BULL. AUSTRAL. MATH. SOC. 20G15
VOL. 22 (1980), 339-364.

CONJUGACY CLASSES IN PROJECTIVE AND SPECIAL LINEAR GROUPS

G.E. WALL

The conjugacy classes in the finite-dimensional projective full linear, special linear and projective special linear groups over an arbitrary commutative field are determined. The results over a finite field are applied to certain enumerative problems.

1. Introduction

One of the first things to establish about a given group is the distribution of its elements into conjugacy classes. In the case of the full linear group $GL_n(F)$, where F is a (commutative) field, this information is supplied by the classical theory of the similarity of matrices. The object of the present paper is to develop the corresponding theory for the groups $PGL_n(F)$, $SL_n(F)$ and $PSL_n(F)$. The methods are direct and elementary, keeping within the usual framework of similarity theory. Special attention is paid to the case of a finite coefficient field, where the results take a particularly simple and transparent form. The fact that the *special* and *projective indices* (defined in (3.17) and (3.18)) enter the relevant formulae in a symmetrical way is the source of the dualities observed by Lehrer ([4], Theorem B) and Macdonald ([5], Remark after (4.6)).

Macdonald ([5]) also develops the conjugacy theory over finite fields, although by somewhat different methods (for example, greater emphasis is placed on a certain partition of n called the *type*). Reading his paper

Received 29 April 1980.

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

stimulated me to work out several further results, which appear here as (4.17)-(4.21).

2. General principles

Let F be a (commutative) field. Denote by F^* the multiplicative group of its non-zero elements. We consider the full linear group $GL_n(F)$. The homomorphism

$$\det : GL_n(F) \to F^*$$

maps $GL_n(F)$ onto F^* and has kernel $SL_n(F)$, so that $GL/SL \cong F^*$. Hence, if G is a subgroup of GL,

$$G/SG \cong \det G ,$$

where

$$SG = G \cap SL .$$

The non-zero scalar matrices λI form a central subgroup $Z = Z(n, F)$ of GL isomorphic to F^*. The canonical homomorphism

$$P : GL \to PGL = GL/Z$$

carries each subgroup G of GL onto its projective counterpart $PG = GZ/Z$.

The subgroup G acts by conjugation on GL. The *G-class* of a nonsingular matrix A is defined as its orbit under this action, namely,

(2.1) $$(A)_G = \{TAT^{-1} : T \in G\} .$$

The G-classes of elements of G are just the conjugacy classes of G.

Similarly, PG acts by conjugation on PGL. The PG-class of a non-singular matrix A is defined to be

(2.2) $$(A)_{PG} = \{\lambda(TAT^{-1}) : \lambda \in F^*, T \in G\} .$$

In other words,

$$(A)_{PG} = P^{-1} \text{ (orbit of } PA \text{ under } PG) ,$$

so that there is a canonical one-one correspondence between PG-classes and

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

orbits under the action of PG on PGL .

Let us now compare $(A)_G$, $(A)_{SG}$, $(A)_{PG}$ and $(A)_{PSG}$. For this purpose we introduce the following groups:

(2.3) $C = C_G(A) = \{T \in G : TAT^{-1} = A\}$,

(2.4) $\Gamma = \Gamma_G(A) = \{T \in G : TAT^{-1} = \text{scalar multiple of } A\}$,

(2.5) $L = L_G(A) = \{\lambda \in F^* : (\lambda A)_G = (A)_G\}$,

(2.6) $\Lambda = \Lambda_G(A) = L_{SG}(A)$.

We shall apply again and again the simple principle that the elements in the orbit of a given point correspond one-one to the left cosets of the stabilizer of that point.

First, G acts by conjugation on the set of all SG-classes and $(A)_G$ is the union of the SG-classes in the orbit of $(A)_{SG}$. The stabilizer of $(A)_{SG}$ is clearly $(SG)C$. In view of the isomorphism $G/(SG)C \cong \det G/\det C$, we have:

(2.7) *the* SG*-classes into which* $(A)_G$ *splits correspond one-one to the elements of* $\det G/\det C$.

A similar argument gives:

(2.8) *the* PSG*-classes into which* $(A)_{PG}$ *splits correspond one-one to the elements of* $\det G/\det \Gamma$.

Next, F^* acts on the set of all G-classes by the rule

(2.9) $$\lambda \circ (A)_G = (\lambda A)_G ,$$

and $(A)_{PG}$ is the union of the G-classes in the orbit of $(A)_G$. Since the stabilizer of $(A)_G$ is L , we have:

(2.10) *the* G*-classes into which* $(A)_{PG}$ *splits correspond one-one to the elements of* F^*/L .

Replacing G by SG in (2.10), we get:

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

(2.11) *the SG-classes into which* $(A)_{PSG}$ *splits correspond one-one to the elements of* F^*/Λ .

Some simple properties of the groups C, Γ, L and Λ corresponding to a given non-singular matrix A may be noted. If $T \in \Gamma$, there exists $\lambda \in F^*$ such that

(2.12) $$TAT^{-1} = \lambda A .$$

Taking determinants, we deduce that

(2.13) *every element of* L *is an* nth *root of unity; thus* L *is a finite cyclic group of order dividing* n .

Again, the mapping

$$\text{mult} : \Gamma \to F^*$$

which assigns to each $T \in \Gamma$ the multiplier λ in (2.12) is a homomorphism with image L and kernel C , so that

$$\Gamma/C \cong L .$$

This implies that $\Gamma/(S\Gamma)C \cong L/\Lambda$. Since also $\Gamma/(S\Gamma)C \cong \det \Gamma/\det C$, we have

