Received 9 August, 2007; revised 13 September, 2007

U. Meierfrankenfeld, Department of Mathematics, Michigan State University, East Lansing, MI 48824, U.S.A.

G. Stroth, Martin-Luther-Universität Halle-Wittenberg, Institut für Mathematik, 06099 Halle, Germany
E-mail: stroth@mathematik.uni-halle.de

J. Group Theory 11 (2008), 495–510
DOI 10.1515/JGT.2008.029

Journal of Group Theory

The expected order of a random unitary matrix

Eric Schmutz

(Communicated by A. Shalev)

Abstract. Let μ_n be the average of the orders of the elements the unitary group $\mathrm{U}(n,q)$. The following conjecture of Fulman is proved: for any fixed q,

$$\log \mu_n = n \log q - \log n + o_q(\log n) \quad \text{as } n \to \infty.$$

1 Introduction

This paper concerns the finite unitary group $\mathrm{U}(n,q)$, and so we begin by reviewing some basic notation and definitions relating to this group. Let $q = p^t$ for some prime number p and some positive integer t. The involution $c \mapsto c^q$ is an automorphism of the finite field $\mathbb{F}_{q^2}$ that fixes the subfield $\mathbb{F}_q$. If $A = (a_{i,j})$ is an $n \times n$ matrix with entries in $\mathbb{F}_{q^2}$, let A^* be the matrix with (i,j)-entry $a_{j,i}^q$ (for $1 \leqslant i, j \leqslant n$). The unitary group $\mathrm{U}(n,q)$ consists of those $n \times n$ matrices A for which $A^{-1} = A^*$. It is well known (see e.g. [4, p. 109]), that, under matrix multiplication, this set of matrices forms a group of order

$$|\mathrm{U}(n,q)| = q^{n^2} \prod_{j=1}^{n} (1 - (-q)^{-j}). \tag{1}$$

For any prime power q, let $\mathrm{GL}(n,q)$ be the group of invertible $n \times n$ matrices with entries in $\mathbb{F}_q$. It is well known that $\mathrm{GL}(n,q)$ has order

$$|\mathrm{GL}(n,q)| = q^{n^2} \prod_{j=1}^{n} (1 - q^{-j}).$$

Note that $\mathrm{U}(n,q)$ is a subgroup of $\mathrm{GL}(n,q^2)$, not $\mathrm{GL}(n,q)$.

For any finite group G, let

$$\mu(G) = \frac{1}{|G|} \sum_{g \in G} \mathbf{V}(g),$$

where $\mathbf{V}(g)$ is the order of g. Stong [8] proved that, for any prime power r,

$$\log \mu(\mathrm{GL}(n,r)) = n\log r - \log n + o_r(\log n) \quad \text{as } n \to \infty.$$

Fulman proposed the analogous problem of estimating $\mu(\mathrm{U}(n,q))$. He proved that

$$\log \mu(\mathrm{U}(n,q)) \geqslant \tfrac{1}{2} n \log q^2 - \log n + o_q(\log n),$$

and he conjectured that this lower bound is sharp in the sense that '$\geqslant$' can be replaced with '$=$'. The goal of this paper is to prove Fulman's conjecture.

The rest of this section contains additional definitions and symbols that are listed in quasi-alphabetical order, and then used throughout without comment.

$|f|$: the degree of the polynomial f.
$\tilde{f}$: if $f(x) = x^d + \sum_{j=0}^{d-1} a_j x^j$ is a monic polynomial of degree d with non-zero constant term a_0, then $\tilde{f}(x) = x^d + \sum_{j=0}^{d-1} a_0^{-q} a_{d-j}^q x^j$.
$[\![z^n]\!]F(z)$: coefficient of z^n in $F(z)$.
$c_i(\pi)$: number of parts of size i in the partition π.
$C_\infty = \prod_{j=1}^{\infty}(1-2^{-j})$.
E_n: expected value with respect to P_n, i.e.

$$E_n(\mathbf{Y}) = \frac{1}{|\mathrm{U}(n,q)|} \sum_{A \in \mathrm{U}(n,q)} \mathbf{Y}(\text{char.poly.}(A))$$

for any real-valued function $\mathbf{Y}$ defined on characteristic polynomials of matrices in $\mathrm{U}(n,q)$.
$\mathscr{I}_{d,r}$: the set of all monic irreducible polynomials of degree d in $\mathbb{F}_r[x]$ (except for $\phi(x) = x$, which is excluded from $\mathscr{I}_1$).
$\mathscr{I}_d = \mathscr{I}_{d,q^2}$, $\mathscr{I} = \bigcup_{d=1}^{\infty} \mathscr{I}_{d,q^2}$.
$\mathscr{J}_d = \{\phi \in \mathscr{I}_d \mid \phi = \tilde{\phi}\}$, $\mathscr{J} = \bigcup_{d=1}^{\infty} \mathscr{J}_d$, $\mathscr{K}_d = \mathscr{I}_{d,q^2} - \mathscr{J}_d$, $\mathscr{K} = \bigcup_{n=1}^{\infty} \mathscr{K}_d$.
$\mathscr{K}_+$, $\mathscr{K}_-$: disjoint subsets of $\mathscr{K}$ such that $\phi \in \mathscr{K}_+$ if and only if $\tilde{\phi} \in \mathscr{K}_-$.
$m_\phi = m_\phi(f)$ = the multiplicity of ϕ in f: the largest integer m such that ϕ^m divides f.
$m_\phi(A) = m_\phi$ (characteristic polynomial of A).
$\mathbf{M} = \max_{\phi \in \mathscr{J}} m_\phi$.
O_s: set of all partitions of s into parts that are odd and distinct.
Ω_n: set of all characteristic polynomials of matrices in $\mathrm{U}(n,q)$, i.e. of monic, degree n, polynomials $f \in \mathbb{F}_{q^2}[x]$ satisfying $m_\phi(f) = m_{\tilde{\phi}}(f)$ for all $\phi \in \mathscr{I}$.
P_n: the probability measure on Ω_n induced by the uniform distribution on $\mathrm{U}(n,q)$, i.e.

$$P_n(\mathscr{S}) = \frac{|\{A \in \mathrm{U}(n,q) : \text{char.poly}(A) \in \mathscr{S}\}|}{|\mathrm{U}(n,q)|}$$

for all $\mathscr{S} \subseteq \Omega_n$.

