Mathematics 74, Springer, New York, 1980.
[5] *J. Eichenauer-Herrmann, H. Grothe, J. Lehn*: On the period length of pseudorandom vector sequences generated by matrix generators. Math. Comput. *52* (1989), 145–148.
[6] *J. Friedlander, H. Iwaniec*: Opera de Cribro. American Mathematical Society Colloquium Publications 57, Providence, 2010.
[7] *P. X. Gallagher*: A larger sieve. Acta Arith. *18* (1971), 77–81.
[8] *G. H. Hardy, E. M. Wright*: An Introduction to the Theory of Numbers (6th rev. ed.). Oxford University Press, Oxford, 2008.
[9] *G. Harman*: Prime-Detecting Sieves. London Mathematical Society Monographs Series 33, Princeton University Press, Princeton, 2007.
[10] *M. N. Huxley*: On the difference between consecutive primes. Invent. Math. *15* (1972), 164–170.
[11] *H. Iwaniec, E. Kowalski*: Analytic Number Theory. American Mathematical Society Colloquium Publications 53, Providence, 2004.
[12] *L. Kuipers, H. Niederreiter*: Uniform Distribution of Sequences. Pure and Applied Mathematics, John Wiley & Sons, New York, 1974.

[13] *P. Kurlberg*: On the order of unimodular matrices modulo integers. Acta Arith. *110* (2003), 141–151.
[14] *P. Kurlberg, L. Rosenzweig, Z. Rudnick*: Matrix elements for the quantum cat map: fluctuations in short windows. Nonlinearity *20* (2007), 2289–2304.
[15] *P. Kurlberg, Z. Rudnick*: On quantum ergodicity for linear maps of the torus. Commun. Math. Phys. *222* (2001), 201–227.
[16] *P. L'Ecuyer*: Uniform random number generation. Ann. Oper. Res. *53* (1994), 77–120.
[17] *W. C. W. Li*: Number Theory with Applications. Series on University Mathematics 7, World Scientific, River Edge, 1996.
[18] *H. Maier*: Primes in short intervals. Michigan Math. J. *32* (1985), 221–225.
[19] *R. A. Mollin*: Advanced Number Theory with Applications. Discrete Mathematics and Its Applications, CRC Press, Boca Raton, 2010.
[20] *H. L. Montgomery*: Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis. CBMS Regional Conference Series in Mathematics 84, AMS, Providence, 1994.
[21] *H. L. Montgomery, R. C. Vaughan*: The large sieve. Mathematika, Lond. *20* (1973), 119–134.
[22] *H. Niederreiter*: Statistical independence properties of pseudorandom vectors produced by matrix generators. J. Comput. Appl. Math. *31* (1990), 139–151.
[23] *H. Roskam*: A quadratic analogue of Artin's conjecture on primitive roots. J. Number Theory *81* (2000), 93–109.
[24] *V. Shoup*: Searching for primitive roots in finite fields. Math. Comput. *58* (1992), 369–380.
[25] *P. J. Stephens*: An average result for Artin's conjecture. Mathematika, Lond. *16* (1969), 178–188.

Authors' address: Fernando Chamizo, Dulcinea Raboso, Department of Mathematics and ICMAT, Faculty of Science, Universidad Autónoma de Madrid, Francisco Tomás y Valiente 7, 280 49 Madrid, Spain, e-mail: `fernando.chamizo@uam.es`, `dulcinea.raboso@uam.es`.