Journal of Algebraic Combinatorics 1 (1992), 133–150

Generating Random Elements in $SL_n(\mathbf{F}_q)$ by Random Transvections

MARTIN HILDEBRAND*
Department of Mathematics, University of Michigan, Ann Arbor, MI 48109–1003

Received June 10, 1991; Revised February 3, 1992

Abstract. This paper studies a random walk based on random transvections in $SL_n(\mathbf{F}_q)$ and shows that, given $\epsilon > 0$, there is a constant c such that after $n + c$ steps the walk is within a distance ϵ from uniform and that after $n - c$ steps the walk is a distance at least $1 - \epsilon$ from uniform. This paper uses results of Diaconis and Shahshahani to get the upper bound, uses results of Rudvalis to get the lower bound, and briefly considers some other random walks on $SL_n(\mathbf{F}_q)$ to compare them with random transvections.

Keywords: transvection, random walk, representation theory, upper bound lemma

1. Introduction

Diaconis and Shahshahani [3] have studied a random walk on S_n, the symmetric group on n elements. This walk involves at each step picking two (possibly nondistinct) elements at random and transposing them if they are distinct. The techniques they used are relatively general, and Diaconis [2] suggested some other processes to which this technique may be applied. The present paper studies one such process, random transvections on $SL_n(\mathbf{F}_q)$, and finds an unusually sharp cutoff phenomenon.

$SL_n(\mathbf{F}_q)$ is the group of $n \times n$ matrices with elements in $\mathbf{F}_q$, a finite field with q elements, and determinant 1. Suzuki [10] defined a transvection on $SL_n(\mathbf{F}_q)$ as an element which is not the identity but does fix all the points in a hyperplane of $(\mathbf{F}_q)^n$. An example of a transvection is $I + aE_{ij}$, where I is the identity, $a \in \mathbf{F}_q^*$ (the multiplicative group of $\mathbf{F}_q$), and E_{ij} is an $n \times n$ matrix with the only nonzero entry being 1 in the (i, j)th position. Transvections are basic building blocks for working in matrix groups, just as transpositions are for permutations (see [1] for an example). If $n > 2$, the transvections form a conjugacy class. The transvections generate $SL_n(\mathbf{F}_q)$ (see [10]).

We wish to pick a transvection at random. We can do so even without enumerating the transvections. A transvection can be represented as a linear transformation $\lambda \mapsto \lambda + f(\lambda)a$, where λ is in $\mathbf{F}_q^n$, a is a nonzero vector in $\mathbf{F}_q^n$, f

*Supported in part by a Rackham Faculty Fellowship at the University of Michigan.

is a nonzero linear transformation $f : \mathbf{F}_q^n \to \mathbf{F}_q$, and $f(a) = 0$. In coordinates a transvection can be represented as $I + b^t a$ with a, b nonzero vectors in $\mathbf{F}_q^n$ such that $ba^t = 0$. Note that if $b^t a = b'^t a'$, then $b = cb'$ and $a' = ca$ for some $c \in \mathbf{F}_q^*$. Thus picking such a and b at random enables us to pick a random transvection.

Pick m transvections with independent identical distributions each of which is uniform over the transvections, and multiply them to get an element of $SL_n(\mathbf{F}_q)$. The Markov process consisting of multiplying a matrix in $SL_n(\mathbf{F}_q)$ by a random transvection is doubly stochastic and hence has the uniform distribution on $SL_n(\mathbf{F}_q)$ for its stationary distribution (see [6] for more details on Markov processes). Thus if m is large enough and if there is no parity problem, then the product of m random transvections will be nearly uniform on $SL_n(\mathbf{F}_q)$. The question we ask is how large does m (as a function of n) have to be for this product to get close to uniform on $SL_n(\mathbf{F}_q)$.

We define the variation distance of a probability distribution P on a finite group G from the uniform distribution U on G by

$$\| P - U \| := \frac{1}{2} \sum_{s \in G} |P(s) - \frac{1}{|G|}|.$$

It is easy to show that

$$\| P - U \| = \max_{A \subseteq G} |P(A) - U(A)|.$$

Let P^{*m} be the probability distribution of the product of m random transvections. The main goal of this paper is to show

THEOREM 1.1. *There exist positive constants A and k such that*

$$\| P^{*m} - U \| < Ae^{-kc}$$

for sufficiently large n and for all $c > 0$, where $c = m - n$.

If $n = 2$, there are parity problems in the case $q = 2$. These problems do not occur if $n > 2$.

A secondary goal is to prove

THEOREM 1.2. *Given $\epsilon > 0$, there exists $c > 0$ such that $\| P^{*m} - U \| > 1 - \epsilon$ for $m = n - c$ and sufficiently large n.*

To prove Theorem 1.1, we use

UPPER BOUND LEMMA (Diaconis and Shahshahani).

$$\| P - U \|^2 \leq \frac{1}{4} \sum\nolimits^{*} d_\rho \mathrm{Tr}(\hat{P}(\rho)\, \hat{P}(\rho)^*),$$

*where the * of a matrix is its conjugate transpose,* $\overset{*}{\sum}$ *means the sum over all nontrivial irreducible representations* ρ *of* G, d_ρ *is the degree of* ρ, *and* $\hat{P}(\rho)$ *is the Fourier transform of* ρ:

$$\hat{P}(\rho) := \sum_{s \in G} P(s)\rho(s).$$

The lemma results from the Cauchy–Schwarz inequality and the Plancherel theorem. This lemma is discussed further in [2], and [2] and [9] present some background from representation theory of finite groups.

This paper uses the representation theory of $GL_n(\mathbf{F}_q)$. We may draw conclusions for the random walk on $SL_n(\mathbf{F}_q)$ due to

LEMMA 1.1. *The distance of the probability distribution of* $T_1 \cdots T_m D$ *from uniform in* $GL_n(\mathbf{F}_q)$ *equals the distance of the probability distribution of* $T_1 \cdots T_m$ *from uniform in* $SL_n(\mathbf{F}_q)$, *where* T_i, $i = 1, \ldots, m$, *are independent random transvections and* D *is an* $n \times n$ *diagonal matrix with the lower-right-hand corner a random element of* $\mathbf{F}_q^*$ *and the other diagonal elements 1.*

The proof of Lemma 1.1 is straightforward and is left to the reader.