(2.14) $$\det \Gamma/\det C \cong L/\Lambda .$$

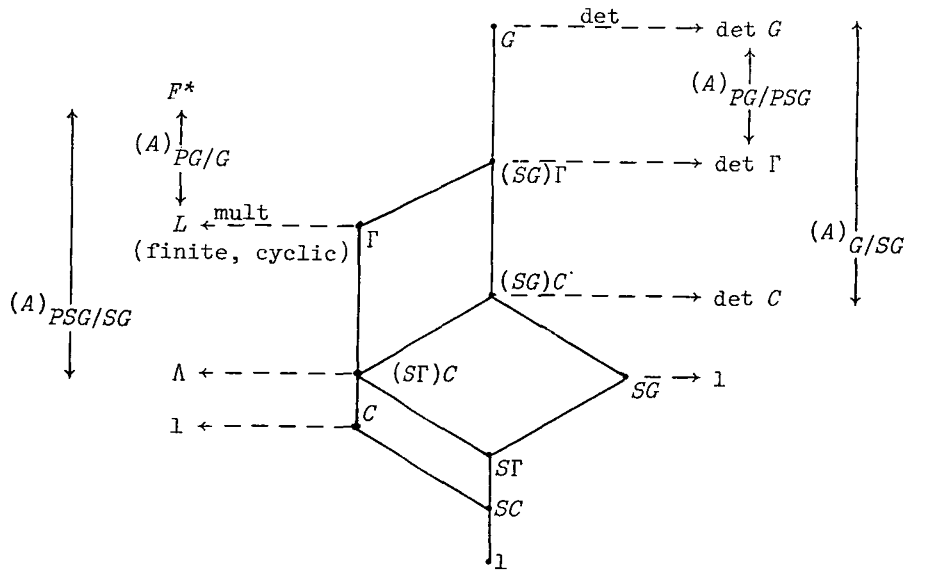

Figure 1

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

Figure 1 is a lattice diagram in which meets and joins are indicated. The broken arrows to left and right show the effect of the homomorphisms mult and det . The label $(A)_{PG/G}$ (for example) indicates that the splitting of $(A)_{PG}$ into G-classes is governed by the quotient group F^*/L in accordance with (2.10).

3. The full linear group

We now specialize to the case $G = GL_n(F)$. We shall determine explicitly the groups $L = L_{GL}(A)$, $\Lambda = \Lambda_{GL}(A)$, $\det C = \det C_{GL}(A)$ and $\det \Gamma = \det \Gamma_{GL}(A)$, where $A \in GL_n(F)$. The term *GL-class* will mean a $GL_n(F)$-class for some (usually unspecified) n . In other words, a GL-class is just a similarity class of non-singular matrices over F . The GL-class of A is denoted by $(A)_{GL}$. We define SL-, PGL- and PSL-classes in the obvious way and use the corresponding notation $(A)_{SL}$, $(A)_{PGL}$, $(A)_{PSL}$.

The similarity class of an $n \times n$ matrix A over F is determined by the elementary divisors of A . However, while this specification is adequate for some of our purposes, the following variant is more convenient for others. Each elementary divisor of A is a power of a monic irreducible polynomial over F . For a given positive integer r , let those elementary divisors of A which are rth powers of monic irreducible polynomials be

$$\pi_1(x)^r, \pi_2(x)^r, \ldots ,$$

each elementary divisor being written down with correct multiplicity. Write

$$f_r(x) = \pi_1(x)\pi_2(x) \ldots .$$

Then the elementary divisors of A , and hence also its similarity class, are uniquely determined by the sequence

$$\sigma(A) = \big(f_1(x), f_2(x), \ldots\big) . \tag{3.1}$$

Macdonald [5] uses essentially the same specification of the similarity

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

classes.

Note that

(3.2) $$n(A) = \deg f_1 + 2 \deg f_2 + \dots ,$$

(3.3) $$\det A = \delta(f_1)\delta(f_2)^2 \dots ,$$

where $n(A) = n$ and $\delta(f)$ denotes the product of the roots of f. In particular, A is non-singular if, and only if, no $f_r(x)$ is divisible by x.

From now on, we assume that A is non-singular. The groups $\det C$, L and $\det \Gamma$ are determined in succession in Theorems 1, 2 and 3. Since L is a finite cyclic group, its subgroup Λ is then determined by the general isomorphism (2.14). Finally, in Theorem 4, these results are further specialized to the case where F is a finite field.

The following result will be required in the proof of Theorem 1.

LEMMA. *Let $X \in GL_n(R)$, where R is a commutative local ring with nilpotent maximal ideal. Then there exist products of $n \times n$ unipotent matrices P, Q over R such that PXQ is diagonal.*

Proof. Let J be the maximal ideal, and $\bar{R} = R/J$ the residue class field, of R. Let $M_n(R) \to M_n(\bar{R})$, $Y \mapsto \bar{Y}$, denote the homomorphism induced by the canonical homomorphism $R \to \bar{R}$. An $n \times n$ matrix is called *elementary* if all its diagonal elements are 1 and all except one of its off-diagonal elements are 0. Since $\bar{X} \in GL_n(\bar{R})$ and $\bar{R}$ is a field, there exist products of elementary matrices P_1, Q_1 over R such that $\bar{P}_1\bar{X}\bar{Q}_1 = \bar{D}$, where D is diagonal. Thus, $P_1XQ_1 = D + Y$, where $Y \in M_n(J)$. Since $\bar{D}$ is non-singular, every diagonal element of D lies in $R - J$ and so is a unit; hence $D \in GL_n(R)$ and $P_1XQ_1 = D(I+Z)$, where $Z = D^{-1}Y$. But $Z \in M_n(J)$ and J is nilpotent, so that Z is nilpotent and thus $I + Z$ unipotent. Therefore $PXQ = D$, where $P = P_1$ and $Q = Q_1(I+Z)^{-1}$ are both products of unipotent matrices. This completes the proof.

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

Notation. (a) If H is a subgroup of F^* and d a positive integer, then

$$H^d = \{h^d : h \in H\} , \tag{3.4}$$

$$H_d = \{h : h \in H,\ h^d = 1\} . \tag{3.5}$$

(b) If π is an irreducible polynomial in $F[x]$ and K the field obtained by adjoining a root of π to F, then $\Delta(\pi)$ denotes the image of K^* under the norm homomorphism $N_{K/F} : K^* \to F^*$.

THEOREM 1. *Let* $A \in GL_n(F)$ *, where* F *is a field. Then*

$$\det C_{GL}(A) = \prod \Delta(\pi)^r , \tag{3.6}$$

where the product is taken over the elementary divisors π^r *of* A .

The proof will be carried out in terms of linear transformations rather than matrices. Let T be a non-singular linear transformation on a finite-dimensional vector space V over F. We turn V into an $F[x]$-module in the usual way by defining $f(x)v = f(T)v$. Let E denote the ring of module endomorphisms of V and E^* the group of units of E. Then the assertion of the theorem is that

$$\det E^* = \prod \Delta(\pi)^r , \tag{3.7}$$

where the product is taken over the elementary divisors π^r of T.

