ELEMENTARY DIVISORS AND DETERMINANTS OF RANDOM MATRICES OVER A LOCAL FIELD

STEVEN N. EVANS

Abstract. We consider the elementary divisors and determinant of a uniformly distributed $n \times n$ random matrix with entries in the ring of integers of an arbitrary local field. We show that the sequence of elementary divisors is in a simple bijective correspondence with a Markov chain on the nonnegative integers. The transition dynamics of this chain do not depend on the size of the matrix. As $n \to \infty$, all but finitely many of the elementary divisors are 1, and the remainder arise from a Markov chain with these same transition dynamics. We also obtain the distribution of the determinant of M_n and find the limit of this distribution as $n \to \infty$. Our formulae have connections with classical identities for q-series, and the q-binomial theorem in particular.

Department of Statistics #3860
University of California at Berkeley
367 Evans Hall
Berkeley, CA 94720-3860
U.S.A

e-mail: evans@stat.berkeley.edu
URL: http://www.stat.berkeley.edu/users/evans

1. Introduction

A *local field* (see Section 2) is a locally compact, non-discrete, totally disconnected, topological field (a locally compact, non-discrete, topological field that is not totally disconnected is necessarily either the real or the complex numbers.) The best known example is the field of *p-adic numbers* – see Section 2. Every local field is either a finite algebraic extension of the p-adic number field for some prime p or a finite algebraic extension of the *p-series field*; that is, the field of formal Laurent series with coefficients drawn from the finite field with p elements.)

There has been considerable interest in recent years in probability on local fields. We refer the reader to [Eva01] for an indication of the literature that is most relevant to this paper. Here we investigate a particular class of random matrices over a local field. This class is described as follows.

Any local field $\mathcal{K}$ has a maximal compact subring $\mathcal{D}$ called the *ring of integers*. For example, the p-adic numbers arise as a particular completion of the rationals and the ring of integers in this case is just the closure of the integers. The ring $\mathcal{D}$

Date: July 8, 2002.

2000 *Mathematics Subject Classification.* Primary: 15A52, 60B99; Secondary: 12J25, 05A30.

Key words and phrases. local field, p-adic, p-series, random matrix, elementary divisor, Gaussian elimination, determinant, q-binomial coefficient, partition.

Research supported in part by NSF grant DMS-0071468.

has a unique maximal ideal, and this ideal can be written as $\rho\mathcal{D}$ for some $\rho \in \mathcal{D}$. For the p-adic numbers we can take $\rho = p$ and the maximal ideal is just the closure of the integers that are divisible by p.

An $n \times n$ matrix A with entries in $\mathcal{D}$ that is full rank (that is, is invertible over $\mathcal{K}$) can be represented as

$$A = U \operatorname{diag}(\rho^{k_1}, \rho^{k_2}, \ldots, \rho^{k_n})\, V,$$

where U and V are $\mathcal{D}$-valued matrices with $\mathcal{D}$-valued inverses and $0 \le k_1 \le k_2 \le \ldots \le k_n$. The sequence $(\rho^{k_1}, \rho^{k_2}, \ldots, \rho^{k_n})$ is unique and is called *sequence of elementary divisors* of A (see Section 3).

Let M_n be an $n \times n$ random matrix with independent entries distributed according to normalised Haar measure on $\mathcal{D}$ and elementary divisors $(\rho^{K_1^n}, \rho^{K_2^n}, \ldots, \rho^{K_n^n})$. Set $L_k^n = \#\{1 \le i \le n : K_i^n = k\}$, $k \ge 0$. We show in Section 3 that the process n, $n \Leftrightarrow L_0^n$, $n \Leftrightarrow L_0^n \Leftrightarrow L_1^n, \ldots$ is a Markov chain with an explicitly given transition matrix that does not depend on n. Moreover, $n \Leftrightarrow L_0^n$ converges in distribution as $n \to \infty$, so that in the limit all but finitely many of the elementary divisors are 1 and the distribution of the remainder is described in terms of a simple Markov chain.

We also find explicitly the distribution of $\det M_n$ in Section 4 and show that this distribution converges as $n \to \infty$ to one that has a simple density with respect to Haar measure on $\mathcal{D}$. The special case of this result for algebraic extensions of the p-series field (that is, the case of non-zero characteristic) is given in [AG00].

Our results have some of the flavour of the body of work on ranks and determinants of random matrices over finite fields – see, for example, [Koz66, Bal68, Muk84, Wat87, BM87, BW87, Lev91, BKW97, Coo00b, Coo00a]. In particular, our proofs involve solving recursions that are very similar to those appearing in that area, and this leads to interesting connections with q-series and the q-binomial theorem in particular. In this connection we should also mention the work that has been done on the cycle structure and characteristic polynomial of Haar distributed random invertible matrices over finite fields — see, for example, [Kun81, Sto88, HS93, Ful99, Ful02]. We note that the common transition matrix of the Markov chains n, $n \Leftrightarrow L_0^n$, $n \Leftrightarrow L_0^n \Leftrightarrow L_1^n, \ldots$ is a particular case of a class of transition matrices that appears in [Ful02] in a description of the conjugacy class of a randomly chosen element of the group of $m \times m$ matrices over the field with q elements — see Remark 3.8 for more details.. This more general class of chains is used to give a transparent proof of the Rogers–Ramanujan identities. Finally, we note that there are numerous other connections between q-series and probability, particularly those arising from Blomqvist's absorption process and related structures — see, for example, [Raw98, Raw97].

2. Local fields

This section is a summary of background that can be found in numerous sources such as [Tai75, Sch84], and we refer the reader to these works for a fuller account. We begin with prototypical example of a local field: the field of p-adic numbers.