Q_b: the set of all partitions of b into distinct parts.
τ_ϕ: multiplicative order of the roots of the irreducible polynomial ϕ (in a splitting field).
$\mathbf{T}(f) = \mathrm{lcm}\{\tau_\phi : \phi \text{ is an irreducible factor of } f\}$.
$\mathbf{X}_1(f) = \mathrm{lcm}(\{q^{|\phi|}+1 : m_\phi(f) > 0, \phi \in \mathscr{J}\})$.
$\mathbf{X}_2(f) = \mathrm{lcm}(\{q^{2|\phi|}-1 : m_\phi(f) > 0, \phi \in \mathscr{K}_+\})$.
$\mathbf{X}_1(\pi) = \mathrm{lcm}(\{q^d+1 : \pi \text{ has a part of size } d\})$.
$\mathbf{X}_2(\lambda) = \mathrm{lcm}(\{q^{2d}-1 : \lambda \text{ has a part of size } d\})$.
$\mathbf{X}(f) = \mathrm{lcm}(\mathbf{X}_1(f), \mathbf{X}_2(f))$.
$\mathbf{X}(A) = \mathbf{X}$ (characteristic polynomial of A).

2 Reduction from V to X

There is a close relationship between the order of a matrix $A \in \mathrm{GL}(n,q^2)$ and the orders of its eigenvalues (as multiplicative units in a splitting field for the characteristic polynomial). Hence we begin this section with a simple lemma about the orders of the roots of irreducible polynomials. We also state, for future reference, Fulman's formula for the number of unitary matrices with a given characteristic polynomial. These facts are used to bound the maximum order, and to prove that most matrices in $\mathrm{U}(n,q)$ do not have eigenvalues of large algebraic multiplicity. This in turn enables us to reduce the problem of estimating $E_n(\mathbf{V})$ to the easier problem of estimating $E_n(\mathbf{X})$.

Recall that, if $\phi(x) = x^d + \sum_{j=0}^{d-1} a_j x^j$ is a monic polynomial of degree d with non-zero constant term a_0, then $\tilde{\phi}(x) = x^d + \sum_{j=0}^{d-1} a_0^{-q} a_{d-j}^q x^j$.

Lemma 1. *Suppose that $\phi \in \mathscr{I}_d$, and that τ_ϕ and $\tau_{\tilde{\phi}}$ are respectively the orders of the roots of ϕ and $\tilde{\phi}$ (as multiplicative units in $\mathbb{F}_{q^{2d}}$). Then*

(a) $\tau_\phi = \tau_{\tilde{\phi}}$.

(b) *If $\phi = \tilde{\phi}$, then τ_ϕ is a divisor of q^d+1.*

Proof. Observe that ρ is a root of ϕ if and only if ρ^{-q} is a root of $\tilde{\phi}$:

$$\tilde{\phi}(\rho^{-q}) = a_0^{-q}\rho^{-dq}\sum_{k=0}^{d} a_k^q \rho^{kq} = a_0^{-q}\rho^{-dq}\left(\sum_{k=0}^{d} a_k \rho^k\right)^q.$$

The multiplicative order of ρ^{-q} equals the order of ρ^q, which in turn equals the order of ρ (since q and $q^{2d}-1$ are coprime). This proves (a).

Now assume that $\phi = \tilde{\phi}$ and let ρ be a root of ϕ. Then ρ^{-q} is a root of ϕ. But the roots of ϕ are $\rho^{q^2}, \rho^{q^4}, \ldots, \rho^{q^{2d-2}}, \rho^{q^{2d}} = \rho$. Hence, for some $j \leqslant d$, we have $\rho^{q^{2j}} = \rho^{-q}$, and consequently $\rho^{q(q^{2j-1}+1)} = 1$. This proves that τ_ϕ divides $q(q^{2j-1}+1)$. But τ_ϕ also divides $q^{2d}-1$, and $\gcd(q, q^{2d}-1) = 1$. Therefore τ_ϕ divides $q^{2j-1}+1$. Let m be the smallest positive integer such that τ_ϕ divides q^m+1. If $\tau_\phi = 2$, then it is clear that τ_ϕ divides q^d+1 since τ_ϕ divides $q^{2d}-1 = (q^d+1)(q^d-1)$ and both

factors are even. We may therefore assume that $\tau_\phi > 2$. Using [10, Proposition 1] (with $s = 2d$), and the fact that $\tau_\phi \mid q^{2d} - 1$, we get $2d = 2lm$ for some positive integer l. We know that d is odd (from Fulman [2, Theorem 9]), and therefore l must also be odd. Again using Yucas and Mullen [10, Proposition 1] (this time with $s = d$), we get $\tau_\phi \mid q^d + 1$. □

Let Ω_n be the set of polynomials that are characteristic polynomials of matrices in $U(n, q)$. A beautiful characterization of these polynomials is known. A monic polynomial f is in Ω_n if and only if $m_\phi(f) = m_{\tilde{\phi}}(f)$ for all $\phi \in \mathscr{F}$; the multiplicity of ϕ is the same as the multiplicity of $\tilde{\phi}$ for all irreducible polynomials ϕ. With the notational convention that $|U(0, r)| = |GL(0, r)| = 1$ for all prime powers r, we have the following theorem of Fulman [2]:

Theorem 2. *If $f \in \Omega_n$, then*

$$P_n(\{f\}) = \prod_{\phi \in \mathscr{F}} \frac{q^{|\phi|(m_\phi^2 - m_\phi)}}{|U(m_\phi, q^{|\phi|})|} \cdot \prod_{\theta \in \mathscr{H}_+} \frac{q^{2|\theta|(m_\theta^2 - m_\theta)}}{|GL(m_\theta, q^{2|\theta|})|}.$$

Theorem 2 was just one application of powerful generating function techniques that Fulman developed for $U(n, q)$ and other finite classical groups. Related work can be found in Kung [5], Stong [9], and recent work of Fulman, Neumann and Praeger, e.g. [3].