We begin with the simplest case of all, where T has a single irreducible elementary divisor π with multiplicity 1. Let K be the field obtained by adjoining a root α of π to F. Then K is a finite-dimensional vector space over F and for each $\beta \in K$ the mapping $\hat{\beta} : K \to K$, $u \mapsto \beta u$, is F-linear. We may take $V = K$ and $T = \hat{\alpha}$, and it is easy to see that E consists of the $\hat{\beta}$. The well-known formula $\det \hat{\beta} = N_{K/F}(\beta)$ now gives $\det E^* = \Delta(\pi)$, as required.

We now proceed to the next simplest case, where T has a single elementary divisor π^r with multiplicity 1. Since V is a cyclic module, the elements of E are the polynomials $f(T)$. Let $\overline{T}$ be the linear transformation induced by T on the quotient module $\overline{V} = V/\pi(x)V$

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

and let $\bar{E}$ be the ring of module endomorphisms of $\bar{V}$. Then $\bar{T}$ has the single elementary divisor π with multiplicity 1, $\bar{E}$ consists of the polynomials $f(\bar{T})$ and, by what we have already proved, $\det \bar{E}^* = \Delta(\pi)$.

We observe now that

$$V \supset \pi(x)V \supset \ldots \supset \pi(x)^r V = \{0\}$$

is a composition series for the module V in which all quotient modules

$$\pi(x)^{i-1}V/\pi(x)^i V \quad (i = 1, \ldots, r)$$

are isomorphic. (Indeed, multiplication by $\pi(x)^{i-1}$ gives an isomorphism of $V/\pi(x)V$ onto $\pi(x)^{i-1}V/\pi(x)^i V$.) It follows that

$$\det f(T) = \left(\det f(\bar{T})\right)^r ,$$

whence

$$\det E^* = (\det \bar{E}^*)^r = \Delta(\pi)^r ,$$

as required.

We turn now to the general case. Write $M\left(\pi^r\right)$ for the indecomposable $F[x]$-module $F[x]/\pi(x)^r F[x]$. Let the elementary divisors of T be $\pi_1^{r_1}, \ldots, \pi_k^{r_k}$ with respective multiplicities $m_1, \ldots, m_k$. Then we may assume that

$$V = V_1 \oplus \ldots \oplus V_m \quad \left(m = \sum m_i\right) , \tag{3.8}$$

where

$$V_1 = \ldots = V_{m_1} = M\left(\pi_1^{r_1}\right) ,$$

$$V_{m_1+1} = \ldots = V_{m_1+m_2} = M\left(\pi_2^{r_2}\right) ,$$

and so on. In view of what we have proved already, (3.7) can be rewritten as

$$\det E^* = \prod_{i=1}^{m} \det E_i^* , \tag{3.9}$$

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

where E_i is the ring of module endomorphisms of V_i .

In the present paragraph we take advantage of the direct decomposition (3.8) to *identify* E with the ring of all $m \times m$ matrices $S = (s_{ij})$, where $s_{ij} \in \mathrm{Hom}_{F[x]}(V_i, V_j)$ for all i, j . Such a matrix S can be written as a $k \times k$ block matrix $(S_{\lambda\mu})$, where $S_{\lambda\mu}$ is an $m_\lambda \times m_\mu$ matrix for all λ, μ . Notice that

$$S_{\lambda\lambda} \in M_{m_\lambda}\left(R\left(\pi_\lambda^{r_\lambda}\right)\right) ,$$

where

$$R\left(\pi_\lambda^{r_\lambda}\right) = \mathrm{End}\ M\left(\pi_\lambda^{r_\lambda}\right)$$

is isomorphic to the (local) ring $F[x]/\pi_\lambda(x)^{r_\lambda}F[x]$. We introduce the block diagonal matrix

$$S' = \mathrm{diag}\left(S_{11}, \ldots, S_{kk}\right) .$$

Then (see Jacobson [2], Chapter 4, Theorem 8)

$$S' \equiv S \pmod{\mathrm{rad}\ E} .$$

Suppose now that S is invertible. Then $S' = S(I+N)$, where $N \in \mathrm{rad}\ E$ and so $I + N$ is unipotent. In particular, S' is invertible and so each of its diagonal blocks $S_{\lambda\lambda}$ is invertible. Applying the lemma to each of these diagonal blocks, we deduce that there exist products of (block diagonal) unipotent matrices P_1, Q_1 such that $P_1S'Q_1 = D$, where D is diagonal. Then

$$PSQ = D , \tag{3.10}$$

where $P = P_1$ and $Q = (I+N)Q_1$ are also products of unipotent matrices.

Let us now regard the elements of E once more as linear transformations on V . Then (3.10) implies that every element of E^* has the same determinant as some element of E^* which maps every V_i onto itself. Since the determinants of the latter elements of E^* obviously form the

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

group $\prod \det E_i^*$, our result (3.9) follows. This completes the proof of the theorem.

Consider the set M of all monic polynomials $f(x) \in F[x]$ which are not divisible by x . Then F^* acts on M by the rule

$$(3.11)\qquad (\lambda \circ f)(x) = \lambda^m f\left(\lambda^{-1}x\right) \quad (m = \deg f) .$$

It is easily verified that if $f, g, \ldots$ are the elementary divisors of A then $\lambda \circ f, \lambda \circ g, \ldots$ are those of λA . Hence, in the notation (3.1),

$$(3.12)\qquad \sigma(\lambda A) = \left(\lambda \circ f_1, \lambda \circ f_2, \ldots\right) .$$

THEOREM 2. *Let* $A \in GL_n(F)$, *where* F *is a field, and let* (3.1) *be the corresponding sequence of polynomials. Then*

$$(3.13)\qquad L_{GL}(A) = (F^*)_\delta ,$$

where δ *is the greatest positive integer such that* $f_r(x) \in F\left[x^\delta\right]$ *for all* r .

Proof. By (2.13), L is finite. Let ε be a primitive dth root of unity in F . By (3.12), $\varepsilon \in L$ if, and only if,

(a) $\varepsilon \circ f_r = f_r$ for all r .

We prove the theorem by showing that (a) is equivalent to

(b) $f_r(x) \in F\left[x^d\right]$ for all r .