Example 2.1. Fix a positive prime p. We can write any non-zero rational number $r \in \mathbb{Q}\backslash\{0\}$ uniquely as $r = p^s(a/b)$ where a and b are not divisible by p. Set

$|r| = p^{-s}$. If we set $|0| = 0$, then the map $|\cdot|$ has the properties:

$$\begin{aligned} |x| = 0 &\Leftrightarrow x = 0 \\ |xy| &= |x||y| \\ |x+y| &\leq |x| \vee |y|. \end{aligned} \tag{2.1}$$

The map $(x, y) \mapsto |x \Leftrightarrow y|$ defines a metric on $\mathbb{Q}$ and we denote the completion of $\mathbb{Q}$ in this metric by $\mathbb{Q}_p$. The field operations on $\mathbb{Q}$ extend continuously to make $\mathbb{Q}_p$ a topological field called *p-adic numbers*. The map $|\cdot|$ also extends continuously and the extension has properties (2.1). The closed unit ball around 0, $\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x| \leq 1\}$, is the closure in $\mathbb{Q}_p$ of the integers $\mathbb{Z}$, and is thus a ring (this is also apparent from (2.1)), called the *p-adic integers*. As $\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x| < p\}$, the set $\mathbb{Z}_p$ is also open. Any other ball around 0 is of the form $\{x \in \mathbb{Q}_p : |x| \leq p^{-k}\} = p^k\mathbb{Z}_p$ for some integer k. Such a ball is the closure of the rational numbers divisible by p^k, and is thus a $\mathbb{Z}_p$ - module (this is again also apparent from (2.1)). In particular, such a ball is an additive subgroup of $\mathbb{Q}_p$. Arbitrary balls are translates (= cosets) of these closed and open subgroups. In particular, the topology of $\mathbb{Q}_p$ has a base of closed and open sets, and hence $\mathbb{Q}_p$ is totally disconnected. Further, each of these balls is compact, and hence $\mathbb{Q}_p$ is also locally compact.

From now on, we let $\mathcal{K}$ be a fixed local field. There is a real-valued mapping on $\mathcal{K}$ which we denote by $x \mapsto |x|$. This map has the properties (2.1). A map with properties (2.1) is called a non-archimedean valuation. The third of these properties is known as the *ultrametric inequality* or the *strong triangle inequality.* The mapping $(x, y) \mapsto |x \Leftrightarrow y|$ on $\mathcal{K} \times \mathcal{K}$ is a metric on $\mathcal{K}$ which gives the topology of $\mathcal{K}$. A consequence of of the strong triangle inequality is that if $|x| \neq |y|$, then $|x + y| = |x| \vee |y|$. This latter result implies that for every "triangle" $\{x, y, z\} \subset \mathcal{K}$ we have that at least two of the lengths $|x \Leftrightarrow y|$, $|x \Leftrightarrow z|$, $|y \Leftrightarrow z|$ must be equal and is therefore often called the *isosceles triangle property.*

The valuation takes the values $\{q^k : k \in \mathbb{Z}\} \cup \{0\}$, where $q = p^c$ for some prime p and positive integer c (so that for $\mathcal{K} = \mathbb{Q}_p$ we have $c = 1$). Write $\mathcal{D}$ for $\{x \in \mathcal{K} : |x| \leq 1\}$ (so that $\mathcal{D} = \mathbb{Z}_p$ when $\mathcal{K} = \mathbb{Q}_p$). Fix $\rho \in \mathcal{K}$ so that $|\rho| = q^{-1}$ (such a ρ is called a *prime element*). Then

$$\rho^k\mathcal{D} = \{x : |x| \leq q^{-k}\} = \{x : |x| < q^{-(k-1)}\}$$

for each $k \in \mathbb{Z}$ (so that for $\mathcal{K} = \mathbb{Q}_p$ we could take $\rho = p$). The set $\mathcal{D}$ is the unique maximal compact subring of $\mathcal{K}$ (the so-called *ring of integers* of $\mathcal{K}$). Each of the sets $\rho^k\mathcal{D}$, $k \in \mathbb{Z}$, is a compact $\mathcal{D}$- submodule of $\mathcal{K}$ and every non-trivial compact $\mathcal{D}$-submodule of $\mathcal{K}$ is of this form. The set $\rho\mathcal{D}$ is the unique maximal ideal in $\mathcal{D}$ and is called the *prime ideal.* The quotient $\mathcal{D}/\rho\mathcal{D}$ is the finite field with q elements.

For $\ell < k$ the additive quotient group $\rho^\ell\mathcal{D}/\rho^k\mathcal{D}$ has order $q^{k-\ell}$. Consequently, $\mathcal{D}$ is the union of q disjoint translates of $\rho\mathcal{D}$. Each of these components is, in turn, the union of q disjoint translates of $\rho^2\mathcal{D}$, and so on. We can thus think of the collection of balls contained in $\mathcal{D}$ as being arranged in an infinite rooted q-ary tree: the root is $\mathcal{D}$ itself, the nodes at level k are the balls of radius q^{-k} (= cosets of $\rho^k\mathcal{D}$), and the q "children" of such a ball are the q cosets of $\rho^{k+1}\mathcal{D}$ that it contains. We can uniquely associate each point in $\mathcal{D}$ with the sequence of balls that contain it, and so we can think of the points in $\mathcal{D}$ as the boundary of this tree.