If the eigenvalues of a matrix are distinct, then the order of the matrix is just the least common multiple of the orders of the eigenvalues. The general case is more complicated because the Jordan form includes off-diagonal elements. This leads to Theorem 3 below. This convenient inequality is an immediate consequence of the slightly stronger inequality in the introduction of Stong's paper [8]. (See also Lidl and Niederreiter [6, p. 80]):

Theorem 3. *For all $A \in GL(n, q^2)$, we have* $\mathbf{V} \leqslant p\mathbf{MT}$.

An immediate consequence of Theorem 3 is a bound on the maximum order:

Corollary 4. *For all $A \in GL(n, q^2)$, we have* $\mathbf{V} < pnq^{2n}$.

However a stronger inequality holds for $U(n, q)$.

Corollary 5. *For all $A \in U(n, q)$, we have* $\mathbf{V} \leqslant 3p\mathbf{M}q^n$.

Proof. By Theorem 3, it suffices to prove that $\mathbf{T} \leqslant 3q^n$. Suppose that the characteristic polynomial of A is

$$\prod_{i=1}^{r} \phi_i^{m_{\phi_i}} \prod_{j=1}^{s} (\phi_{r+j}\tilde{\phi}_{r+j})^{m_{\phi_{r+j}}}$$

where $\phi_i \in \mathscr{F}$ for $i \leqslant r$ and $\phi_{r+j} \in \mathscr{H}_+$ for $j \leqslant s$. To simplify notation, let $d_i = |\phi_i|$ and $\tau_i = \tau_{\phi_i}$. Then by Lemma 1, τ_i divides $(q^{d_i} + 1)$ for $i \leqslant r$ and τ_i divides $q^{2d_i} - 1$ for $r < i \leqslant r + s$. Hence

$$\begin{aligned}\mathbf{T}(A) &= \operatorname{lcm}(\tau_1, \tau_2, \ldots, \tau_{r+s}) \\ &\leqslant \operatorname{lcm}(q^{d_1} + 1, q^{d_2} + 1, \ldots, q^{d_r} + 1) \cdot \operatorname{lcm}(q^{2d_{r+1}} - 1, \ldots, q^{2d_{r+s}} - 1).\end{aligned}$$

Without loss of generality, we may assume that $d_i \neq d_j$ for $1 \leqslant i < j \leqslant r$. (If two degrees are equal, then we can remove one of the arguments to the least common multiple function without changing its value.) Then

$$\begin{aligned}\mathbf{T}(A) &\leqslant \prod_{i=1}^{r}(q^{d_i} + 1) \cdot \prod_{j=1}^{s} q^{2d_{r+j}} = q^n \prod_{i=1}^{r}(1 + q^{-d_i}) \\ &\leqslant q^n \prod_{i=1}^{r}(1 + q^{-i}) = q^n \frac{\prod_{i=1}^{r}(1 - q^{-2i})}{\prod_{i=1}^{r}(1 - q^{-i})}.\end{aligned}$$

Neumann and Praeger [7] used a classical identity of Euler to prove that, for any $\alpha \geqslant 2$ and any $r \geqslant 2$,

$$(1 - \alpha^{-1})^2 < \prod_{i=1}^{r}(1 - \alpha^{-i}) < 1 - \alpha^{-1}. \tag{2}$$

Using the upper bound of (2) with $\alpha = q^2$, and the lower bound with $\alpha = q$, we get

$$\mathbf{T}(A) \leqslant q^n \frac{1 - q^{-2}}{(1 - q^{-1})^2} = q^n\left(1 + \frac{2}{q - 1}\right) \leqslant 3q^n. \quad \square$$

We have a bound on the maximum order, but we still need to prove that the maximum multiplicity $\mathbf{M}$ is usually small. If $\xi = \xi(n) \to \infty$, then with high probability, no irreducible factor has multiplicity larger than ξ.

Lemma 6. *For all positive integers n and all $\xi > 2$, we have* $P_n(\mathbf{M} > \xi) \leqslant 40q^{1-\xi}$.

Proof. Suppose that d is a positive integer with $d \leqslant n$ and let $\psi = \tilde{\psi} \in J_d$. Note that, for $f \in \Omega_n$, we have $m_\psi(f) = l$ if and only if $f = \psi^l g$ for some $g \in \Omega_{n-dl}$ such that $m_\psi(g) = 0$. Hence, by Theorem 2,

$$P_n(m_\psi = l) = \frac{q^{dl^2 - dl}}{|U(l, q^d)|} P_{n-dl}(m_\psi = 0) \leqslant \frac{q^{dl^2 - dl}}{|U(l, q^d)|}. \tag{3}$$

Using (1), we get

$$\frac{q^{dl^2}}{|\mathrm{U}(l,q^d)|} = \frac{1}{\prod_{j=1}^{l}(1-(-1)^j q^{-dj})} < \frac{1}{\prod_{j=1}^{l}(1-q^{-dj})} < \frac{1}{C_\infty}.$$

Using (2) again, we get

$$\frac{1}{C_\infty} \leqslant \frac{1}{(1-q^{-1})^2} = \left(1+\frac{1}{q-1}\right)^2 \leqslant 4.$$

Thus

$$\frac{q^{dl^2}}{|\mathrm{U}(l,q^d)|} \leqslant 4.$$

Putting this back into the right-hand side of (3), and summing over l, we get

$$P_n(m_\psi \geqslant \xi) = \sum_{l \geqslant \xi} P_n(m_\psi = l) \leqslant 4 \sum_{l \geqslant \xi} q^{-ld} \leqslant 8q^{-d\xi}. \tag{4}$$

Similarly, for any $\psi \in \mathscr{K}_d$, we have

$$\begin{aligned} P_n(m_\psi = l) &= \frac{q^{2d(l^2-l)}}{|\mathrm{GL}(l,q^{2d})|} P_{n-2dl}(m_\psi = 0) \\ &\leqslant \frac{q^{2d(l^2-l)}}{|\mathrm{GL}(l,q^{2d})|} = q^{-2dl}\left(\prod_{j=1}^{l}(1-q^{-2dj})\right)^{-1} \\ &< q^{-2dl}\left(\prod_{j=1}^{\infty}(1-2^{-2j})\right)^{-1} < 2q^{-2dl}, \end{aligned}$$

and consequently

$$P_n(m_\psi \geqslant \xi) \leqslant 4q^{-2d\xi}. \tag{5}$$

Now, given a real number $\xi > 2$, let $\mathbf{N}_\xi$ be the number of irreducible factors having multiplicity greater than ξ. Then $\mathbf{M} > \xi$ if and only if $\mathbf{N}_\xi > 0$, and it suffices to show that $P_n(\mathbf{N}_\xi > 0) \leqslant 40q^{1-\xi}$.