Let P_1 be the probability distribution of random transvections in $GL_n(\mathbf{F}_q)$, let Q be the probability distribution of D, and let R be the probability distribution of $T_1 \cdots T_m D$ in $GL_n(\mathbf{F}_q)$. Since P_1 is constant on conjugacy classes, $\hat{P}_1(\rho)$ is a constant times the identity (see [2]). By taking the trace, we see that

$$\hat{P}_1(\rho) = \frac{\chi_\rho(\tau)}{d_\rho} I,$$

where $\chi_\rho(\tau)$ is the character of ρ on the transvections. We can express $\hat{R}$ in terms of $\hat{P}_1$ and $\hat{Q}$:

$$\hat{R}(\rho) = \hat{Q}(\rho)\, \hat{P}_1(\rho)^m$$

(see [2]).

Thus we may conclude

LEMMA 1.2.

$$\| R - U \|^2 \leq \frac{1}{4} \overset{*}{\sum} d_\rho \left| \frac{\chi_\rho(\tau)}{d_\rho} \right|^{2m} \mathrm{Tr}(\hat{Q}(\rho)\, \hat{Q}(\rho)^*).$$

To use this expression, we need to find $\chi_\rho(\tau)$ and d_ρ. Using Macdonald's book [7], we do so in Section 2. In Section 3 we deal with the factor $\mathrm{Tr}(\hat{Q}(\rho)\, \hat{Q}(\rho)^*)$. In Section 4 we put some bounds on $\chi_\rho(\tau)/d_\rho$, and in Section 5 we use these bounds to prove Theorem 1.1. In Section 6 we prove Theorem 1.2. In Section 7 we consider some other random processes on $SL_n(\mathbf{F}_q)$.

2. The characters of $GL_n(\mathbf{F}_q)$ on transvections

The characters of $GL_n(\mathbf{F}_q)$ have been determined in earlier work. Both Green [5] and Zelevinsky [11] have provided expressions which determine the value of characters of $GL_n(\mathbf{F}_q)$ on transvections, but these expressions do not seem to be directly useful for the asymptotics we want. We shall develop a different expression which works well with asymptotics. In doing so, our notation will follow that of Macdonald [7].

Let k be a finite field. Let Φ be the set of all irreducible monic polynomials in $k[t]$ except for the polynomial t. Each conjugacy class of $GL_n(k)$ corresponds to a partition-valued function μ on Φ such that

$$\| \mu \| := \sum_{f \in \Phi} d(f) \, |\mu(f)| = n.$$

Furthermore, each partition-valued function μ on Φ such that $\| \mu \| = n$ determines a conjugacy class.

By using the Jordan canonical form for the matrix [7, p. 140], we can determine which conjugacy class a given partition-valued function μ corresponds to.

If μ corresponds to the conjugacy class of the identity, then

$$\mu = \begin{cases} (1^n) & \text{if } f = f_1 \\ 0 & \text{otherwise,} \end{cases}$$

where $f_1 = t - 1$.

If μ corresponds to the conjugacy class of a transvection, then

$$\mu = \begin{cases} (21^{n-2}) & \text{if } f = f_1 \\ 0 & \text{otherwise.} \end{cases}$$

To see this, just note that

$$\begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

is a transvection and is in Jordan canonical form.

Let k_n be the unique extension of k of degree n in $\bar{k}$, let M_n be the multiplicative group of k_n, and let $\hat{M}_n$ be the character group of M_n. Define $L = \varinjlim \hat{M}_n$. Observe that $F : \xi \mapsto \xi^q$ acts on L. Let Θ denote the set of F-orbits in L. The irreducible characters of $GL_n(k)$ are in a bijective correspondence with partition-valued functions $\boldsymbol{\lambda}$ of Θ with

$$\| \boldsymbol{\lambda} \| := \sum_{\varphi \in \Theta} d(\varphi) \, |\boldsymbol{\lambda}(\varphi)| = n.$$

Let $\boldsymbol{\mu}$ correspond to the conjugacy class of the transvections, let $\boldsymbol{\lambda}$ correspond to an irreducible representation, let $d_{\boldsymbol{\lambda}}$ be the degree of the representation, and let $\chi_{\boldsymbol{\mu}}^{\boldsymbol{\lambda}}$ be the value of this representation's character on the transvections. We wish to show

THEOREM 2.1.

$$\frac{\chi_{\boldsymbol{\mu}}^{\boldsymbol{\lambda}}}{d_{\boldsymbol{\lambda}}} = \frac{q^{(n^2/2)-(n/2)+1}}{q^{n(n+1)/2}} \frac{\varphi_{n-2}(q^{-1})\varphi_1(q^{-1})}{\varphi_n(q^{-1})} \sum_{\boldsymbol{\lambda}_1 \in B(\boldsymbol{\lambda})} \frac{\delta(S_{\boldsymbol{\lambda}_1})(q-1)^{-1}}{\delta(S_{\boldsymbol{\lambda}})}$$

$$- \frac{q^{(n^2/2)-(n/2)+1}}{q^{n(n+1)/2}} \frac{\varphi_{n-2}(q^{-1})\varphi_1(q^{-1})}{\varphi_n(q^{-1})} \sum_{i=0}^{n-1} q^{-i},$$

where $B(\boldsymbol{\lambda}) = \{\boldsymbol{\lambda}_1 : |\boldsymbol{\lambda}_1(\varphi_a)| = |\boldsymbol{\lambda}(\varphi_a)| - 1$ *for some* $\varphi_a \in \Phi$ *of degree* 1, $\boldsymbol{\lambda}_1(\varphi) \subset \boldsymbol{\lambda}(\varphi)$ *for all* $\varphi \in \Phi$, *and* $\| \boldsymbol{\lambda}_1 \| = \| \boldsymbol{\lambda} \| - 1\}$ *and* $\varphi_m(t) = \prod_{i=1}^{m}(1 - t^i)$. *(Do not confuse the function* $\varphi_m(t)$ *with elements of* Φ.*) As in Macdonald,*

$$\delta(S_{\boldsymbol{\lambda}}) = \prod_{\varphi \in \Phi} \delta(s_{\boldsymbol{\lambda}(\varphi)}(\varphi))$$

with

$$\delta(s_{\lambda}(\varphi)) = q_{\varphi}^{n(\lambda')} \tilde{H}_{\lambda}(q_{\varphi})^{-1},$$

$n(\lambda) = \sum (i-1)\lambda_i$, $q_{\varphi} = q^{d(\varphi)}$, *and* $\tilde{H}_{\lambda}(q_{\varphi}) = \prod_{x \in \lambda}(q_{\varphi}^{h(x)} - 1)$, *where* $h(x)$ *is the hook length of* x.