There is a unique Borel measure λ on $\mathcal{K}$ for which

$$\begin{aligned} \lambda(x + A) &= \lambda(A), \quad x \in \mathcal{K}, \\ \lambda(xA) &= |x|\lambda(A), \quad x \in \mathcal{K}, \\ \lambda(\mathcal{D}) &= 1. \end{aligned}$$

The measure λ is a suitably normalised Haar measure on the additive group of $\mathcal{K}$. In the case of $\mathbb{Q}_p$, the restriction of λ to $\mathbb{Z}_p$ is the weak limit as $n \to \infty$ of the sequence of probability measures that at the n-th stage assigns mass p^{-n} to each of the points $\{0, 1, \ldots, p^n \Leftrightarrow 1\}$.

Equip the $\mathcal{K}$-vector space $\mathcal{K}^n$ with the norm $\|\cdot\|$ given by

$$\|(x_1, \ldots, x_n)\| = \bigvee_{i=1}^{n} |x_i|.$$

Note that

$$\begin{aligned} \|x\| &= 0 \Leftrightarrow x = 0; \\ \|\lambda x\| &= |\lambda|\|x\|, \quad \lambda \in \mathcal{K}, \\ \|x + y\| &\le \|x\| \vee \|y\|. \end{aligned}$$

The balls around 0 in this space are all of the form

$$\begin{aligned} &\{(x_1, \ldots, x_n) : \|(x_1, \ldots, x_n)\| \le q^{-k}\} \\ &\quad = \{(x_1, \ldots, x_n) : |x_i| \le q^{-k}, \ 1 \le i \le n\} \\ &\quad = (\rho^k \mathcal{D})^n = \rho^k \mathcal{D}^n \end{aligned}$$

for some integer k.

We take our normalisation of Haar measure on the additive group of $\mathcal{K}^n$ to be such that we have the product of n-copies of λ. With a slight abuse of notation, we also denote this measure by λ if the context is clear.

For reasons explained in [Eva01], the natural analogue of (centered) Gaussian measures on $\mathcal{K}^n$ are the normalised Haar measures on compact $\mathcal{D}$-submodules of $\mathcal{K}^n$ (that is, additive subgroups of $\mathcal{K}^n$ that are also closed under multiplication by scalars in $\mathcal{D}$.) Such probability measures are call *$\mathcal{K}$-Gaussian*. When the supporting $\mathcal{D}$-submodule doesn't lie in a lower dimensional subspace (equivalently, is open), then the $\mathcal{K}$-Gaussian measure is just the normalised restriction of λ. In the case $n = 1$, the $\mathcal{K}$-Gaussian measure with support $\mathcal{D}$ is called the *standard* $\mathcal{K}$-Gaussian measure. We note that if Z is a random variable with the standard $\mathcal{K}$-Gaussian distribution, then the conditional distribution of Z given $\{|Z| \le q^{-k}\}$, $k \ge 0$, is the distribution of $\rho^k Z$.

3. Elementary Divisors

Fix $n \in \mathbb{N}$. For $1 \le i, j \le n$, let E_{ij} denote the $n \times n$ $\mathcal{D}$-valued matrix with 1 in the (i, j) position and 0 elsewhere. For $1 \le i, j \le n$, $x \in \mathcal{D}$, and $y \in \mathcal{D}$ with $|y| = 1$, define the $n \times n$ $\mathcal{D}$-valued *elementary matrices*

$$\begin{aligned} A_{ij}(x) &= I + xE_{ij}, \\ B_i(y) &= I + (y \Leftrightarrow 1)E_{ii}, \\ C_{ij} &= I \Leftrightarrow E_{ii} \Leftrightarrow E_{jj} + E_{ij} + E_{ji}. \end{aligned}$$

- Left (resp. right) multiplication of a $n \times n$ matrix H by $A_{ij}(x)$ has the effect of replacing the i-th row (resp. column) of H by the sum of the i-th row (resp. column) and x times the j-th row (resp. column).
- Left (resp. right) multiplication of an $n \times n$ matrix H by $B_i(y)$ multiplies the i-th row (resp. column) of H by y.
- Left (resp. right) multiplication of an $n \times n$ matrix h by C_{ij} interchanges the i-th and j-th rows (resp. columns) of H.

Write $GL(n, \mathcal{D})$ for the group of $n \times n$ $\mathcal{D}$-valued matrices with $\mathcal{D}$-valued inverses. By Cramer's rule, $GL(n, \mathcal{D})$ consists of $n \times n$ $\mathcal{D}$-valued matrices U such that $|\det U| = 1$. It is not hard to see that the elementary matrices belong to $GL(n, \mathcal{D})$.

Theorem 3.1. *Given an $n \times n$ invertible $\mathcal{D}$-valued matrix H, there exist integers $0 \le k_1 \le k_2 \le \cdots \le k_n$ and $U, V \in GL(n, \mathcal{D})$ such that*

$$H = U \operatorname{diag}(\rho^{k_1}, \rho^{k_2}, \ldots, \rho^{k_n}) V.$$

The integers $k_1, \ldots, k_n$ are unique, and U and V may be taken to be products of elementary matrices.