Combining (4) and (5), we get

$$\begin{aligned} P_n(\mathbf{N}_\xi > 0) \leqslant E(\mathbf{N}_\xi) &= \sum_{d=1}^{\lfloor n/d\xi \rfloor} \sum_{\phi \in \mathscr{J}_{d,q^2}} P_n(m_\phi > \xi) \\ &= \sum_{d=1}^{\lfloor n/d\xi \rfloor} \left(\sum_{\phi \in \mathscr{J}_d} P_n(m_\phi > \xi) + \sum_{\phi \in \mathscr{K}_d} P_n(m_\phi > \xi) \right) \\ &\leqslant \sum_{d=1}^{\infty} (|\mathscr{J}_d| 8q^{-d\xi} + |\mathscr{K}_d| 4q^{-2d\xi}). \end{aligned} \tag{6}$$

It is well known (see e.g. [1, p. 80]) that, for any prime power r,

$$|\mathscr{J}_{d,r}| = \frac{1}{d}\sum_{k|d} \mu(k) r^{d/k} \leqslant \frac{r^d}{d}.$$

Since $\mathscr{K}_d \subseteq \mathscr{J}_{d,q^2}$, it follows that

$$|\mathscr{K}_d| \leqslant \frac{q^{2d}}{d}. \tag{7}$$

We need a similar estimate for $|\mathscr{J}_d|$. Fulman proved that

$$|\mathscr{J}_d| = \begin{cases} 0 & \text{if } d \text{ is even,} \\ \frac{1}{d}\sum_{k|d} \mu(k)(q^{d/k}+1) & \text{otherwise.} \end{cases} \tag{8}$$

For all $d > 1$ we have $\sum_{k|d} \mu(k) = 0$. Therefore, for all odd $d > 1$,

$$|\mathscr{J}_d| = \frac{1}{d}\sum_{k|d} \mu(k) q^{d/k} = |\mathscr{J}_{d,q}| \leqslant \frac{q^d}{d}. \tag{9}$$

(We note that $|\mathscr{J}_d|$ is equal to $|\mathscr{J}_{d,q}|$, even though the two sets are not equal.) For $d = 1$ we have $|\mathscr{J}_d| = q + 1 \leqslant 2q$, and so for all $d \geqslant 1$ we certainly have

$$|\mathscr{J}_d| \leqslant \frac{2q^d}{d}. \tag{10}$$

For $0 < x < \frac{1}{2}$, we have $-\log(1-x) < 2x$, and for $\xi > 2$, we have $q^{1-\xi} < \frac{1}{2}$. Therefore, by putting (10) and (7) into (6), we get

$$P_n(\mathbf{N}_\xi > 0) \leqslant \sum_{d=1}^{\infty}\left(\frac{16q^{d-d\xi}}{d} + \frac{4q^{2d-2d\xi}}{d}\right) \leqslant -20\log(1-q^{1-\xi}) \leqslant 40q^{1-\xi}. \quad \square$$

Now that Lemma 6 is available, we can reduce the problem of estimating $E_n(\mathbf{V})$ to the slightly easier task of estimating $E_n(\mathbf{X})$.

Lemma 7. $\log E_n(\mathbf{V}) \leqslant \log E_n(\mathbf{X}) + O_q(\log\log n)$.

Proof. By Lemma 1, $\mathbf{T}(A)$ divides $\mathbf{X}(A)$ for all A. It therefore suffices to prove that

$$\log E_n(\mathbf{V}) \leqslant \log E_n(\mathbf{T}) + O_q(\log\log n).$$

For any ξ, we have

$$E_n(\mathbf{V}) = P_n(\mathbf{M} \leqslant \xi)E_n(\mathbf{V}|\mathbf{M} \leqslant \xi) + P_n(\mathbf{M} > \xi)E_n(\mathbf{V}|\mathbf{M} > \xi). \tag{11}$$

To estimate the second of the two terms on the right-hand side in (11), we use Corollary 5 and Lemma 6 with $\xi = \log^2 n$:

$$P_n(\mathbf{M} > \xi)E_n(\mathbf{V}|\mathbf{M} > \xi) \leqslant (40q^{1-\log^2 n})(3pnq^n) = q^{n-\log^2 n(1+o(1))}. \tag{12}$$

For the first term on the right-hand side of (11), we are conditioning on $\mathbf{M} \leqslant \xi$ and so we can use the inequality $\mathbf{M} \leqslant \xi$ together with the inequality $\mathbf{V} \leqslant p\mathbf{MT}$ from Theorem 3:

$$\begin{aligned} P_n(\mathbf{M} \leqslant \xi)E_n(\mathbf{V}|\mathbf{M} \leqslant \xi) &\leqslant p\xi P_n(\mathbf{M} \leqslant \xi)E_n(\mathbf{T}|\mathbf{M} \leqslant \xi) \\ &\leqslant p\xi(P_n(\mathbf{M} \leqslant \xi)E_n(\mathbf{T}|\mathbf{M} \leqslant \xi) + P_n(\mathbf{M} > \xi)E_n(\mathbf{T}|\mathbf{M} > \xi)) \\ &= p\xi E_n(\mathbf{T}). \end{aligned} \tag{13}$$

Finally, putting (13) and (12) back into (11), we get

$$E_n(\mathbf{V}) \leqslant (p\log^2 n)E_n(\mathbf{T})\left(1 + \frac{q^{n-\log^2 n(1+o(1))}}{E_n(\mathbf{T})}\right). \tag{14}$$

Since $E_n(\mathbf{T}) \geqslant q^{n-\log n+o(\log n)}$ for all sufficiently large n by Fulman [2, Theorem 26], the lemma follows from (14) by taking logarithms. □

3 Key factorization

There is a second factorization of characteristic polynomials that is crucial for this paper. The idea is to factorize the characteristic polynomial f as $f = gh$ where $\mathbf{X}(f) = \mathbf{X}(g)$ and g is easier to work with than f, and with g and h themselves characteristic polynomials of unitary matrices.