To find this ratio, we shall use two results from [7, p. 151].

The degree of the irreducible representation is given by

$$d_{\lambda} = \psi_n(q) \prod_{\varphi \in \Phi} q_{\varphi}^{n(\lambda(\varphi)')} \tilde{H}_{\lambda(\varphi)}(q_{\varphi})^{-1}, \tag{2.1}$$

where $\psi_n(q) = \prod_{i=1}^{n}(q^i - 1)$.

Two symmetric functions are related by characters of $GL_n(\mathbf{F}_q)$:

$$\tilde{Q}_{\mu} = \sum_{\lambda : \|\lambda\| = n} \chi_{\mu}^{\lambda} \overline{S_{\lambda}}, \tag{2.2}$$

where the symmetric functions are as in [7]; the conjugate $\bar{u}$ of $u = \sum u_\mu \tilde{P}_\mu$ is $\bar{u} := \bar{u}_\mu \tilde{P}_\mu$. Since $P_\mu(x; t)P_\nu(x; t) = \sum_\lambda f^\lambda_{\mu\nu}(t)P_\lambda(x; t)$, where $f^\lambda_{\mu\nu} \in \mathbf{Z}[t]$ (see [7, p. 110]), $\bar{u}\,\bar{v} = \overline{uv}$.

Let $\boldsymbol{\mu}$ correspond to the conjugacy class of the transvections. By definition,

$$\begin{aligned} \tilde{Q}_\mu &= \tilde{Q}_{(21^{n-2})}(f_1) \\ &= a_{(21^{n-2})}(q_{f_1})q_{f_1}^{-n(21^{n-2})}P_{(21^{n-2})}(X_{f_1}; q_{f_1}^{-1}) \\ &= a_{(21^{n-2})}(q)q^{-n(21^{n-2})}P_{(21^{n-2})}(X_{f_1}; q^{-1}), \end{aligned} \tag{2.3}$$

where $a_\lambda(q)$ and $P_\lambda(x; t)$ are as in [7].

The transition matrix between the Hall–Littlewood functions and the Schur functions gives

$$P_{(21^{n-2})}(X_{f_1}; t) = \sum_{\text{partitions } \nu} w_{(21^{n-2})\nu}(t)s_\nu(X_{f_1}).$$

The functions $w_{(21^{n-2})\nu}(t)$ are given as follows:

LEMMA 2.1.

$$w_{(21^{n-2})\nu}(t) = \begin{cases} 1 & \text{if } \nu = (21^{n-2}) \\ -\sum_{i=1}^{n-1} t^i & \text{if } \nu = (1^n) \\ 0 & \text{otherwise.} \end{cases}$$

Proof. All cases except $\nu = (1^n)$ are shown in [7, p. 105], where it is shown that the transition matrix is strictly upper unitriangular.

Let $K(t) = M(s, P)$ be the transition matrix between the s_λ's and the P_λ's. (Note that here λ is a partition. The boldface $\boldsymbol{\lambda}$ denotes the partition-valued functions.)

The $w_{\lambda\mu}$'s are coefficients in $M(P, s) = K(t)^{-1}$. Observe that $K(t)$ and $K(t)^{-1}$ are strictly upper unitriangular.

$$w_{(21^{n-2})(21^{n-2})}(t)K_{(21^{n-2})(1^n)}(t) + w_{(21^{n-2})(1^n)}(t)K_{(1^n)(1^n)}(t) = 0.$$

Thus $-K_{(21^{n-2})(1^n)}(t) = w_{(21^{n-2})(1^n)}(t)$.

By a theorem of Lascoux and Schützenberger,

$$K_{\lambda\mu}(t) = \sum_T t^{c(T)},$$

where the sum is over all tableaux T of shape λ and weight μ (see [7, p. 129]). $c(T)$ is the charge of T, which is defined as follows.

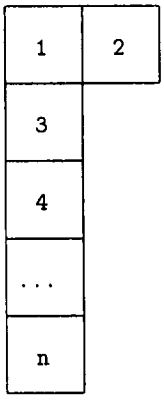

Fig. 1. Tableau illustrating the definition of a word.

One defines the word of a tableau by reading the numbers in the tableau from right to left and then from top to bottom. For instance, the word of the tableau displayed in Figure 1 is $2134\ldots n$. If w is a standard word, i.e., contains the numbers 1 through n exactly once (and it will be so for all tableaux of shape (21^{n-2}) and weight (1^n)), attach an index to each element of w. The number 1 has index 0. If r has index i, then $r+1$ has index i or $i+1$ according to whether it lies to the right or left of r. $c(w)$ is defined to be the sum of the indices. Here it is $n-1$ because the numbers 2 through n each have index 1. Here $c(T)$ is just defined to be $c(w)$.