Proof. See Theorem 3.8 of [Jac85]. The proof is essentially just Gaussian elimination adapted to this setting. □

Remark 3.2. i) The vector $(\rho^{k_1}, \ldots, \rho^{k_n})$ is called the vector of *elementary divisors* of H.
ii) If $H \in GL(n, \mathcal{D})$, then clearly $k_1 = k_2 = \cdots = k_n = 0$, and so H is a product of elementary matrices. That is, $GL(n, \mathcal{D})$ is generated by the elementary matrices.
iii) It is easy to see that the elementary matrices are isometries of $(\mathcal{K}^n, \|\cdot\|)$: if F is elementary, then $\|xF\| = \|Fx\| = \|x\|$ for all $x \in \mathcal{K}^n$. Thus the elements of $GL(n, \mathcal{D})$ are isometries. Conversely, if the matrix H is an isometry, then $\operatorname{diag}(\rho^{k_1}, \ldots, \rho^{k_n}) = U^{-1} H V^{-1}$ is also an isometry, and hence $k_1 = \cdots = k_n = 0$. Thus $GL(n, \mathcal{D})$ coincides with the group of (linear) isometries of $(\mathcal{K}^n, \|\cdot\|)$.

Notation 3.3. Set

$$\begin{aligned}\Pi_n &= (1 \Leftrightarrow q^{-1})(1 \Leftrightarrow q^{-2}) \ldots (1 \Leftrightarrow q^{-n}) \\ &= (1 \Leftrightarrow q^{-1})^n (1 + q^{-1}) \ldots (1 + q^{-1} + \cdots + q^{-(n-1)}),\end{aligned}$$

so that

$$\frac{\Pi_n}{\Pi_k \Pi_{n-k}} = \begin{bmatrix} n \\ k \end{bmatrix}_{q^{-1}},$$

the usual q^{-1}-binomial coefficient (see, for example, Ch. 10 of [AAR99]). For simplicity, we write $\begin{bmatrix} n \\ k \end{bmatrix}$ for $\begin{bmatrix} n \\ k \end{bmatrix}_{q^{-1}}$. Set

$$\Pi_\infty = \lim_{n \to \infty} \Pi_n = (1 \Leftrightarrow q^{-1})(1 \Leftrightarrow q^{-2}) \ldots$$

Notation 3.4. Let M_n be an $n \times n$ random matrix with independent standard $\mathcal{K}$-Gaussian entries. Write $(\rho^{K_1^n}, \rho^{K_2^n}, \ldots, \rho^{K_n^n})$ for the elementary divisors of M_n and set $L_k^n = \#\{1 \le l \le n : K_l^n = k\}$, $k \in \mathbb{N}$.

Theorem 3.5. *The sequence of* $\mathbb{N}$*-valued random variables* n, $n \Leftrightarrow L_0^n$, $n \Leftrightarrow L_0^n \Leftrightarrow L_1^n, \dots$ *is a Markov chain on* $\mathbb{N}$ *with transition matrix*

$$P(s,t) = \begin{cases} q^{-t^2} \frac{\Pi_s}{\Pi_t} \begin{bmatrix} s \\ t \end{bmatrix}, & 0 \le t \le s, \\ 0, & \textit{otherwise.} \end{cases}$$

Consequently,

$$\mathbb{P}\{L_0^n = l_0, L_1^n = l_1, \dots\} = q^{-((n-l_0)^2+(n-l_0-l_1)^2+\cdots)} \frac{\Pi_n^2}{\Pi_{l_0}\Pi_{l_1}\dots}, \quad \textit{for} \sum_k l_k = n.$$

Proof. Our proof uses conditionings similar to those used in the proof of Theorem 3.1 of [BM87]. For integers $0 \le r \le n$ and $l_0, l_1, \dots \ge 0$ such that $\Sigma_{k=0}^{\infty} l_k = n$, set

$$\pi_{n,r}(l_0, l_1, \dots) = \mathbb{P}\left\{ L_0^n = l_0,\, L_1^n = l_1, \dots \;\middle|\; \bigvee_{i=1}^{n} \bigvee_{j=1}^{n-r} |M_n(i,j)| \le q^{-1} \right\}$$

so that

$$\pi_{n,n}(l_0, l_1, \dots) = \mathbb{P}(L_0^n = l_0, L_1^n = l_1, \dots).$$

Note that conditional on the event $\left\{ \bigvee_{i=1}^{n} \bigvee_{j=1}^{n-r} |M_n(i,j)| \le q^{-1} \right\}$ the entries of M_n are still independent, with the entries in the first $(n \Leftrightarrow r)$ columns being $\mathcal{K}$-Gaussian supported on $\rho\mathcal{D}$ while the entries in the remaining r columns are standard $\mathcal{K}$-Gaussian. Let $M_{n,r}$ be a random matrix with this conditional distribution. Thus

$$\pi_{n,r}(l_0, l_1, \dots) = \mathbb{P}\left\{ L_0^{n,r} = l_0,\, L_1^{n,r} = l_1, \dots \right\},$$

where we write $L_k^{n,r}$ for the number of elementary divisors of $M_{n,r}$ of the form ρ^k.

The event

$$\left\{ \bigvee_{i=1}^{n} |M_{n,r}(i, n \Leftrightarrow r+1)| \le q^{-1} \right\}$$

has probability q^{-n} and conditional on this event the distribution of $M_{n,r}$ is that of $M_{n,r-1}$. Let $\tilde{M}_{n,r}$ be a random matrix with distribution that of $M_{n,r}$ conditioned on the complementary event

$$\left\{ \bigvee_{i=1}^{n} |M_{n,r}(i, n \Leftrightarrow r+1)| = 1 \right\}.$$

The columns of $\tilde{M}_{n,r}$ are independent, the entries in the first $(n \Leftrightarrow r)$ columns are i.i.d. $\mathcal{K}$-Gaussian on $\rho\mathcal{D}$ and the entries in the last $(r \Leftrightarrow 1)$ columns are i.i.d. standard $\mathcal{K}$-Gaussian.