To this end, define $\mathscr{D}(f)$ to be the set of polynomials g that satisfy the following three conditions:

$$g \in \Omega, \quad g \text{ divides } f, \quad \mathbf{X}(g) = \mathbf{X}(f).$$

The set $\mathscr{D}(f)$ is non-empty since $f \in \mathscr{D}(f)$. Because $\mathscr{D}(f)$ is a non-empty finite set that is partially ordered by divisibility, we can choose a minimal element $\pi(f)$.

Suppose that we have chosen, for each $f \in \Omega_n$, a factor $g = \pi(f)$ that is minimal in $\mathscr{D}(f)$. It is clear that, no matter how the minimal element is chosen, it will have the following useful properties:

(1) For all ϕ in $\mathscr{I}$, $m_\phi(g) = 0$ or 1.

(2) For all positive integers d, $\pi(f)$ has zero, one or two irreducible factors of degree d. If there is one such irreducible factor ϕ, then $\phi \in \mathscr{J}_d$. If there are two, and ϕ is one of them, then $\bar{\phi}$ is the other and both are in $\mathscr{K}_d$.

The third property we need is less obvious; it is proved in the following lemma.

Lemma 8. *If $f \in \Omega_n$ and $g = \pi(f)$ has degree $|g| < n$, and if $h = f/\pi(f)$, then*

$$P_n(\{f\}) \leqslant P_{|g|}(\{g\})P_{n-|g|}(\{h\}).$$

Proof. We consider each factor of $P_n(\{f\})$ in the factorization of Theorem 2 and show that it is bounded above by the corresponding factors in the product $P_{|g|}(\{g\})P_{|h|}(\{h\})$.

Suppose first that $\phi \in \mathscr{J}_d$ for some d and that ϕ divides g. To simplify notation, let $m = m_\phi(f)$. In Theorem 2, the factor of $P_n(\{f\})$ corresponding to ϕ is

$$\begin{aligned} \frac{q^{d(m^2-m)}}{|\mathrm{U}(m, q^d)|} &= \frac{q^{-dm}}{\prod_{j=1}^m (1-(-1)^j q^{-dj})} \\ &= \frac{q^{-d}}{(1-(-1)^m q^{-dm})} \frac{q^{-d(m-1)}}{\prod_{j=1}^{m-1}(1-(-1)^j q^{-dj})} \\ &\leqslant \frac{q^{-d}}{(1-q^{-d})} \frac{q^{-d(m-1)}}{\prod_{j=1}^{m-1}(1-(-1)^j q^{-dj})}. \end{aligned} \tag{15}$$

Since ϕ divides g, we have $m_\phi(g) = 1$ and $m_\phi(h) = m - 1$. Therefore the factors of $P_{|g|}(g)$ and $P_{|h|}(h)$ that correspond to ϕ are respectively

$$\frac{q^{-d}}{(1-q^{-d})} \quad \text{and} \quad \frac{q^{-d(m-1)}}{\prod_{j=1}^{m-1}(1-(-1)^j q^{-dj})}.$$

These are precisely the two factors on the right-hand side of (15).

Similarly, if $\phi \in \mathscr{K}^+$ has degree d and ϕ divides g, then the factor of $P_n(\{f\})$ that corresponds to ϕ is

$$\frac{q^{2d(m^2-m)}}{|\mathrm{GL}(m, q^{2d})|} = \frac{q^{-2dm}}{\prod_{j=1}^m (1-q^{-2dj})} \leqslant \frac{q^{-2d}}{(1-q^{-2d})} \frac{q^{-2d(m-1)}}{\prod_{j=1}^{m-1}(1-q^{-2dj})}. \tag{16}$$

Again $m_\phi(g) = 1$ and the factor of $P_{|g|}(g)$ that corresponds to ϕ is

$$\frac{q^{-2d}}{(1-q^{-2d})}.$$

Likewise $m_\phi(h) = m - 1$, and the factor of $P_{|h|}(\{h\})$ that corresponds to ϕ is

$$\frac{q^{-2d(m-1)}}{\prod_{j=1}^{m-1}(1-q^{-2dj})}.$$

Again these two expressions are precisely the factors on the right-hand side of (16).

Finally, if ϕ does not divide g, then $m_\phi(g) = 0$ and $m_\phi(f) = m_\phi(h)$. In this case, the factor of $P_{|g|}(g)$ that corresponds to ϕ is 1, and the factor of $P_{|h|}(h)$ that corresponds to ϕ is exactly the same as the factor $P_n(\{f\})$ that corresponds to ϕ. □

4 Estimating $E_n(\mathbf{X})$

We now have all of the tools necessary to prove the main result:

Theorem 9. $\log E_n(\mathbf{V}) = n\log q - \log n + o_q(\log n)$.