The upper-left corner of a tableau with shape (21^{n-2}) and weight (1^n) is always 1. The upper-right corner can take on any value x between 2 and n. The remaining $n-2$ elements increase as one goes down the column, and no values occur twice since the weight is (1^n). The word is thus $x12\ldots\ \hat{x}\ \ldots n$, where $\hat{x}$ means omit x. Thus $c(w) = n - x + 1 = n - (x-1)$. So $K_{(21^{n-2})(1^n)} = \sum_{i=1}^{n-1} t^i$. This completes the proof of the lemma. □

Since $s_\nu = \det(e_{\nu_i'-i+j})_{1\le i,j\le m}$, we may conclude

COROLLARY.

$$\begin{aligned} P_{(21^{n-2})}(X_{f_1}; t) &= e_{n-1}(f_1)e_1(f_1) - e_n(f_1) + \left(-\sum_{i=1}^{n-1} t^i\right) e_n(f_1) \\ &= e_{n-1}(f_1)e_1(f_1) - \left(\sum_{i=0}^{n-1} t^i\right) e_n(f_1). \end{aligned} \tag{2.4}$$

The following lemma describes the transition between $e_{n-1}(f_1)e_1(f_1)$ and S_λ:

LEMMA 2.2.

$$e_{n-1}(f_1)e_1(f_1) = \sum_{\lambda:\|\lambda\|=n} \gamma(S_\lambda)\overline{S_\lambda},$$

where

$$\gamma(S_\lambda) = \sum_{\lambda_1 \in B(\lambda)} \delta(S_{\lambda_1})(q-1)^{-1}$$

and $B(\lambda)$ *is as in the statement of Theorem 2.1.*

Proof. By examining (2.1) and (2.2), observe that

$$e_k(f_1) = \sum_{\lambda:\|\lambda\|=k} \gamma(S_\lambda)\overline{S_\lambda}. \tag{2.5}$$

Thus

$$e_{n-1}(f_1)e_1(f_1) = \left(\sum_{\lambda_1:\|\lambda_1\|=n-1} \delta(S_{\lambda_1})\overline{S_{\lambda_1}}\right)\left(\sum_{\lambda_2:\|\lambda_2\|=1} \delta(S_{\lambda_2})\overline{S_{\lambda_2}}\right).$$

Since $(a\overline{S_{\lambda_1}})(b\overline{S_{\lambda_2}}) = ab\overline{(S_{\lambda_1})(S_{\lambda_2})}$, we may conclude

$$e_{n-1}(f_1)e_1(f_1) = \sum_{\lambda_1:\|\lambda_1\|=n-1} \sum_{\lambda_2:\|\lambda_2\|=1} \delta(S_{\lambda_1})\delta(S_{\lambda_2})\overline{S_{\lambda_1}S_{\lambda_2}}.$$

There are $q-1$ partition-valued functions λ_2 such that $\|\lambda_2\| = 1$. Let $\varphi_a \in \Phi$ be such that $d(\varphi_a) = 1$ and $\lambda_2(\varphi_a) \neq 0$. It is known that

$$s_{\lambda_1(\varphi_a)}(\varphi_a)s_{\lambda_2(\varphi_a)}(\varphi_a) = \sum_{\text{partitions } \lambda} c^{\lambda}_{\lambda_1(\varphi_a)\lambda_2(\varphi_a)}s_\lambda(\varphi_a).$$

To find the values $c^{\lambda}_{\lambda_1(\varphi_a)\lambda_2(\varphi_a)}$, apply the Littlewood-Richardson rule. It says that $c^{\lambda}_{\lambda_1(\varphi_a)\lambda_2(\varphi_a)}$ is the number of tableaux T of shape $\lambda - \lambda_1(\varphi_a)$ and weight $\lambda_2(\varphi_1)$ such that $w(T)$ is a lattice permutation. If $|\lambda| \neq |\lambda_1(\varphi_a)| + |\lambda_2(\varphi_a)|$ or $\lambda_1(\varphi_a) \not\subset \lambda$, then $c^{\lambda}_{\lambda_1(\varphi_a)\lambda_2(\varphi_a)} = 0$. Otherwise, since $\lambda_2(\varphi_a) = (1)$, there is exactly one such tableau.

Thus $S_{\lambda_1}S_{\lambda_2} = \sum_{\lambda\in A(\lambda_1)} S_\lambda$, where $A(\lambda_1) := \{\lambda : \lambda(\varphi_a) \supset \lambda_1(\varphi_a), |\lambda(\varphi_a)| - |\lambda_1(\varphi_a)| = 1, \text{ and } \lambda_1(\varphi) = \lambda(\varphi) \text{ if } \varphi \neq \varphi_a\}$.

Lemma 2.2 follows from this statement. □

Theorem 2.1 follows from Lemma 2.2 and equations (2.3), (2.4), and (2.5).

3. Bounds on $\mathrm{Tr}(\hat{Q}(\rho)\,\hat{Q}(\rho)^*)$

The following two lemmas enable us to bound $\mathrm{Tr}(\hat{Q}(\rho)\,\hat{Q}(\rho)^*)$ from Lemma 1.2.

LEMMA 3.1. *If ρ is an irreducible representation of $GL_n(\mathbf{F}_q)$ and Q is as in Section 1, then*

$$\mathrm{Tr}(\hat{Q}(\rho)\,\hat{Q}(\rho)^*) \leq d_\rho.$$

Proof. Observe that ρ on the matrices M such that $Q(M) > 0$ may be viewed as a representation on $\mathbf{F}_q^*$. Thus in some basis,

$$\hat{Q}(\rho) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{d_\rho} \end{pmatrix}.$$

Because $\det(D)$ is uniform on $\mathbf{F}_q^*$, $a_i \in \{0, 1\}$ for $i = 1, 2, \ldots, d_\rho$. The lemma follows. □

For certain representations, we can improve the result:

LEMMA 3.2. *If $\lambda(\varphi_a) = (1^n)$ and λ is nontrivial, then* $\mathrm{Tr}(\hat{Q}(\lambda)\,\hat{Q}(\lambda)^*) = 0$.

Proof. It can be readily shown that in this case $\chi_\mu^\lambda = d_\lambda = 1$. From the proof of Lemma 3.1, we see $\hat{Q}(\lambda) = (a_1)$, where $a_1 \in \{0, 1\}$. If $\hat{Q}(\lambda) = 1$, then $\lambda(N) = (1)$ for all $N \in GL_n(\mathbf{F}_q)$ and λ is trivial. Since λ is nontrivial, $\hat{Q}(\lambda) = (0)$ and hence $\mathrm{Tr}(\hat{Q}(\lambda)\,\hat{Q}(\lambda)^*) = 0$. □

Combining Lemmas 1.2, 3.1, and 3.2, we conclude

LEMMA 3.3.

$$\| R - U \|^2 \leq \frac{1}{4} \sum^{**} d_\lambda^2 \left| \frac{\chi_\mu^\lambda}{d_\lambda} \right|^{2m},$$

where $\sum^{**}$ means sum over all irreducible representations λ except for λ such that $\lambda(\varphi_a) = (1^n)$ for some $\varphi_a \in \Phi$ with $d(\varphi_a) = 1$.