Now, $f_r(x)$ has the form $x^m + a_1x^{m-1} + \ldots + a_m$, where $a_m \neq 0$. The equation $\varepsilon \circ f_r = f_r$ means that $a_t = \varepsilon^t a_t$ for all t and thus that $a_t = 0$ except when $d|t$. However, since $a_m \neq 0$, this is equivalent to $f_r(x) \in F\left[x^d\right]$. Thus, (a) and (b) are equivalent and the theorem is proved.

THEOREM 3. *Let* $A \in GL_n(F)$, *where* F *is a field. Then* $\det \Gamma_{GL}(A)$ *is generated by* $\det C_{GL}(A)$ *and* $(-1)^{n(n-1)/l}$, *where* $l = |L_{GL}(A)|$.

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

Proof. We have

$$(-1)^{n(n-1)/l} = \varepsilon^{\binom{n}{2}} ,$$

where ε is a primitive lth root of unity in F. The theorem will be established by proving the existence of a matrix T over F such that

$$(3.14) \qquad TAT^{-1} = \varepsilon A , \quad \det T = \varepsilon^{\binom{n}{2}} .$$

Consider an elementary divisor h_1 of A. Let $h_1, \ldots, h_s$ be the distinct members of its orbit under the action of $L = L_{GL}(A)$ given by (3.11). Since $\varepsilon \in L$, A is similar to εA and so all h_i have the same multiplicity as elementary divisors of A. Since the h_i are relatively prime in pairs, the direct sum of the companion matrices of the h_i is similar to the companion matrix of their product $h = h_1 \ldots h_s$. Clearly, $h(x)$ has the form $g\left(x^l\right)$, where $g(x) \in F[x]$. These considerations show that we may assume that A is the block diagonal matrix $\operatorname{diag}\left(A_1, \ldots, A_t\right)$, where each A_r is the companion matrix of a polynomial $g_r\left(x^l\right)$ with $g_r(x) \in F[x]$.

Suppose that, for each r, we have found a matrix T_r such that

$$T_r A_r T_r^{-1} = \varepsilon A_r , \quad \det T_r = \varepsilon^{\binom{n_r}{2}} ,$$

where $A_r \in GL_{n_r}(F)$. Then the block diagonal matrix $T = \operatorname{diag}\left(T_1, \ldots, T_t\right)$ satisfies $TAT^{-1} = \varepsilon A$ and

$$\det T = \prod \varepsilon^{\binom{n_r}{2}} = \varepsilon^{\binom{n}{2}} ,$$

since $n = \sum n_r$ and each n_r is divisible by l. Thus, it is sufficient to establish the existence of a matrix T satisfying (3.14) when A itself is the companion matrix of a polynomial $f(x) = g\left(x^l\right)$, where

$g(x) \in F[x]$.

Let K be the commutative ring obtained by adjoining a root α of $f(x)$ to F . Then K is an n-dimensional vector space over F with basis $1, \alpha, \ldots, \alpha^{n-1}$. The mapping $\hat{\alpha} : K \to K$, $u \mapsto \alpha u$, is an F-linear transformation and its matrix with respect to the above basis is A . It is therefore sufficient to prove that $\hat{\alpha}$ is similar to $\varepsilon\hat{\alpha}$ by a linear transformation τ of determinant $\varepsilon^{\binom{n}{2}}$. Now, since $f(x) \in F[x^l]$, there is an automorphism of K which carries α to $\varepsilon\alpha$ and fixes the elements of F . We take τ to be this automorphism. Since $\tau(\alpha^i) = \varepsilon^i\alpha^i$, the determinant of τ is

$$\varepsilon^{1+2+\ldots+(n-1)} = \varepsilon^{\binom{n}{2}} ,$$

as required. This completes the proof.

COROLLARY. *The quotient group* $\det \Gamma_{GL}(A)/\det C_{GL}(A)$ *has order* 1 *or* 2 . *It has order* 2 *if, and only if,* $-1 \notin \det C_{GL}(A)$ *and* n/l *is odd.*

Proof. Theorem 3 shows at once that if $-1 \in \det C$ then $\det C = \det \Gamma$. Suppose therefore that $-1 \notin \det C$. By (2.13), n/l is an integer. If n/l is even, then $(-1)^{n(n-1)/l} = 1$ and so, by Theorem 3, $\det C = \det \Gamma$. Suppose therefore that n/l is odd. Then $(-1)^{n(n-1)/l} = (-1)^{n-1}$. Since $-I \in C$ but $-1 \notin \det C$, it follows that n is even and thus that $(-1)^{n-1} = -1$. Since $-1 \notin \det C$, Theorem 3 now shows that $\det C$ has index 2 in $\det \Gamma$. This completes the proof.

We have just shown that the quotient group $\det \Gamma/\det C$ has order 1 or 2 . The following example shows that both values are possible. Recall that $\det \Gamma/\det C \cong L/\Lambda$.

EXAMPLE. Take

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in SL_2(F) .$$

It is easy to see that $L = \{1, -1\}$. Now, the general solution of

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

$TA = -AT$ is

$$T = \begin{bmatrix} a & b \\ b & -a \end{bmatrix} ,$$

where clearly $\det T = -(a^2+b^2)$. If $F = \mathbb{R}$, $\det T = 1$ has no solution a, b in F and thus $\Lambda = \{1\}$. On the other hand, if $F = \mathbb{C}$, $\det T = 1$ has a solution in F and so $\Lambda = L$.

To conclude this section, we consider the case of a finite coefficient field.

THEOREM 4. *Let $A \in GL_n(\mathbb{F}_q)$ and let* (3.1) *be the corresponding sequence of polynomials. Then*

(3.15) $$\det C_{GL}(A) = \det \Gamma_{GL}(A) = (\mathbb{F}_q^*)^{d(A)} ,$$

(3.16) $$L_{GL}(A) = \Lambda_{GL}(A) = (\mathbb{F}_q^*)_{\delta(A)} ,$$

where

(3.17) *$d(A)$ is the greatest divisor d of $q - 1$ such that $f_r(x) = 1$ whenever r is not a multiple of d , and*

(3.18) *$\delta(A)$ is the greatest divisor δ of $q - 1$ such that $f_r(x) \in F[x^\delta]$ for all r .*

Proof. In view of (2.14), it is sufficient to prove that

(a) $\det C = (F^*)^{d(A)}$,

(b) $L = (F^*)_{\delta(A)}$,

(c) $\det C = \det \Gamma$,

where $F = \mathbb{F}_q$.