Proof. By Lemma 7, it suffices to prove that $\log E_n(\mathbf{X}) = n\log q - \log n + o_q(\log n)$. Recall the factorizations $f = \pi(f)h$, and define

$$\mathscr{G}_n = \{g : g = \pi(f) \text{ for some } f \in \Omega_n\}.$$

Then

$$E_n(\mathbf{X}) = \sum_{f\in\Omega_n} \mathbf{X}(\{f\})P_n(\{f\}) = \sum_{g\in\mathscr{G}_n} \mathbf{X}(\{g\}) \sum_{\{h:\pi(gh)=g\}} P_n(gh).$$

By Lemma 8, this is less than or equal to

$$\sum_{g\in\mathscr{G}_n} \mathbf{X}(\{g\})P_{|g|}(\{g\}) \sum_{\{h:\pi(gh)=g,\,|h|=n-|g|\}} P_{n-|g|}(\{h\}).$$

The inner sum is bounded by 1 since $P_{n-|g|}$ is a probability measure. Hence

$$E_n(\mathbf{X}) \leqslant \sum_{g\in\mathscr{G}_n} \mathbf{X}(\{g\})P_{|g|}(\{g\}). \tag{17}$$

To estimate the sum in (17), we need an upper bound for $P_{|g|}(\{g\})$. Note that $|\mathrm{U}(1,q^d)| = q^d + 1$ and $|\mathrm{GL}(1,q^{2d})| = q^{2d} - 1$ for all d. Recall that, for $g \in \mathscr{G}_n$, we have $m_\phi(g) \leqslant 1$ for all $\phi \in \mathscr{J}$. Therefore, by Theorem 2, we have

$$\begin{aligned} P_{|g|}(\{g\}) &= q^{-|g|} \prod_{\{\phi\in\mathscr{J}:m_\phi(g)=1\}} \frac{1}{1+q^{-|\phi|}} \prod_{\{\theta\in\mathscr{K}_+:m_\theta(g)=1\}} \frac{1}{1-q^{-2|\phi|}} \\ &\leqslant q^{-|g|} \prod_{d=1}^{\infty} \frac{1}{1-q^{-2d}} \leqslant 2q^{-|g|}. \end{aligned}$$

Thus

$$E_n(\mathbf{X}) \leqslant 2\sum_{g\in\mathscr{G}_n} q^{-|g|}\mathbf{X}(\{g\}) = 2\sum_{m=1}^{n} q^{-m} \sum_{\{g\in\mathscr{G}_n:|g|=m\}} \mathbf{X}(g).$$

Factorize each $g \in \mathscr{G}_n$ as $g = g_1g_2$, where g_1 and g_2 respectively are the products of the irreducible factors in $\mathscr{J}$ and $\mathscr{K}$:

$$g_1 = \prod_{\{\phi\in\mathscr{J}:m_\phi(g)=1\}} \phi \quad \text{and} \quad g_2 = \prod_{\{\theta\in\mathscr{K}_+:m_\theta(g)=1\}} \theta\tilde{\theta}.$$

We certainly have

$$\mathbf{X}(g) = \mathrm{lcm}(\mathbf{X}_1(g_1), \mathbf{X}_2(g_2)) \leqslant \mathbf{X}_1(g_1)\mathbf{X}_2(g_2),$$

and so

$$E_n(\mathbf{X}) \leqslant 2\sum_{m=1}^{n} q^{-m} \sum_{\{g\in\mathscr{G}_n:|g|=m\}} \mathbf{X}_1(g_1)\mathbf{X}_2(g_2). \tag{18}$$

The degrees of of the irreducible factors of g_1 form a partition of the integer $|g_1|$ into distinct odd parts. Let $\mathscr{S}_1(\pi)$ be the set of polynomials g_1 with partition π. Similarly, the degrees of the factors of g_2 from $\theta \in \mathscr{K}_+$ form a partition of $s := |g_2|/2$ into distinct parts, and we let $S_2(\lambda)$ be the set of all g_2 with partition λ. Writing Q_s for the set of all partitions of s into distinct parts and O_b for the set of all partitions of b into distinct odd parts, we get

$$\sum_{\{g\in\mathscr{G}_n:|g|=m\}} \mathbf{X}_1(g_1)\mathbf{X}_2(g_2) = \sum_{s=1}^{\lfloor m/2\rfloor} \sum_{\pi\in O_{m-2s}} \sum_{\lambda\in Q_s} |S_1(\pi)|\,|S_2(\lambda)|\mathbf{X}_1(\pi)\mathbf{X}_2(\lambda). \tag{19}$$

Using the inequalities (7) and (9), we get

$$|S_1(\pi)| \leqslant \frac{q^{|\pi|}}{\pi_1\pi_2\ldots} \quad \text{and} \quad |S_2(\lambda)| \leqslant \frac{q^{2|\lambda|}}{\lambda_1\lambda_2\ldots} \tag{20}$$

(where $\pi_1, \pi_2, \ldots$ are the parts of π and similarly for λ). Putting (19) and (20) back into the right-hand side of (18), we get

$$E_n(\mathbf{X}) \leqslant 2\sum_{m=1}^{n} \left(\sum_{s=1}^{\lfloor m/2\rfloor} \sum_{\lambda\in O_{m-2s}} \frac{\mathbf{X}_1(\pi)}{\pi_1\pi_2\ldots} \sum_{\lambda\in Q_s} \frac{\mathbf{X}_2(\lambda)}{\lambda_1\lambda_2\ldots} \right). \tag{21}$$

Let

$$\sigma_2(s) = \sum_{\lambda\in Q_s} \frac{\mathbf{X}_2(\lambda)}{\lambda_1\lambda_2\ldots}$$

be the innermost sum. This sum was estimated by Stong at the end of [8]. The conclusion was that

$$\sigma_2(s) \leqslant \frac{(q^2)^{s+o(\log s)}}{s} \quad \text{as } s \to \infty.$$

For each positive integer b define

$$\sigma_1(b) = \sum_{\pi \in O_b} \frac{\mathbf{X}_1(\pi)}{\pi_1 \pi_2 \dots}.$$

We show next that it is sufficient to prove that

$$\sigma_1(b) \leqslant b^{-1}(q^{b+o(\log b)}). \tag{22}$$

Assume for now that (22) holds. (This will be verified later.) For integers k let $(k)^+ = \max(k, 1)$. Using the partial fraction decomposition

$$\frac{1}{s(m-2s)} = \frac{1}{ms} + \frac{2}{m(m-2s)}$$

we have

$$\sum_{s=1}^{\lfloor m/2 \rfloor} \frac{1}{s} \frac{1}{(m-2s)^+} = O\left(\frac{\log m}{m}\right).$$