Multiplying $\tilde{M}_{n,r}$ on the left and right by random elementary matrices we can successively:

- interchange the 1-st and $(n \Leftrightarrow r+1)$-st column of $\tilde{M}_{n,r}$ to produce a matrix $\tilde{M}'_{n,r}$;
- interchange the 1-st and m-th row of $\tilde{M}'_{n,r}$, where

  $$|\tilde{M}'_{n,r}(m,1)| = |\tilde{M}_{n,r}(m, n \Leftrightarrow r+1)| = 1,$$

  to produce a matrix $\tilde{M}''_{n,r}$ with $|\tilde{M}''_{n,r}(1,1)| = 1$ (note that the entries of $\tilde{M}''_{n,r}$ outside the first row and column are independent, and independent of those in the first row and column, those in columns 2 through $(n \Leftrightarrow r+1)$ are $\mathcal{K}$-Gaussian on $\rho\mathcal{D}$, and those in the remaining $(r \Leftrightarrow 1)$ columns are standard $\mathcal{K}$-Gaussian);

- subtract $\tilde{M}''_{n,r}(i,1)/\tilde{M}''_{n,r}(1,1)$ times the 1-st row of $\tilde{M}''_{n,r}$ from the i-th row of $\tilde{M}''_{n,r}$ for $2 \leq i \leq n$ to produce a matrix $\tilde{M}^*_{n,r}$ with
$$|\tilde{M}^*_{n,r}(1,1)| = 1,$$
and
$$\tilde{M}^*_{n,r}(2,1) = \cdots = \tilde{M}^*_{n,r}(n,1) = 0$$
(note that $\tilde{M}''_{n,r}(1,j) \in \rho\mathcal{D}$ for $2 \leq j \leq n \Leftrightarrow r+1$ so that the entries outside the first row and column of $\tilde{M}^*_{n,r}$ have the same distribution as those outside the first row and column of $\tilde{M}''_{n,r}$ and these entries are independent of the first row and column of $\tilde{M}^*_{n,r}$);
- subtract $\tilde{M}^*_{n,r}(1,j)/\tilde{M}^*_{n,r}(1,1)$ times the 1-st column of $M^*_{n,r}$ from the j-th column of $\tilde{M}^*_{n,r}$ to produce a matrix $\tilde{M}^{**}_{n,r}$ with
$$|M^{**}_{n,r}(1,1)| = 1,$$
$$\tilde{M}^{**}_{n,r}(2,1) = \cdots = \tilde{M}^{**}_{n,r}(n,1) = 0,$$
and
$$\tilde{M}^{**}_{n,r}(1,2) = \cdots = \tilde{M}^{**}_{n,r}(1,n) = 0$$
(note that $\tilde{M}^{**}_{n,r}(i,j) = \tilde{M}^*_{n,r}(i,j)$ for $2 \leq i,j \leq n$);
- multiply the 1-st row of $\tilde{M}^{**}_{n,r}$ by $\tilde{M}^{**}_{n,r}(1,1)^{-1}$ to produce a matrix with the same distribution as
$$\begin{pmatrix} 1 & 0 \\ 0 & M_{n-1,r-1} \end{pmatrix}.$$

The elementary divisors of $\tilde{M}_{n,r}$ thus have the same distribution as
$$(1, \rho^{K_0^{n-1,r-1}}, \rho^{K_1^{n-1,r-1}}, \ldots, \rho^{K_{j-1}^{n-1,r-1}}),$$
where
$$(\rho^{K_0^{n-1,r-1}}, \ldots, \rho^{K_{n-1}^{n-1,r-1}})$$
are the elementary divisors of $M_{n-1,r-1}$.

Putting these observations together gives the recursion

$$\begin{aligned} &\pi_{n,r}(l_0, l_1, \ldots) \\ (3.1) \qquad &= \begin{cases} q^{-n}\pi_{n,r-1}(l_0, l_1, \ldots) + (1 \Leftrightarrow q^{-n})\pi_{n-1,r-1}(l_0 \Leftrightarrow 1, l_1, \ldots), \\ \qquad 1 \leq r \leq n, \ l+_0 \geq 1, \\ q^{-n}\pi_{n,r-1}(0, l_1, \ldots), \ 1 \leq r \leq n, \ l_0 = 0 \end{cases} \end{aligned}$$

provided we adopt the convention $\pi_{0,0}(0,0,\ldots) = 1$. We also have the boundary conditions
$$\pi_{m,0}(l'_0, l'_1, \ldots) = 0, \ m \geq 1, \ l'_0 \geq 1,$$
$$\pi_{m,0}(0, l'_1, \ldots) = \pi_{m,m}(l'_1, l'_2, \ldots), \ m \geq 1,$$
because $M_{m,0}$ has the same distribution as $\rho M_{m,m}$.

Thus
$$\mathbb{P}\{L_0^n = l_0, L_1^n = l_1, \ldots\} = \gamma_{n,l_0}\mathbb{P}\{L_0^{n-l_0} = l_1, L_1^{n-l_0} = l_2, \ldots\}$$

for a certain constant γ_{n,l_0} if we adopt the convention $L_k^0 = 0$ for all k. Rewriting the recursion (3.1) as

$$\Pi_n^{-1}\pi_{n,r}(l_0,l_1,\dots) = \begin{cases} q^{-n}\Pi_n^{-1}\pi_{n,r-1}(l_0,l_1,\dots) + \Pi_{n-1}^{-1}\pi_{n-1,r-1}(l_0 \Leftrightarrow 1, l_1,\dots), \\ \qquad 1 \le r \le n,\ l_0 \ge 1, \\ q^{-n}\Pi_n^{-1}\pi_{n,r-1}(0,l_1,\dots),\ 1 \le r \le n,\ l_0 = 0 \end{cases}$$

it follows that the constant γ_{n,l_0} can be represented graphically as

$$\gamma_{n,l_0} = \frac{\Pi_n}{\Pi_{n-l_0}} \sum_\sigma q^{-\beta(\sigma)}$$

where the sum is over all planar lattice paths σ of length n from $(0,0)$ to $(n \Leftrightarrow l_0, l_0)$ that consist of steps of the form $(x,y) \to (x+1,h)$ or $(x,y) \to (x,y+1)$ and $\beta(\sigma)$ denotes the area in the plane above σ and below the line $\{(x,y) : 0 \le x \le n \Leftrightarrow l_0,\ y = n\}$.