Proof of (a). Let the elementary divisors of A be $\pi_1^{r_1}, \pi_2^{r_2}, \ldots$, where $\pi_1, \pi_2, \ldots$ are irreducible. Then

$$d(A) = (q-1, d) ,$$

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

where d is the greatest common divisor of $r_1, r_2, \ldots$. Moreover, since F is finite, $\Delta(\pi_i) = F^*$ for all i . Therefore, by Theorem 1,

$$\det C = \prod_i (F^*)^{r_i} = (F^*)^d = (F^*)^{d(A)} .$$

Proof of (b). We have

$$\delta(A) = (q-1, \delta) ,$$

where δ is the greatest positive integer such that $f_r(x) \in F[x^\delta]$ for all r . By Theorem 2,

$$L = (F^*)_\delta = (F^*)_{\delta(A)} .$$

Proof of (c). By the Corollary to Theorem 3, it is sufficient to prove that if $-1 \notin \det C$ then $n/\delta(A)$ is even. Now, if $-1 \notin \det C$, then $d(A)$ is even by (a). Thus, it will be sufficient to prove that

(3.19) $$d(A)\delta(A) \mid n .$$

In the formula (3.2) for $n = n(A)$, $\delta(A) \mid \deg f_r$ for all r , and $\deg f_r = 0$ unless $r \mid d(A)$. Thus, all terms $r \deg f_r$ are divisible by $d(A)\delta(A)$ and (3.19) follows. This proves the theorem.

COROLLARY. *If* $A \in GL(n, \mathbb{F}_q)$ *, then*

(3.20). $$\begin{cases} (A)_{GL} \text{ splits into } d(A) \text{ } SL\text{-classes,} \\ (A)_{PGL} \text{ splits into } d(A) \text{ } PSL\text{-classes,} \\ (A)_{PGL} \text{ splits into } (q-1)/\delta(A) \text{ } GL\text{-classes,} \\ (A)_{PSL} \text{ splits into } (q-1)/\delta(A) \text{ } SL\text{-classes.} \end{cases}$$

Proof. This follows at once from the theorem and (2.7), (2.8), (2.10), (2.11).

4. Enumeration, duality

In this section, F will be the finite field $\mathbb{F}_q$. Let $A \in GL_n = GL_n(F)$. With (3.20) in mind, we call $d(A)$ the *special index*, or *s-index*, of A and $\delta(A)$ the *projective index*, or *p-index*, of A. By (3.17)-(3.19), both are divisors of $q-1$ and their product is a divisor of n. Since all matrices in the one *PGL*-class have the same *s*- and *p*-indices, we may speak of the *s*- and *p*-indices of a *PGL*-class or of any subset such as a *GL*-class.

We begin with a detailed discussion of the results to be proved. Write

$$(4.1)\qquad C(t) = \prod_{r=1}^{\infty} \left((1-t^r)/(1-qt^r)\right) .$$

Feit and Fein [1] showed that

$$(4.2)\qquad C(t) = 1 + \sum_{n=1}^{\infty} c_n t^n ,$$

where c_n is the number of conjugacy classes in GL_n. Our first result is that, if d and δ are divisors of $q-1$,

$$(4.3)\qquad 1 + \sum_{n=1}^{\infty} \gamma_n(d, \delta)t^n = C\left(t^{d\delta}\right) ,$$

where $\gamma_n(d, \delta)$ is the number of conjugacy classes in GL_n with *s*-index divisible by d and *p*-index divisible by δ. In other words, if d and δ are divisors of $q-1$,

$$(4.4)\qquad \gamma_n(d, \delta) = \begin{cases} c_{n/d\delta} & \text{if } d\delta \mid n , \\ 0 & \text{otherwise.} \end{cases}$$

This result enables us to determine the number, $c_n(d, \delta)$, of conjugacy classes in GL_n with *s*-index d and *p*-index δ. Assuming still that d and δ are divisors of $q-1$, we have

$$(4.5)\qquad \gamma_n(d, \delta) = \sum_{\substack{D,\Delta \\ d|D|q-1 \\ \delta|\Delta|q-1}} c_n(D, \Delta) ,$$

whence, by the Möbius inversion formula,

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

(4.6) $$1 + \sum_{n=1}^{\infty} c_n(d, \delta)t^n = \sum_{\substack{D,\Delta \\ d|D|q-1 \\ \delta|\Delta|q-1}} \mu(D/d)\mu(\Delta/\delta)C\left(t^{D\Delta}\right) .$$

More explicitly, if $d\delta$ is not a divisor of n, then

(4.7) $$c_n(d, \delta) = 0 ,$$

and if $d\delta$ is a divisor of n,

(4.8) $$c_n(d, \delta) = \sum_{\substack{D,\Delta \\ d|D|(n,q-1) \\ \delta|\Delta|(n,q-1) \\ D\Delta|n}} \mu(D/d)\mu(\Delta/\delta)c_{n/D\Delta} .$$

In particular, we see that

(4.9) $$c_n(d, \delta) = c_n(\delta, d) .$$

This simple result is the source of the later results on duality.

The above formulae can be generalised. Let D, E be divisors of $q - 1$. We introduce the following subgroups of GL_n :

(4.10) $$P^D = P^D(n) = \{A : \det A \in (F^*)^D\} ,$$

(4.11) $$Z_E = Z_E(n) = \{\lambda I : \lambda \in (F^*)_E\} .$$

If, in addition, $DE|n(q-1)$, then $Z_E \subseteq P^D$ and we may form the quotient group

(4.12) $$P_E^D = P_E^D(n) = P^D/Z_E .$$

Each subgroup G of GL_n acts on P_E^D by conjugation. Slightly extending the notation of Section 2, we call the orbits *G-classes*. Then the duality theorem of Lehrer cited in the introduction asserts that *if* $DE|q-1$ *then the numbers of* GL_n*-classes in* $P_E^D(n)$ *and* $P_D^E(n)$ *are the same.*

In practice it is convenient to deal, not with the G-classes in P_E^D themselves, but rather with their inverse images under the canonical

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

homomorphism $P^D \to P^D_E$. These are the sets

(4.13) $$(A)_{E;G} \quad \left(A \in P^D\right) ,$$

where

(4.14) $$(A)_{E;G} = \left\{\lambda TAT^{-1} : \lambda \in (F^*)_E, \ T \in G\right\} .$$