4. Preliminary bounds on $\chi_\mu^\lambda / d_\lambda$

The exact expressions for χ_μ^λ and d_λ are too cumbersome to substitute directly in the right side of the inequality in Lemma 3.3 and get useful general results. We wish to find bounds useful for asymptotics as $n \to \infty$.

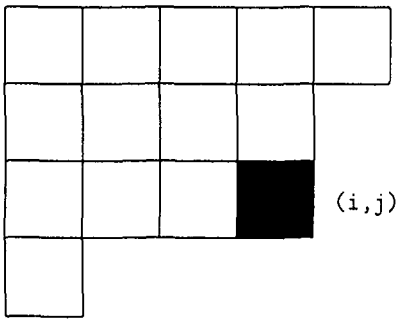

Fig. 2. Tableau used in the proof of Lemma 4.1.

In the expression in Theorem 2.1 for $\chi^\lambda_\mu/d_\lambda$, the first term is positive and the second term simplifies to $-1/(q^{n-1}-1)$. Thus

$$\frac{-1}{q^{n-1}-1} \le \frac{\chi^\lambda_\mu}{d_\lambda}.$$

The following upper bounds will be useful:

LEMMA 4.1. *If* $\lambda(\varphi_a) \neq (1^n)$ *for all* $\varphi_a \in \Phi$ *with* $d(\varphi_a) = 1$, *then*

$$\frac{\chi^\lambda_\mu}{d_\lambda} < \frac{q(1-q^{-1})}{q^n(1-q^{-(n-1)})(1-q^{-n})} \sum_{\varphi_a : d(\varphi_a)=1} \sum_{(i',j)\in C(\lambda,\varphi_a)} \frac{(q^j-1)(q^{i'}-1)}{q^{j-1}(q-1)^2},$$

where $C(\lambda, \varphi_a) = \{(i, j) : (i, j) \in \lambda(\varphi_a), (i+1, j) \notin \lambda(\varphi_a), (i, j+1) \notin \lambda(\varphi_a)\}$.

COROLLARY. *For the same* λ,

$$\frac{\chi^\lambda_\mu}{d_\lambda} < \frac{1}{(q^{n-1}-1)(1-q^{-n})} \sum_{\varphi_a : d(\varphi_a)=1} \sum_{(i',j)\in C(\lambda,\varphi_a)} \frac{(q^{i'}-1)}{(q-1)}.$$

Proof of Lemma 4.1. By Theorem 2.1,

$$\frac{\chi^\lambda_\mu}{d_\lambda} < \frac{q(1-q^{-1})}{q^n(1-q^{-(n-1)})(1-q^{-n})} \sum_{\varphi_a : d(\varphi_a)=1} \sum_{\lambda_1 \in B(\lambda,\varphi_a)} \frac{q^{n(\lambda_1(\varphi_a)')}\tilde{H}_{\lambda_1(\varphi_a)}(q)^{-1}(q-1)^{-1}}{q^{n(\lambda(\varphi_a)')}\tilde{H}_{\lambda(\varphi_a)}(q)^{-1}},$$

where $B(\lambda, \varphi_a) = \{\lambda_1 : |\lambda_1(\varphi_a)| = |\lambda(\varphi_a)| - 1, \lambda_1(\varphi) \subset \lambda(\varphi)$ for all $\varphi \in \Phi$, and $\| \lambda_1 \| = \| \lambda \| - 1\}$.

Suppose one obtains $\lambda_1(\varphi_a)$ from $\lambda(\varphi_a)$ by removing the element at the end of row i (see Figure 2.) Then

$$\frac{q^{n(\lambda_1(\varphi_a)')}}{q^{n(\lambda(\varphi_a)')}} = \frac{1}{q^{j-1}}.$$

If $a_1, \ldots, a_{j-1}$ and $b_1, \ldots, b_{i-1}$ are hook lengths in $\lambda_1(\varphi_a)$'s row i and column j, respectively, then

$$\frac{\tilde{H}_{\lambda(\varphi_a)}(q)}{\tilde{H}_{\lambda_1(\varphi_a)}(q)} = \frac{(q-1)(q^{a_1+1}-1)\cdots(q^{a_{j-1}+1}-1)(q^{b_1+1}-1)\cdots(q^{b_{i-1}+1}-1)}{(q^{a_1}-1)\cdots(q^{a_{j-1}}-1)(q^{b_1}-1)\cdots(q^{b_{i-1}}-1)}.$$

Note that

$$\frac{q^{x+1}-1}{q^x-1} = q + \frac{q-1}{q^x-1}$$

for positive integers x. If $1 \le x \le y$, then $(q-1)/(q^x-1) \ge (q-1)/(q^y-1)$ and hence

$$\frac{q^{y+1}-1}{q^y-1} \le \frac{q^{x+1}-1}{q^x-1}.$$

Thus we may conclude

$$\begin{aligned}\frac{\tilde{H}_{\lambda(\varphi_a)}(q)}{\tilde{H}_{\lambda_1(\varphi_a)}(q)} &\le (q-1)\frac{(q^2-1)}{(q-1)}\cdots\frac{(q^j-1)}{(q^{j-1}-1)}\frac{(q^2-1)}{(q-1)}\cdots\frac{(q^i-1)}{(q^{i-1}-1)} \\ &= \frac{(q^j-1)(q^i-1)}{(q-1)}.\end{aligned}$$

Lemma 4.1 follows by substitution. □

5. Upper bound

In this section we show Theorem 1.1. First we show

LEMMA 5.1.

$$\sum_{\lambda:\chi_\mu^\lambda \le 0} d_\lambda^2 \left|\frac{\chi_\mu^\lambda}{d_\lambda}\right|^{2(n+c)} < A_1 e^{-k_1 c}$$

for some positive constants A_1 and k_1 and for sufficiently large n.