As explained in Ch. 10 of [AAR99], the evaluation of this sum is a consequence of the non-commutative q-binomial theorem of [Sch53] (see also [Pól69]). Writing $\delta(\sigma)$ for the area below the path σ (and above the line $\{(x,y) : 0 \le x \le n \Leftrightarrow l_0,\ y = 0\}$),

$$\begin{aligned} \gamma_{n,l_0} &= \frac{\Pi_n}{\Pi_{n-l_0}} q^{-n(n-l_0)} \sum_\sigma q^{\delta(\sigma)} \\ &= \frac{\Pi_n}{\Pi_{n-l_0}} q^{-n(n-l_0)} \frac{(1 \Leftrightarrow q)\dots(1 \Leftrightarrow q^n)}{(1 \Leftrightarrow q)\dots(1 \Leftrightarrow q^{n-l_0})(1 \Leftrightarrow q)\dots(1 \Leftrightarrow q^{l_0})} \\ &= \frac{\Pi_n}{\Pi_{n-l_0}} q^{-n(n-l_0)} q^{l_0(n-l_0)} \begin{bmatrix} n \\ n \Leftrightarrow l_0 \end{bmatrix} \\ &= q^{-(n-l_0)^2} \frac{\Pi_n}{\Pi_{n-l_0}} \begin{bmatrix} n \\ n \Leftrightarrow l_0 \end{bmatrix}. \end{aligned}$$

The theorem now follows immediately. □

Remark 3.6. It follows from Theorem 3.5 that the joint distribution of the elementary divisors $(\rho^{K_1^n}, \rho^{K_2^n}, \dots, \rho^{K_n^n})$ of M_n conditional on $L_0^n = l_0$ does not depend on n. The following corollary is immediate.

Corollary 3.7. *As $n \to \infty$, the sequence of $\mathbb{N}$-valued random variables $n \Leftrightarrow L_0^n$, $n \Leftrightarrow L_0^n \Leftrightarrow L_1^n$, $n \Leftrightarrow L_0 \Leftrightarrow L_1 \Leftrightarrow L_2, \dots$ converges in distribution to a Markov chain on $\mathbb{N}$ with initial distribution*

$$p(s) = q^{-s^2} \frac{\Pi_\infty}{\Pi_s^2}, \quad s \ge 0,$$

and transition matrix $P(\cdot,\cdot)$ as in Theorem 3.5.

Remark 3.8. i) A consequence of Corollary 3.7 is that $n \Leftrightarrow L_0^n$, the number of elementary divisors of M_n that are not 1, converges in distribution. Also, $\max\{k : L_k^n \neq 0\}$ converges in distribution, so that K_n^n converges in distribution. Moreover,

$$|\det M_n| = q^{-(K_1^n + \cdots + K_n^n)} = q^{-\Sigma_k k L_k^n}$$

also converges in distribution to an almost surely strictly positive limit. We obtain the distribution of $\det M_n$ and the limit explicitly in the next section.
ii) It follows from Corollary 3.7 that

$$\sum_{s=0}^\infty \frac{q^{-s^2}}{\Pi_s^2} = \frac{1}{\Pi_\infty}.$$

This identity is due to Euler (see the comments after Corollary 10.9.4 in [AAR99]). iii) Consider the group $GL(m,q)$ of $m \times m$ invertible matrices over the finite field $\mathbb{F}_q$ with q elements. The conjugacy class of an element of $GL(m,q)$ is determined by its rational canonical form. The form is in turn specified by a map $\phi \mapsto \lambda_\phi$ from the set of monic non-constant irreducible polynomials over $\mathbb{F}_q$ into the set of partititions of non-negative integers. The only restrictions on the λ_ϕ are that $|\lambda_z| = 0$ (that is, the partition corresponding to the polynomial $z \mapsto z$ is the unique partition of the integer 0) and that $\sum_\phi \deg(\phi)|\lambda_\phi| = m$. The map $\phi \mapsto \lambda_\phi$ for a uniformly chosen group element is studied in great detail in [Ful02]. In particular, as $m \to \infty$ the random partition $\lambda_{z-1,1} \ge \lambda_{z-1,2} \ge \ldots$ associated with the polynomial $z \mapsto z \Leftrightarrow 1$ for a uniformly chosen group element converges in distribution to the Markov chain of Corollary 3.7. More generally, the random partition associated with a monic irreducible polynomial of degree d converges to a similarly defined chain with q replaced by q^d, and these random partitions are asympotically independent.