Since $(A)_{E;G} \subseteq (A)_{PGL}$, we may speak of the s- and p-indices of $(A)_{E;G}$ and hence of the corresponding G-class. Our second result is that

(4.15) *if D, E, d, δ, Δ are divisors of $q - 1$ such that $DE|n(q-1)$, then the number of P^Δ-classes in $P^D_E(n)$ with s-index d and p-index δ is*

$$\frac{(d,D)(\delta,E)}{DE}(d, \Delta)c_n(d, \delta) .$$

Some special cases are of interest. When $D = \Delta$, we get the number of conjugacy classes in $P^D_E(n)$ with s-index d and p-index δ . Specialising even further, we get the numbers of such conjugacy classes in PGL_n, SL_n and PSL_n . The *total* numbers of conjugacy classes in the latter groups are discussed in detail by Macdonald in [5].

Again, taking $\Delta = 1$ in (4.15) and using (4.9), we deduce that

(4.16) *if d, δ, D, E are divisors of $q - 1$ such that $DE|n(q-1)$, then the number of GL_n-classes in $P^D_E(n)$ with s-index d and p-index δ is equal to the number of GL_n-classes in $P^E_D(n)$ with s-index δ and p-index d .*

This clearly implies Lehrer's theorem.

Another consequence of (4.15) is the following:

(4.17) *Let D, E, Δ be divisors of $q - 1$ and suppose that $DE|n(q-1)$. Then the total number of P^Δ-classes in $P^D_E(n)$ is*

$$(DE)^{-1} \sum_{\substack{d,\delta \\ \delta|E, d|[D,\Delta] \\ d\delta|n}} \phi(d_1)\phi_2(d_2)d_3^2\phi(\delta)c_{n/d\delta} ,$$

where $d_1 = d/(d, D, \Delta)$, d_2 *is the largest divisor of* d *relatively prime to* d_1 *and* $d_1d_2d_3 = d$.

Here, $\phi(m) = \phi_1(m)$ and $\phi_2(m)$ are the Eulerian functions defined by

$$\phi_r(m) = m^r \prod_p \left(1-\left(1/p^r\right)\right) ,$$

where summation is over the distinct prime divisors p of m . We shall pass over the proof of (4.17) except to mention that it depends on summing $\sum_{d|m} \mu(m/d)\left(d, m_1\right)\left(d, m_2\right)$ in closed form.

Taking $\Delta = 1$ in (4.17), we get the following formula for the number of GL_n-classes in $P_E^D(n)$:

$$(DE)^{-1} \sum_{\substack{d,\delta \\ d|D, \delta|E \\ d\delta|n}} \phi(d)\phi(\delta)c_{n/d\delta} .$$

It follows that, for given n and q , *the number of* GL_n*-classes in* $P_E^D(n)$ *depends only on* DE *and* (D, E) . This is a slight generalisation of Lehrer's theorem.

Similar results can be proved by similar methods for the groups of F-rational points of the connected algebraic groups isogenous to $SL_n(\overline{F})$, where $\overline{F}$ is the algebraic closure of F . These have the same order as SL_n and include both SL_n and PGL_n as special cases. In the formulation (but not the notation) of Macdonald [5], they appear as the quotient groups

(4.18) $$Q_e = Q_e(n) = R_e/S_e \ (e|n) ,$$

where R_e, S_e are the following subgroups of $GL_n \times F^*$:

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

$$R_e = \{(X, \lambda) : \det X = \lambda^e\} ,$$

$$S_e = \{(\alpha I, \alpha^{n/e}) : \alpha \in F^*\} .$$

The action of GL_n on Q_e is defined *via* the embedding $GL_n \to GL_n \times F^*$, $X \mapsto (X, 1)$. If G is a subgroup of GL_n, the orbits under the action of G on Q_e are again called G-classes. The duality theorem of Macdonald referred to in the introduction asserts that *if* $ef = n$ *then the numbers of* GL_n*-classes in* $Q_e(n)$ *and* $Q_f(n)$ *are the same.* Indeed the explicit formula used to prove this result yields the slightly stronger result that *if* $ef = n$ *then the numbers of* GL_n*-classes in* $Q_e(n)$ *and* $Q_{(e,f,q-1)}(n)$ *are the same.*

The p- and s-indices of the G-class of an element $(X, \lambda)S_e$ of Q_e are defined as the p- and s-indices of X. The following result is the analogue of (4.15).

(4.19) *Let* d, δ, Δ *be divisors of* $q - 1$ *and* e, f *positive integers such that* $ef = n$. *Then the number of* P^Δ*-classes in* $Q_e(n)$ *with s-index* d *and p-index* δ *is*

$$\frac{(d,e)(\delta,f)}{(q-1)} (d, \Delta)c_n(d, \delta) .$$

Special cases are again of interest. When $\Delta = (e, p-1)$, the P^Δ-classes in $Q_e(n)$ become the conjugacy classes of $Q_e(n)$. Again, taking $\Delta = 1$ we get the following analogue of (4.16).

(4.20) *Let* $d, \delta | q-1$ *and* $ef = n$. *Then the number of* GL_n*-classes in* $Q_e(n)$ *with s-index* d *and p-index* δ *is equal to the number of* GL_n*-classes in* $Q_f(n)$ *with s-index* δ *and p-index* d.

Macdonald's duality theorem is an immediate consequence. Finally, we have the following analogue of (4.17).