Then inside the parentheses of (21) we have

$$\begin{aligned}\sum_{s=1}^{\lfloor m/2 \rfloor} \sigma_1(m-2s)\sigma_2(s) &= \sum_{s=1}^{\lfloor m/2 \rfloor} \frac{q^{m-2s+o(\log(m-2s))}}{(m-2s)^+} \frac{q^{2s+o(\log s)}}{s} \\ &= q^{m+o(\log m)} \sum_{s=1}^{\lfloor m/2 \rfloor} \frac{1}{s} \frac{1}{(m-2s)^+} \\ &= \frac{q^{m+o(\log m)}}{m}.\end{aligned}$$

Note that q^m/m is an increasing function of m. So if we let $\omega = \lfloor \log n \rfloor$, then we can easily finish estimating (21):

$$\sum_{m=1}^{n} \frac{q^m}{m} = \sum_{m=1}^{n-\omega} \frac{q^m}{m} + \sum_{m=n-\omega+1}^{n} \frac{q^m}{m} \leqslant (n-\omega)\frac{q^{n-\omega}}{n-\omega} + \omega\frac{q^n}{n} = \frac{q^{n+o(\log n)}}{n}.$$

To complete the proof of Theorem 9, all that remains is to prove that $\sigma_1(b) = b^{-1}(q^{b+o(\log b)})$. The sum σ_1 is somewhat similar to σ_2, and we can estimate it with techniques like those that Stong used in estimating σ_2. The cyclotomic polynomials satisfy a simple identity: if π_i is odd, then

$$q^{\pi_i} + 1 = \frac{q^{2\pi_i} - 1}{q^{\pi_i} - 1} = \prod_{d \mid \pi_i} \Phi_{2d}(q).$$

Define

$\Lambda = \Lambda(\pi) = \{d : \text{for some } i,\ d \text{ divides } \pi_i\}$,
$\nu_d(\pi) = \sum_{k \equiv 0(d)} c_k(\pi) =$ the number of parts that are multiples of d, and
$w_d(\pi) = \max(0, \nu_d - 1)$.

Then

$$\operatorname{lcm}(q^{\pi_1} + 1, q^{\pi_2} + 1, \dots) \leqslant \prod_{d \in \Lambda} \Phi_{2d} \tag{23}$$

$$= \frac{\prod_i (q^{\pi_i} + 1)}{\prod_d \Phi_{2d}^{w_d}}. \tag{24}$$

If π is a partition of b into distinct parts, then for the numerator of (24) we have

$$\prod_i (q^{\pi_i} + 1) = q^b \prod_i (1 + q^{-\pi_i}) < q^b \prod_{i=1}^{\infty} (1 + q^{-i}) < 4q^b. \tag{25}$$

An upper bound is obtained if, in the denominator of (24), we restrict d to a finite set of primes. For each i, let p_i denote the ith prime. Given a positive integer ξ, let

$$\mathscr{P} = \mathscr{P}(\xi) = \{p_i : \xi \leqslant i \leqslant e^{\xi}\} = \{p_\xi, p_{\xi+1}, \dots, p_{\lfloor e^\xi \rfloor}\}.$$

Let $\kappa_\xi = \prod_{p \in \mathscr{P}} \Phi_{2p}(q)$. Then, for any $\pi \in O_b$,

$$\operatorname{lcm}(q^{\pi_1} + 1, q^{\pi_2} + 1, \dots) \leqslant \frac{4\kappa_\xi q^b}{\prod_{p \in \mathscr{P}} \Phi_{2p}^{\nu_p}}.$$

Define

$$G(k) = \begin{cases} \displaystyle\prod_{\{p : p \in \mathscr{P} \text{ and } p \mid k\}} \frac{1}{\Phi_{2p}(q)} & \text{if } k \text{ is divisible by at least one prime in } \mathscr{P}, \\ 1 & \text{otherwise.} \end{cases}$$

For any partition π, let

$$z_\pi = \left(\prod_{i=1}^{\infty} c_i! i^{c_i}\right)^{-1},$$

where $c_i = c_i(\pi)$ is the number of parts of size i in π. Thus $z_\pi = (\pi_1 \pi_2 \ldots)^{-1}$ for $\pi \in O_b$. We get an upper bound for $\sigma_1(b)$ if we sum over all partitions of b (not just those in O_b). Hence and from (23) we have

$$\sigma_1(b) \leqslant 4\kappa_\xi q^b \sum_{\pi \vdash b} \frac{z_\pi}{\prod_{p \in \mathscr{P}} \Phi_{2p}^{\nu_p}} = 4\kappa_\xi q^b \sum_{\pi \vdash b} z_\pi \prod_{k=1}^{\infty} G(k)^{c_k}.$$

In the well-known cycle index identity

$$1 + \sum_{b=1}^{\infty} \sum_{\pi \vdash b} z_\pi \prod_k x_k^{c_k} z^b = \exp\left(\sum_{k=1}^{\infty} \frac{x_k z^k}{k} \right),$$

we can make the substitution $x_k = G(k)$ for $k = 1, 2, \ldots$ to get

$$\sigma_1(b) \leqslant 4\kappa_\xi q^b [\![z^b]\!] \exp\left(\sum_{k=1}^{\infty} \frac{G(k)}{k} z^k \right).$$

Following Stong, we note that $G(k)$ is a periodic function of k with period $N = \prod_{p \in \mathscr{P}} p$. Hence we have the Fourier expansion

$$G(k) = a_0 + \sum_{l=1}^{N-1} a_l \omega^{lk},$$

where $\omega = e^{2\pi i/N}$ and the constants a_l are the Fourier coefficients:

$$a_l = \frac{1}{N} \sum_{v=0}^{N-1} G(v) \omega^{-lv}. \tag{26}$$

Thus

$$\exp\left(\sum_{k=1}^{\infty} \frac{G(k)}{k} z^k \right) = \exp\left(\sum_{l=0}^{N-1} a_l \sum_{k=1}^{\infty} \frac{(\omega^l z)^k}{k} \right) = (1-z)^{-a_0} \prod_{l=1}^{N-1} (1 - \omega^l z)^{-a_l}. \tag{27}$$

Let $\alpha = \prod_{l=1}^{N-1} (1 - \omega^l)^{-a_l}$. Because $G(k) > 0$ for all k, it is clear from (26) that $|a_0| > |a_j|$ for all $j > 0$. Hence the coefficient of z^n in (27) is asymptotic to

$$\alpha [\![z^n]\!] (1-z)^{-a_0} = O\left(\frac{1}{n^{1-a_0}} \right).$$

It therefore suffices to verify that a_0 can be made arbitrarily small by choosing ξ sufficiently large.