4. Distribution of the determinant

The obvious approach to obtaining the distribution of the determinant of M_n is to note for $n \ge 2$ and $h \ge 0$ that we have the recursion

$$\begin{aligned}
\mathbb{P}\{|\det M_n| = q^{-h}\} &= \mathbb{P}\left\{ |\det M_n| = q^{-h} \;\middle|\; \bigvee_{i=1}^{n} |M_n(i,1)| \le q^{-1} \right\} q^{-n} \\
&\quad + \mathbb{P}\left\{ |\det M_n| = q^{-h} \;\middle|\; \bigvee_{i=1}^{n} |M_n(i,1)| = 1 \right\} (1 \Leftrightarrow q^{-n}) \\
&= \mathbb{P}\{|\det M_n| = q^{-(h-1)}\} q^{-n} \\
&\quad + \mathbb{P}\{|\det M_{n-1}| = q^{-h}\}(1 \Leftrightarrow q^{-n}),
\end{aligned}$$

because conditioned on the event $\left\{\bigvee_{i=1}^n |M_n(i,1)| \le q^{-1}\right\}$ the conditional distribution of M_n is that of the random matrix obtained by multiplying the first column of M_n by ρ, and conditioned on the event $\{\bigvee_{i=1}^n |M_n(i,1)| = 1\}$ the conditional distribution of $|\det M_n|$ is that of

$$\left| \det \begin{pmatrix} 1 & 0 \\ 0 & M_{n-1} \end{pmatrix} \right| .$$

We also have the boundary values

$$\mathbb{P}\{|\det M_1| = q^{-h}\} = q^{-h} \Leftrightarrow q^{-(h+1)}, \quad h \ge 0,$$

and, by Theorem 3.5,

$$\mathbb{P}\{|\det M_n| = 1\} = \mathbb{P}\{L_0^n = n,\, 0 = L_1^n = L_2^n = \ldots\} = \Pi_n, \quad n \ge 1.$$

This sort of recursion is solved in Theorem 2.1 of [BM87], apparently by guessing and then verifying the form of the solution. The solution in [BM87] can be manipulated to establish the following result. We provide a somewhat more illuminating proof that follows the first proof of Theorem 1 in [AG00].

Theorem 4.1. *For $h \ge 0$,*

$$\mathbb{P}\{|\det M_n| = q^{-h}\} = \Pi_n q^{-h} \begin{bmatrix} n + h \Leftrightarrow 1 \\ h \end{bmatrix} .$$

Consequently, the distribution of $\det M_n$ *has density*

$$x \mapsto \frac{1 \Leftrightarrow q^{-n}}{1 \Leftrightarrow q^{-1}}(1 \Leftrightarrow |x|q^{-1}) \dots (1 \Leftrightarrow |x|q^{-(n-1)}), \quad x \in \mathcal{D},$$

with respect to λ.

Proof. Set

$$I = \min \left\{ 1 \le i \le n : |M_n(i,1)| = \bigvee_{i'=1}^{n} |M_n(i',1)| \right\}.$$

Note that

$$\mathbb{P}\{|M_n(I,1)| = q^{-h}\} = (q^{-hn} \Leftrightarrow q^{-(h+1)n}), \quad h \ge 0.$$

Multiplying M_n be random elementary matrices we can successively:

- interchange the 1-st and I-th row of M_n to produce a random matrix M_n';
- subtract $M_n'(i,1)/M_n'(1,1)$ times the first row of M_n' from the i-th row of M_n' for $2 \le i \le n$ to produce a random matrix M_n'';
- divide the first row of $M_n''(1,1)/\rho^H$, where $|M_n''(1,1)| = |M_n'(1,1)| = |M_n(I,1)| = q^{-H}$.

This produces a random matrix with the same distribution as

$$\begin{pmatrix} \rho^H & Z_{n-1} \\ 0 & M_{n-1} \end{pmatrix},$$

where H, Z_{n-1} and M_{n-1} are independent and Z_{n-1} is a vector of $n \Leftrightarrow 1$ standard $\mathcal{K}$-Gaussian random variables. In particular, $|\det M_n|$ has the same distribution as $q^{-H}|\det M_{n-1}|$.

Thus

$$\begin{aligned} \mathbb{P}\{|\det M_n| = q^{-h}\} &= \sum_{h_1 + \cdots + h_n = h} \prod_{i=1}^{n} (q^{-h_i(n+1-i)} \Leftrightarrow q^{-(h_i+1)(n+1-i)}) \\ &= \Pi_n \sum_{h_1 + \cdots + h_n = h} q^{-(nh_1 + (n-1)h_2 + \cdots + h_n)} \\ &= \Pi_n \sum_{l=0}^{\infty} Q(n,h,l) q^{-l}, \end{aligned}$$

where $Q(n,h,l)$ is the number of partitions of l into h parts, each of size at most n. The claimed distribution for $|\det M_n|$ then follows from the known generating function of the partition counts (see, for example, Theorem 11.4.4 of [AAR99]).

Note that multiplying the first column of M_n by a constant $c \in \mathcal{D}$ with $|c| = 1$ gives a matrix with the same distribution as M_n, so that $\det M_n$ and $c \det M_n$ have the same distribution. Consequently, the distribution of $\det M_n$ has a density at $x \in \mathcal{D}$ given by

$$\frac{\mathbb{P}\{|\det M_n| = |x|\}}{\lambda\{y : |y| = |x|\}},$$

and this is readily seen to be given by the stated formula. □

Remark 4.2. It follows from Theorem 4.1 that

$$\sum_{h=0}^{\infty} q^{-h} \begin{bmatrix} n + h \Leftrightarrow 1 \\ h \end{bmatrix} = \frac{1}{\Pi_n}.$$

This identity is a classical consequence of the (commutative) q-binomial theorem — see Corollary 10.2.2(d) of [AAR99].