(4.21) *Let* $\Delta | q-1$ *and* $ef = n$. *Then the total number of*

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

P^{Δ}-classes in $Q_e(n)$ *is*

$$(q-1)^{-1} \sum_{\substack{d,\delta | q-1 \\ \delta | f, d | [e,\Delta] \\ d\delta | n}} \phi(d_1)\phi_2(d_2)d_3^2\phi(\delta)c_{n/d\delta} ,$$

where $d_1 = d/(d, e, \Delta)$ *,* d_2 *is the largest divisor of* d *relatively prime to* d_1 *and* $d_1d_2d_3 = d$ *.*

EXAMPLE. Suppose that $e|n|q-1$ and write $ef = n$, $mn = q - 1$. Each of the groups $Q_e(n)$ and $P^e_{(q-1)/e}(n)$ is an extension of SL_n/Z_f by a cyclic group of order f . Ketter and Lehrer [3] carried out computer calculations to determine the numbers of GL_n-classes in these groups for certain e, n and q . Now, (4.21) and (4.17) show that these numbers are, respectively,

$$M_e = \sum_{d|e,\delta|f} \phi(d)\phi(\delta)c'_{n/d\delta} ,$$

$$N_e = \sum_{d|e,\delta|(m,e/d)f} \phi(d)\phi(\delta)c'_{n/d\delta} ,$$

where $c'_r = c_r/(q-1)$. Macdonald [5] proved the formula for M_e and tabulated c'_r for $r \le 12$. Splitting d, δ into their prime-powers, one see that

$$M_e = M_{(e,f)}$$

and so, in particular, $M_e = M_f$. Further

$$N_e \ge M_e$$

and, as Ketter and Lehrer observed,

$$N_e = M_e \text{ if } (m, e) = 1 .$$

For $n = 4$, we have

$$M_1 = N_1 = M_4 = q^3 + q^2 + 2q + 3 ,$$

$$M_2 = M_1 + q ,$$

$N_2 - M_2 = 0$ or 2 according as $(m, 2) = 1$ or 2 ,

$N_4 - M_4 = 0, q + 2$ or $q + 4$ according as $(m, 4) = 1, 2$ or 4 .

Our results agree with those of Ketter and Lehrer except in one case: the values when $n = 6$, $q = 13$, should be

$$M_1 = M_2 = M_3 = M_6 = N_1 = N_3 = 402\ 432 ,$$

$$N_2 = N_6 = 402\ 616 .$$

The proofs of the results for the groups $Q_e(n)$ will be omitted. It remains to prove the key results (4.3) and (4.15). In each case, some preparation is necessary.

Let P be a set of monic polynomials over F such that $1 \in P$. The *generating function for* P is defined to be the power series

$$g_P(t) = \sum_{n=0}^{\infty} g_n(P)t^n ,$$

where $g_n(P)$ denotes the number of elements of P of degree n . (Notice that $g_0(P) = 1$ since $1 \in P$.) In the same way, if Q is a set of similarity classes of square matrices over F , the *generating function for* Q is the power series

$$G_Q(t) = 1 + \sum_{n=1}^{\infty} G_n(Q)t^n ,$$

where $G_n(Q)$ denotes the number of similarity classes of $n \times n$ matrices in Q . The following enumerative principle is due to Feit and Fein [1].

LEMMA. *Given sets* $P_1, P_2, \ldots$ *of monic polynomials over* F *such that* $1 \in P_r$ *for all* r , *let* Q *be the set formed by those similarity classes of matrices over* F *whose associated sequences of polynomials* $(f_1(x), f_2(x) \ldots)$ (*in the sense of* (3.1)) *satisfy* $f_r(x) \in P_r$ *for all*

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

r . *Then*

(4.22) $$G_Q(t) = \prod_{r=1}^{\infty} g_{P_r}(t^r) .$$

Proof. It follows from (3.2) and the definition of Q that

$$G_n(Q) = \sum_{\substack{n_1, n_2, \ldots \geq 0 \\ \Sigma r n_r = n}} g_{n_1}(P_1) g_{n_2}(P_2) \ldots ,$$

which is equivalent to (4.22).

Proof of (4.3). If M is the set of all monic polynomials in $F[x^\delta]$ which are not divisible by x , then

$$g_M(t) = 1 + (q-1)t^\delta + (q-1)qt^{2\delta} + \ldots$$
$$= (1-t^\delta)/(1-qt^\delta) .$$

Let us now choose $P_1, P_2, \ldots$ in the lemma as follows:

$$P_r = \begin{cases} M & \text{if } d | r , \\ \{1\} & \text{otherwise.} \end{cases}$$

Then, by (3.17) and (3.18), the resulting set Q is the set S_δ^d of all GL-classes with s-index divisible by d and p-index divisible by δ . The Lemma gives

$$G_{S_\delta^d}(t) = \prod_{\substack{r \\ d|r}} g_M(t^r) = C(t^{d\delta}) ,$$

which is just another way of writing (4.3). This completes the proof.

Further preparation is needed for the proof of (4.15). We introduce certain unions of GL-classes within which the distribution of GL-classes according to determinant can be simply described.

Consider the sequence of polynomials (3.1) associated with a given $A \in GL_n$. Write each component polynomial down explicitly in the form:

(4.23) $$f_r(x) = x^{n_r} + (-1)^{n_r - i} a_{ri} x^i + (-1)^{n_r - j} a_{rj} x^j + \ldots ,$$

where $n_r > i > j > \ldots$ and all coefficients $a_{ri}, a_{rj}, \ldots$ are non-zero. The index set

$$\Omega_r(A) = \{n_r, i, j, \ldots\}$$

is finite and non-empty with greatest and least members n_r and 0, and $\Omega_r(A) = \{0\}$ for almost all r. The sequence

(4.24) $$\Omega(A) = (\Omega_1(A), \Omega_2(A), \ldots)$$

will be called the *support* of A. Since all matrices in the one *PGL*-class have the same support, we may speak of the support of a *PGL*-class or of any subset such as a *GL*-class.

Consider now the set $T(\Omega)$ of all matrices A having a given support

(4.25) $$\Omega = (\Omega_1, \Omega_2, \ldots) ,$$

where each Ω_r is a finite set such that $0 \in \Omega_r$ and where almost all (but *not* all) Ω_r are $\{0\}$. By (3.2), all matrices in $T(\Omega)$ have the same dimension, namely,

(4.26) $$n(\Omega) = \sum_r r n_r(\Omega) ,$$

where $n_r(\Omega)$ denotes the greatest member of Ω_r. By (3.17) and (3.18) we have

(4.27) $$d(A) = \big(d(\Omega(A)), q-1\big) , \quad \delta(A) = \big(\delta(\Omega(A)), q-1\big) ,$$

where $d(\Omega)$ denotes the greatest common divisor of the indices r for which $\Omega_r \neq \{0\}$ and $\delta(\Omega)$ the greatest common divisor of the elements of $\bigcup_r \Omega_r$.