Proof. We know

$$\frac{\chi_\mu^\lambda}{d_\lambda} \ge \frac{-1}{q^{n-1}-1}.$$

Furthermore,

$$\sum_{\lambda:\chi_\mu^\lambda \neq 0} d_\lambda^2 \leq \sum_{\lambda:\|\lambda\|=n} d_\lambda^2 = |GL_n(\mathbf{F}_q)| < q^{n^2}.$$

Thus

$$\sum_{\lambda:\chi_\mu^\lambda \neq 0} d_\lambda^2 \left| \frac{\chi_\mu^\lambda}{d_\lambda} \right|^{2(n+c)} < q^{n^2}(q^{n-1}-1)^{-2(n+c)},$$

and the result follows. □

Next we show

LEMMA 5.2.

$$\sum_{\lambda \in F} d_\lambda^2 \left| \frac{\chi_\mu^\lambda}{d_\lambda} \right|^{2(n+c)} < A_2 e^{-k_2 c}$$

for some positive constants A_2 and k_2 and sufficiently large n and where $F := \{\lambda : (\lambda(\varphi_a)')_1 \leq n - n^{0.6}$ for all $\varphi_a \in \Phi$ with $d(\varphi_a) = 1\}$.

Proof. By Lemma 5.1, we need consider only λ such that $\chi_\mu^\lambda > 0$. Thus we may apply an even power to both sides of the inequality in the corollary to Lemma 4.1.

Let $x(\varphi_a) = (\lambda(\varphi_a)')_1$, where $\varphi_a \in \Phi$ with $d(\varphi_a) = 1$. Observe

$$\frac{\chi_\mu^\lambda}{d_\lambda} < \sum_{\varphi_a : d(\varphi_a)=1} \frac{nq^{x(\varphi_a)}(1+\epsilon)}{q^{n-1}} < \frac{qnq^{x_{\max}}(1+\epsilon)}{q^{n-1}},$$

where $\epsilon < 1$ for sufficiently large n and $x_{\max} = \max_{\varphi_a : d(\varphi_a)=1} x(\varphi_a)$. The term $nq^{x(\varphi_a)}$ comes from

$$\begin{aligned} \sum_{(i',j) \in C(\lambda, \varphi_a)} (q^{i'} - 1) &< \sum_{(i',j) \in C(\lambda, \varphi_a)} (q^{x(\varphi_a)}) \\ &< nq^{x(\varphi_a)}. \end{aligned}$$

We shall sum over values of $n - x(\varphi_a)$. Consider

$$\sum_{i=n^{0.6}}^{n} \sum_{\lambda:(\lambda(\varphi_a)')_1 = n-i} d_\lambda^2 \left(\frac{qn(1+\epsilon)}{q^{i-1}} \right)^{2(n+c)}$$

To bound this expression, we shall relate d_λ with $d_{\tilde{\lambda}}$, where $\tilde{\lambda}$ is a partition-valued function with $\| \tilde{\lambda} \| = i$. This relation is analogous to a technique used in the problem involving random transpositions on S_n. The following two lemmas do so for the random transvection problem.

LEMMA 5.3. *If* $(\lambda(\varphi_a)')_1 = n - i$, *then*

$$d_\lambda \leq \frac{q^i \prod_{j=i+1}^{n}(q^j - 1)}{\prod_{j=1}^{n-i}(q^j - 1)} d_{\bar{\lambda}},$$

where $\bar{\lambda}$ *is defined by* $\bar{\lambda}(\varphi) = \lambda(\varphi)$ *if* $\varphi \neq \varphi_a$ *and*

$$(\bar{\lambda}(\varphi_a)) := \max(0, (\lambda(\varphi_a))_j - 1).$$

In other words, $\bar{\lambda}$ is almost identical to λ; the only difference is that the first column is removed from $\lambda(\varphi_a)$.

Proof. This inequality comes from (2.1). One term comes from

$$\frac{\psi_n(q)}{\psi_i(q)} = \prod_{j=i+1}^{n}(q^j - 1).$$

Another term results from

$$\begin{aligned} \frac{H_{\lambda(\varphi_a)}(q_{\varphi_a})^{-1}}{H_{\bar{\lambda}_1(\varphi_a)}(q_{\varphi_a})^{-1}} &= \frac{1}{\prod_{x \in \text{FirstCol}(\lambda(\varphi_a))}(q^{h(x)} - 1)} \\ &\leq \frac{1}{\prod_{j=1}^{n-i}(q^j - 1)}, \end{aligned}$$

where FirstCol(λ) are the boxes in the first column of the Ferrers diagram of λ.

Finally,

$$\frac{q^{n(\lambda(\varphi_a)')}}{q^{n(\bar{\lambda}(\varphi_a)')}} = q^{|\bar{\lambda}(\varphi_a)|} \leq q^i.$$

Lemma 5.3 follows. □

LEMMA 5.4. *Given* $\varphi_a \in \Phi$ *with* $d(\varphi_a) = 1$,

$$\sum_{\lambda:(\lambda(\varphi_a)')_1 = n-i} d_\lambda^2 \leq C^2 q^{2in} q^{-i^2} q^{2i}$$

for some constant C.