Corollary 4.3. *As $n \to \infty$, the distribution of* $\det M_n$ *converges to a probability measure on $\mathcal{D}$ with density against λ given by*

$$x \mapsto \frac{1}{(1 \Leftrightarrow q^{-1})}(1 \Leftrightarrow |x|q^{-1})(1 \Leftrightarrow |x|q^{-2})\dots, \quad x \in \mathcal{D}.$$

Combining Theorem 3.5 and Theorem 4.1 gives the following identity.

Corollary 4.4. *For $n \geq 1$ and $h \geq 0$,*

$$\sum q^{-((n-l_0)^2+(n-l_0-l_1)^2+\cdots)}\frac{\Pi_n}{\Pi_{l_0}\Pi_{l_1}\dots} = q^{-h}\begin{bmatrix} n+h \Leftrightarrow 1 \\ h \end{bmatrix},$$

where the sum is over all $l_0, l_1, \dots$ such that $\sum_k l_k = n$ and $\sum_k k l_k = h$.

Acknowledgment: The author thanks Jason Fulman and the referee for pointing out that the Markov chain in Section 3 also appears in Fulman's work on the conjugacy class of a random element of $GL(m,q)$.

References

[AAR99] George E. Andrews, Richard Askey, and Ranjan Roy. *Special functions.* Cambridge University Press, Cambridge, 1999.

[AG00] Khaled A. S. Abdel-Ghaffar. The determinant of random power series matrices over finite fields. *Linear Algebra Appl.*, 315(1-3):139–144, 2000.

[Bal68] G. V. Balakin. The distribution of the rank of random matrices over a finite field. *Teor. Verojatnost. i Primenen.*, 13:631–641, 1968.

[BKW97] Johannes Blömer, Richard Karp, and Emo Welzl. The rank of sparse random matrices over finite fields. *Random Structures Algorithms*, 10(4):407–419, 1997.

[BM87] Richard P. Brent and Brendan D. McKay. Determinants and ranks of random matrices over Z_m. *Discrete Math.*, 66(1-2):35–49, 1987.

[BW87] Joseph P. Brennan and John Wolfskill. Remarks: "On the probability that the determinant of an $n \times n$ matrix over a finite field vanishes" [Discrete Math. **51** (1984), no. 3, 311–315; MR 86a:15012] by A. Mukhopadhyay. *Discrete Math.*, 67(3):311–313, 1987.

[Coo00a] C. Cooper. On the distribution of rank of a random matrix over a finite field. In *Proceedings of the Ninth International Conference "Random Structures and Algorithms" (Poznan, 1999)*, volume 17, pages 197–212, 2000.

[Coo00b] C. Cooper. On the rank of random matrices. *Random Structures Algorithms*, 16(2):209–232, 2000.

[Eva01] Steven N. Evans. Local fields, Gaussian measures, and Brownian motions. In *Topics in probability and Lie groups: boundary theory*, pages 11–50. Amer. Math. Soc., Providence, RI, 2001.

[Ful99] Jason Fulman. Cycle indices for the finite classical groups. *J. Group Theory*, 2(3):251–289, 1999.

[Ful02] Jason Fulman. Random matrix theory over finite fields. *Bull. Amer. Math. Soc. (N.S.)*, 39(1):51–85 (electronic), 2002.

[HS93] Jennie C. Hansen and Eric Schmutz. How random is the characteristic polynomial of a random matrix? *Math. Proc. Cambridge Philos. Soc.*, 114(3):507–515, 1993.

[Jac85] Nathan Jacobson. *Basic algebra. I.* W. H. Freeman and Company, New York, second edition, 1985.

[Koz66] M. V. Kozlov. On the rank of matrices with random Boolean elements. *Soviet Math. Dokl.*, 7:1048–1051, 1966.

[Kun81] Joseph P. S. Kung. The cycle structure of a linear transformation over a finite field. *Linear Algebra Appl.*, 36:141–155, 1981.

[Lev91] A. A. Levit·skaya. Limit distribution of the rank of a random Boolean matrix with a fixed number of units in the rows. *Dokl. Akad. Nauk Ukrain. SSR*, (8):49–52, 183, 1991.

[Muk84] A. Mukhopadhyay. On the probability that the determinant of an $n \times n$ matrix over a finite field vanishes. *Discrete Math.*, 51(3):311–315, 1984.

[Pól69] G. Pólya. On the number of certain lattice polygons. *J. Combinatorial Theory*, 6:102–105, 1969.

[Raw97] Don Rawlings. Absorption processes: models for q-identities. *Adv. in Appl. Math.*, 18(2):133–148, 1997.

[Raw98] Don Rawlings. A probabilistic approach to some of Euler's number-theoretic identities. *Trans. Amer. Math. Soc.*, 350(7):2939–2951, 1998.

[Sch53] Marcel Paul Schützenberger. Une interprétation de certaines solutions de l'équation fonctionnelle: $F(x+y) = F(x)F(y)$. *C. R. Acad. Sci. Paris*, 236:352–353, 1953.

[Sch84] W. H. Schikhof. *Ultrametric calculus*. Cambridge University Press, Cambridge, 1984. An introduction to p-adic analysis.

[Sto88] Richard Stong. Some asymptotic results on finite vector spaces. *Adv. in Appl. Math.*, 9(2):167–199, 1988.

[Tai75] M. H. Taibleson. *Fourier analysis on local fields*. Princeton University Press, Princeton, N.J., 1975.

[Wat87] William C. Waterhouse. How often do determinants over finite fields vanish? *Discrete Math.*, 65(1):103–104, 1987.

E-mail address: `evans@stat.Berkeley.EDU`

Department of Statistics #3860, University of California at Berkeley, 367 Evans Hall, Berkeley, CA 94720-3860, U.S.A