LEMMA. *The GL-classes of matrices which make up* $T(\Omega)$ *can be parametrized by the elements of an abelian group* $H(\Omega)$ *in such a way that the mapping which assigns to each element of* $H(\Omega)$ *the determinant of the matrices in the corresponding GL-class is a group homomorphism mapping*

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

$H(\Omega)$ *onto* $(F^*)^{d(\Omega)}$

Proof. The GL-classes $(A)_{GL}$, $A \in T(\Omega)$, are already parametrized by the corresponding rows of polynomials (3.1). Replacing each $f_r(x)$ by the corresponding row of coefficients $(a_{ri}, a_{rj}, \ldots)$ (see (4.23)), we get a row of

$$N(\Omega) = \sum_r \left(|\Omega_r|-1\right)$$

non-zero elements of F , that is, an element of the direct product of $N(\Omega)$ copies of the group F^* . This direct product is the parameter group $H(\Omega)$. Let $r_1, \ldots, r_s$ be the indices r for which $\deg f_r(x) > 0$. Then, by (3.3),

$$\det A = \delta\left(f_{r_1}\right)^{r_1} \ldots \delta\left(f_{r_s}\right)^{r_s} .$$

At the same time, by (4.23), the element of $H(\Omega)$ corresponding to $(A)_{GL}$ has the form

$$\left(\ldots, \delta\left(f_{r_1}\right), \ldots, \delta\left(f_{r_2}\right), \ldots, \delta\left(f_{r_s}\right), \ldots\right) .$$

It follows that the mapping described in the lemma is indeed a homomorphism $H(\Omega) \to F^*$ and that the image of $H(\Omega)$ is

$$\prod_{i=1}^{s} (F^*)^{r_i} = (F^*)^{d(\Omega)} .$$

This proves the lemma.

We are now in a position to prove (4.15). Let D, E, Δ, d, δ be divisors of $q - 1$ with $DE|n(q-1)$. The set $(A)_{E;G}$ in (4.13) with $G = P^\Delta$ becomes

$$(A)_{E,\Delta} = \left\{\lambda TAT^{-1} : \lambda \in (F^*)_E, \det T \in (F^*)^\Delta\right\} .$$

Let $T_\delta^d(n)$ denote the set of all matrices in GL_n with s-index d and p-index δ . Then the numerical restriction $DE|n(q-1)$ guarantees that

$T_\delta^d(n) \cap P^D(n)$ is a disjoint union of sets of the form $(A)_{E,\Delta}$, and (4.15) is equivalent to the assertion that the number of such sets is

$$\frac{(d,D)(\delta,E)}{DE}(d,\ \Delta)c_n(d,\ \delta)\ .$$

We shall prove this last result by showing that:

(a) *if* $A \in T_\delta^d(n)$ *then* $(A)_{E,\Delta}$ *is a union of* $dE/(d,\ \Delta)(\delta,\ E)$ *Sl-classes;*

(b) $T_\delta^d(n) \cap P^D(n)$ *is a union of* $d(d,\ D)c_n(d,\ \delta)/D$ *SL-classes.*

Proof of (a). The direct product $F^* \times F^*$ acts on the set of all *SL*-classes by the rule

$$(\lambda,\ \mu) \circ (A)_{SL} = \left(\lambda TAT^{-1}\right)_{SL}\ ,\ \text{where}\ \det T = \mu\ .$$

The orbit of $(A)_{SL}$ consists of those *SL*-classes which make up $(A)_{PGL}$. Since $A \in T_\delta^d(n)$, it follows from (3.20) that the number of such *SL*-classes is $(q-1)d/\delta$. Since $(F^*)_\delta \times (F^*)^d$ is contained in the stabilizer of $(A)_{SL}$ and has index $(q-1)d/\delta$ in $F^* \times F^*$, it must indeed be the stabilizer.

On the other hand, the *SL*-classes making up $(A)_{E,\Delta}$ form the orbit of $(A)_{SL}$ under the action of $(F^*)_E \times (F^*)^\Delta$. Therefore the number of *SL*-classes into which $(A)_{E,\Delta}$ splits is

$$\left|(F^*)_E : (F^*)_E \cap (F^*)_\delta\right|\left|(F^*)^\Delta : (F^*)^\Delta \cap (F^*)^d\right| = \big(E/(\delta,\ E)\big)\big(d/(d,\ \Delta)\big)$$
$$= dE/(d,\ \Delta)(\delta,\ E)\ .$$

Proof of (b). By (4.26) and (4.27), $T_\delta^d(n)$ is the disjoint union of those sets $T(\Omega)$ in the lemma which satisfy

$$n(\Omega) = n\ ,\quad \big(d(\Omega),\ q-1\big) = d\ ,\quad \big(\delta(\Omega),\ q-1\big) = \delta\ .$$

By that lemma, the proportion of *GL*-classes in such a $T(\Omega)$ having

https://doi.org/10.1017/S0004972700006675 Published online by Cambridge University Press

determinant in $(F^*)^D$ is

$$|(F^*)^{d(\Omega)} : (F^*)^{d(\Omega)} \cap (F^*)^D|^{-1} = |(F^*)^d : (F^*)^d \cap (F^*)^D|^{-1}$$
$$= (d, D)/D .$$

Therefore the number of GL-classes in $T_\delta^d(n) \cap P^D(n)$ is $(d, D)c_n(d, \delta)/D$. On the other hand, by (3.20), each GL-class in $T_\delta^d(n)$ splits into d SL-classes. It follows that $T_\delta^d(n) \cap P^D(n)$ splits into $d(d, D)c_n(d, \delta)/D$ SL-classes, as we had to prove. The proof of (4.15) is now complete.

References

[1] Walter Feit and N.J. Fine, "Pairs of commuting matrices over a finite field", *Duke Math. J.* 27 (1960), 91-94.

[2] Nathan Jacobson, *The theory of rings* (American Mathematical Society Mathematical Surveys, 2. American Mathematical Society, New York, 1943).

[3] T.A. Ketter and G.I. Lehrer, "On conjugacy classes in certain isogenous groups", *Bull. Austral. Math. Soc.* 14 (1976), 371-377.

[4] G.I. Lehrer, "Characters, classes, and duality in isogenous groups", *J. Algebra* 36 (1975), 278-286.

[5] I.G. Macdonald, "Numbers of conjugacy classes in some finite classical groups", *Bull. Austral. Math. Soc.* 23 (1981), 23-48.

Department of Pure Mathematics,
University of Sydney,
Sydney,
New South Wales 2006,
Australia.