Note that $\Phi_{2p} = (q^p + 1)/(q+1)$ for odd primes p. Hence

$$G(k) \leqslant \begin{cases} 1 & \text{if } \gcd(k, N) = 1, \\ \frac{q+1}{q^{p_\xi}+1} & \text{otherwise.} \end{cases} \tag{28}$$

Given $\varepsilon > 0$, choose ξ large enough so that we also have

$$\frac{q+1}{q^{p_\xi}+1} < \varepsilon/2.$$

Let $R_\xi = \{k : \gcd(k, N) = 1 \text{ and } k \leqslant N\}$. By inclusion–exclusion,

$$|R_\xi| = \prod_{i=\xi}^{\lfloor e^\xi \rfloor} (1 - p_i^{-1}) N.$$

By the prime number theorem $p_i \sim i \log i$, and consequently

$$\prod_{i=\xi}^{\lfloor e^\xi \rfloor} (1 - p_i^{-1}) = o(1) \quad \text{as } \xi \to \infty.$$

We can therefore also choose ξ large enough so that $|R_\xi| \leqslant \varepsilon N/2$. But then, by (28), we have

$$a_0 = \frac{1}{N} \sum_{k=1}^{N-1} G(k) \leqslant \frac{|R_\xi|}{N} + \frac{p_\xi + 1}{q^{p_\xi}+1} < \varepsilon. \quad \square$$

Acknowledgement. An anonymous referee suggested using the ideas in [7] for the last estimate in the proof of Corollary 5.

References

[1] E. R. Berlekamp. *Algebraic coding theory*, 2nd edn. (Aegean Park Press, 1984).
[2] J. Fulman. Cycle indices for the finite classical groups. *J. Group Theory* **2** (1999), 251–289.
[3] J. Fulman, P. M. Neumann and C. E. Praeger. A generating function approach to the enumeration of matrices in classical groups over finite fields. *Mem. Amer. Math. Soc.* **176** (2005), no. 830.
[4] L. C. Grove. *Classical groups and geometric algebra*. Graduate Studies in Math. 39 (American Mathematical Society, 2002).
[5] J. P. S. Kung. The cycle structure of a linear transformation over a finite field. *Linear Algebra Appl.* **36** (1981), 141–155.

[6] R. Lidl and H. Niederreiter. *Introduction to finite fields and their applications* (Cambridge University Press, 1994).
[7] P. M. Neumann and C. E. Praeger. Cyclic matrices over finite fields. *J. London Math. Soc.* (2) **52** (1995), 263–284.
[8] R. Stong. The average order of a matrix. *J. Combin. Theory Ser. A* **64** (1993), 337–343.
[9] R. Stong. Some asymptotic results on finite vector spaces. *Adv. in Appl. Math.* **9** (1988), 167–199.
[10] J. L. Yucas and G. L. Mullen. Self-reciprocal irreducible polynomials over finite fields. *Des. Codes Cryptog.* **33** (2004), 275–281.

Received 12 September, 2006; revised 20 January, 2007

Eric Schmutz, Department of Mathematics, Drexel University, Philadelphia, PA 19104, U.S.A.
E-mail: Eric.Jonathan.Schmutz@drexel.edu

J. Group Theory **11** (2008), 511–524
DOI 10.1515/JGT.2008.030

Journal of Group Theory

On the exponent semigroups of finite p-groups

Primož Moravec

(Communicated by F. de Giovanni)

Abstract. In this note we describe the exponent semigroups of finite p-groups of maximal class and finite p-groups of class at most 5. Consequently, sharp bounds for the exponent of the Schur multiplier of a finite p-group of class at most 4 are obtained. Our results extend some well-known results of Jones (1974).

1 Introduction

A group G is said to be *n-abelian* if the map $x \mapsto x^n$ is an endomorphism of G. The study of n-abelian groups was initiated by Levi in [14], and has been a topic of several other papers; see, e.g., [1], [4], [8], [12], [15]. Given a group G, define

$$\mathscr{E}(G) = \{n \in \mathbb{Z} : (xy)^n = x^n y^n \text{ for all } x, y \in G\}.$$

It is clear that $\mathscr{E}(G)$ is a multiplicative subsemigroup of $\mathbb{Z}$ containing 0 and 1. Following Kappe [12], we say that $\mathscr{E}(G)$ is the *exponent semigroup* of G. One of the main results of [12] is a number-theoretic characterization of $\mathscr{E}(G)$ for an arbitrary group G. We have that $\mathscr{E}(G)$ is either $\{0, 1\}$, $\mathbb{Z}$ or a so-called Levi system [12]. When G is a finite p-group, a more refined description of $\mathscr{E}(G)$ can be obtained. It is proved in [16] that for every finite p-group G there exists a non-negative integer r such that $\mathscr{E}(G) = p^{e+r}\mathbb{Z} \cup (p^{e+r}\mathbb{Z} + 1)$, where $\exp G/Z(G) = p^e$. Following [16], we say that r is the *exponential rank* of G, and denote it by $\operatorname{exprank}(G)$. The exponential rank of a finite p-group G, together with $\exp G/Z(G)$, completely determines the endomorphisms of G of the form $x \mapsto x^n$. Moreover, it has been shown in [16] that $\operatorname{exprank}(G)$ can be bounded in terms of $\exp G/Z(G)$. In [16], some classes of finite p-groups having small exponential rank have been exhibited. For instance, every finite abelian p-group clearly has exponential rank zero, and the same is true for regular p-groups [16]. It has also been proved in [16] that powerful p-groups have exponential rank at most 1; when p is odd, then the exponential rank is always zero, and when $p = 2$, the exponential rank is 1, if the group in question is non-abelian.

Roughly speaking, the exponential rank of a finite p-group G can be calculated once the power-commutator structure of G has been determined. A prominent class