Proof. Each $\bar{\lambda}$ corresponds to a representation of $SL_i(\mathbf{F}_q)$, and each $\bar{\lambda}$ such that $(\lambda(\varphi_a)')_1 = n - i$ uniquely determines $\bar{\lambda}$. So

$$\sum_{\lambda:(\lambda(\varphi_a)')_1 = n-i} d_\lambda^2 \leq |GL_i(\mathbf{F}_q)| < q^{i^2}.$$

Note that there exists a positive constant C such that for all n and q,

$$\frac{q}{q-1}\frac{q^2}{q^2-1}\cdots\frac{q^n}{q^n-1} < C.$$

Thus

$$\begin{aligned} d_\lambda &\leq Cq^i \frac{\prod_{j=i+1}^{n} q^j}{\prod_{j=1}^{n-i} q^j} d_\chi \\ &\leq Cq^i (q^i)^{n-i} d_\chi. \end{aligned}$$

Thus we conclude

$$\begin{aligned} \sum_{\lambda:(\lambda(\varphi_a)')_1 = n-i} d_\lambda^2 &\leq C^2 q^{2i(n-i+1)} q^{i^2} \\ &= C^2 q^{2in} q^{-i^2} q^{2i}. \end{aligned}$$

□

Resuming the proof of Lemma 5.2, we see

$$\begin{aligned} \sum_{\lambda \in F} d_\lambda^2 \left| \frac{\chi_\mu^\lambda}{d_\lambda} \right|^{2n} &\leq \sum_{\varphi_a : d(\varphi_a)=1} \sum_{i=n^{0.6}}^{n} \sum_{\lambda:(\lambda(\varphi_a)')_1 = n-i} d_\lambda^2 \left(\frac{qn(1+\epsilon)}{q^{i-1}} \right)^{2n} \\ &\leq \sum_{\varphi_a : d(\varphi_a)=1} \sum_{i=n^{0.6}}^{n} \frac{C^2 q^{2in} q^{-i^2} q^{2i} ((1+\epsilon)n)^{2n}}{(q^{i-2})^{2n}} \\ &= (q-1) \sum_{i=n^{0.6}}^{n} \frac{C^2 q^{2in} q^{-i^2} q^{2i} q^{C'(\log n)n}}{q^{2in} q^{-4n}} \\ &\qquad \text{for some constant } C' > 0 \\ &\leq \sum_{i=n^{0.6}}^{n} C^2 q^{-(i-1)^2} q^2 q^{4n + C'n \log n} \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

Furthermore, since for $i > n^{0.6}$, $n(1+\epsilon)/q^{i-1} < 1/2$ for sufficiently large n, and we may conclude Lemma 5.2. □

Now the only characters of concern are the characters λ with $n > (\lambda(\varphi_a)')_1 > n - n^{0.6}$ for some φ_a with $d(\varphi_a) = 1$. If n is sufficiently large, there can be at most one such φ_a for a given character. The following lemma considers such characters.

LEMMA 5.5. *If* $c \geq 2$, *then*

$$\sum_{\varphi_a : d(\varphi_a)=1} \sum_{i=1}^{n^{0.6}} \sum_{\lambda:(\lambda(\varphi_a)')_1 = n-i} d_\lambda^2 \left| \frac{\chi_\mu^\lambda}{d_\lambda} \right|^{2(n+c)} \leq A_3 e^{-k_3 c}$$

for some positive constants A_3 and k_3 and sufficiently large n.

Proof. By Lemma 5.1, we need consider only λ such that $\chi_\mu^\lambda > 0$.

Examine Lemma 4.1, and pay special attention to the case $j = 1$ with $(i', j) \in C(\lambda, \varphi_a)$. This case corresponds to the first column of $\lambda(\varphi_a)$. Thus $i' = n - i$. So this case corresponds to a term

$$\frac{q(1-q^{-1})}{q^n(1-q^{-(n-1)})(1-q^{-n})}\frac{(q-1)(q^{n-i}-1)}{(q-1)^2} = \frac{a_n}{q^i},$$

where $a_n < 1 + q^{-0.5n}$ for n sufficiently large.

Let $D(\lambda) = \{(i', j, \varphi) : (i', j) \in C(\lambda, \varphi)$, where $d(\varphi) = 1$ and $j > 1$ if $\varphi = \varphi_a\}$. Then

$$\sum_{(i',j,\varphi)\in D(\lambda)} \frac{(q^j-1)(q^{i'}-1)}{q^{j-1}(q-1)^2} \le qq^{n^{0.6}},$$

and for sufficiently large n

$$\frac{q(1-q^{-1})}{q^n(1-q^{-(n-1)})(1-q^{-n})} qq^{n^{0.6}} \le \frac{q^{-0.5n}}{q^i}.$$

Thus

$$\frac{\chi_\mu^\lambda}{d_\lambda} \le q^{-i}(1 + 2q^{-0.5n})$$

for sufficiently large n.

By Lemma 5.4, we conclude

$$\begin{aligned}\sum_{\varphi_a:d(\varphi_a)=1}\sum_{i=1}^{n^{0.6}}\sum_{\lambda:(\lambda(\varphi_a)')_1=n-i} d_\lambda^2\left|\frac{\chi_\mu^\lambda}{d_\lambda}\right|^{2(n+c)} &\le \sum_{\varphi_a:d(\varphi_a)=1}\sum_{i=1}^{n^{0.6}} C^2q^{2in}q^{-i^2}q^{2i}\alpha_n^{2(n+c)}\\ &= (q-1)\sum_{i=1}^{n^{0.6}}\frac{\alpha_n^{2(n+c)}C^2q^{-i^2}q^{2i}}{q^{2ic}}\\ &\le (q-1)\sum_{i=1}^{n^{0.6}}\frac{\alpha_n^{2(n+c)}C^2q^2q^{-i}}{q^{2ic}},\end{aligned}$$

where $\alpha_n := (1 + 2q^{-0.5n})$. Note that $\alpha_n^{2n} \to 1$ as $n \to \infty$ by a calculus exercise.

Thus

$$\begin{aligned}\sum_{\varphi_a:d(\varphi_a)=1}\sum_{i=1}^{n^{0.6}}\sum_{\lambda:(\lambda(\varphi_a)')_1=n-i} d_\lambda^2\left|\frac{\chi_\mu^\lambda}{d_\lambda}\right|^{2(n+c)} &\le \bar{C}\sum_{i=1}^{n^{0.6}}\frac{q^3q^{-i}\alpha_n^{2c}}{q^{2ic}}\\ &\le \bar{C}\sum_{i=1}^{n^{0.6}}\frac{q^3q^{-i}}{q^{1.5ic}}\end{aligned}$$

$$\leq \tilde{C}/q^{1.5c-3}, \quad \text{for sufficiently large } n$$

where $\tilde{C}$ is a constant that does not depend on q. Thus Lemma 5.5 holds. □

The only representations not included in Lemmas 5.1, 5.2, and 5.5 are not included in the sum in Lemma 3.3. Note also that the phrase "for sufficiently large n" and the constants can be made independent of q. Thus we may conclude Theorem 1.1 for $c \geq 2$. By increasing A if necessary, we may remove the restriction $c \geq 2$ since $\| P^{*m} - U \| \leq 1$.

6. Lower bound

Theorem 1.2 is an immediate consequence of a result in [8].

Let G stand for a group, such as $GL_n(\mathbf{F}_q)$, which may vary with n. Define $p^G(k)$ to be the proportion of elements $g \in G$ such that $\dim(\ker(g - I)) = k$. Define $p_\infty^G(k) = \lim_{n\to\infty} p^G(k)$. It is shown in [8] that

$$p^{GL_n(\mathbf{F}_q)}(k) = \frac{1}{|GL_k(q)|} \sum_{m=0}^{n-k} (-1)^m \frac{q^{-km}}{\psi_m(q)},$$

where $\mathbf{F}_q$ has q elements, $|GL_m(q)| = q^{\binom{m}{2}}\psi_m(q)$, $\psi_0(q) = 1$, and $\psi_m(q) = (q^m - 1)(q^{m-1} - 1)\cdots(q - 1)$. The product of $n - c$ transvections fixes a space of dimension at least c.

If $c \geq 3$, then for all q,

$$\begin{aligned} \sum_{k=c}^{n} p^{GL_n(\mathbf{F}_q)}(k) &\leq \sum_{k=c}^{n} \frac{1}{q^k} 2 \\ &\leq \sum_{k=c}^{\infty} \frac{2}{q^k} \\ &\leq 4/q^c. \end{aligned}$$

Furthermore, the proportion of elements g of $SL_n(\mathbf{F}_q)$ with $\dim(\ker(g - I)) \geq c$ is less than $4/q^{c-1}$. For a given value of $\epsilon > 0$, $4/q^{c-1}$ can be made less than ϵ for some c and all q. However, for all elements $g \in SL_n(\mathbf{F}_q)$ such that $P^{*(n-c)} > 0$, $\dim(\ker(g - I)) \geq c$. Thus

$$\| P^{*(n-c)} - U \| > 1 - \epsilon,$$

and Theorem 1.2 is proved.

Note that Theorems 1.1 and 1.2 show that this random process has a sharp transition from having the distance from uniform being close to 1 to having the

distance from uniform being close to 0. Cutoff phenomena have been observed in other random processes on finite groups.

7. Comparisons with other processes on $GL_n(\mathbf{F}_q)$

Diaconis and Shahshahani [4] have presented two other methods for generating random matrices on $GL_n(\mathbf{F}_q)$. One method is simply to choose each of the n^2 entries at random from $\mathbf{F}_q$ and use Gaussian elimination to see if the determinant is 0. The check takes about $\frac{2}{3}n^3$ operations (multiplication and addition). However, one may need to check several matrices before getting an invertible one; the number of matrices is not sharp. The other method is to use the subgroup algorithm. There about $\sum_{k=1}^{n} 4k^2 \approx \frac{4}{3}n^3$ operations suffice, and this number is sharp.

To multiply a random transvection by an arbitrary matrix as described in Section 1 takes about $4n^2$ operations; so to get close to uniform takes about $4n^3$ steps. Although both this approach and the subgroup algorithm take $O(n^3)$ steps, the latter seems to be slower by a factor of about 3. Note that the estimate for random transvections does not take into account the fact that we start with the identity matrix; one may be able to speed up a step or two by such considerations.

Another process is random special transvections. A special transvection is a matrix which is 1 along the diagonal and has exactly one nonzero off-diagonal element. To multiply a special transvection by an arbitrary matrix takes no more than $2n$ steps. How many random special transvections it takes to get a probability distribution on $SL_n(\mathbf{F}_q)$ which is close to uniform is still an open question. $O(n^2/\log n)$ steps are necessary by entropy arguments. If the identity is picked with probability $1/n^2$ and otherwise a random special transvection is picked so that each special transvection is equally likely, then an argument involving eigenvalues and random walks on graphs shows that $O(n^6)$ steps suffice; however, this seems not to be the best bound. For a similar process, where we restrict ourselves to lower unitriangular elements of $SL_n(\mathbf{F}_q)$ and special transvections contained in that subgroup, we can show that $O(n^2 \log n)$ random special transvections suffice to get close to uniform on the lower unitriangular elements.

Acknowledgments

The author thanks Persi Diaconis, his Ph.D. thesis advisor, for suggesting the problem, a method of finding a random transvection, and some references which were helpful. The author also thanks the referees for a number of comments on style. This paper is based in part on the author's Ph.D. thesis.

References

1. E. Artin, *Geometric Algebra*, John Wiley, New York, 1957.
2. P. Diaconis, *Group Representations in Probability and Statistics*, Institute of Mathematical Statistics, Hayward, CA, 1988.
3. P. Diaconis and M. Shahshahani, "Generating a random permutation with random transpositions," *Z. Wahrscheinlichkeitstheorie Verw. Gebeite* **57** (1981), 159–179.
4. P. Diaconis and M. Shahshahani, "The subgroup algorithm for generating uniform random variables," *Probab. Informat. Sci.* **1** (1987), 15–32.
5. J.A. Green, "The characters of the finite general linear groups," *Trans. Amer. Math. Soc.* **80** (1955), 402–447.
6. J.G. Kemeny and J.L. Snell, *Finite Markov Chains*, Springer-Verlag, New York, 1976.
7. I.G. Macdonald, *Symmetric Functions and Hall Polynomials*. Clarendon Press, Oxford, 1979.
8. A. Rudvalis, Unpublished manuscript, 1988.
9. J.P. Serre, *Linear Representations of Finite Groups*, Springer-Verlag, New York, 1977.
10. M. Suzuki, *Group Theory I*, Springer-Verlag, New York, 1982.
11. A.V. Zelevinsky, *Representation Theory of Finite Classical Groups: A Hopf Algebra Approach*, Lecture Notes in Mathematics 869, Springer-Verlag, Berlin, 1981.