PERMUTATION MATRICES AND MATRIX EQUIVALENCE OVER A FINITE FIELD

GARY L. MULLEN

Department of Mathematics
The Pennsylvania State University
Sharon, Pennsylvania 16146

(Received March 21, 1980 and in revised form August 12, 1980)

ABSTRACT. Let \( F = \text{GF}(q) \) denote the finite field of order \( q \) and \( F_{mn}^{mxn} \) the ring of \( m \times n \) matrices over \( F \). Let \( P_n \) be the set of all permutation matrices of order \( n \) over \( F \) so that \( P_n \) is isomorphic to \( S_n \). If \( \Omega \) is a subgroup of \( P_n \) and \( A, B \in F_{mn}^{mxn} \) then \( A \) is equivalent to \( B \) relative to \( \Omega \) if there exists \( P \in P_n \) such that \( AP = B \). In sections 3 and 4, if \( \Omega = P_n \), formulas are given for the number of equivalence classes of a given order and for the total number of classes. In sections 5 and 6 we study two generalizations of the above definition.

KEY WORDS AND PHRASES. Permutation matrix, equivalence, automorphism, finite field.


1. INTRODUCTION.

In a series of papers [1-4,6-8] L. Carlitz, S. Cavior, and the author studied various forms of equivalence of functions over a finite field through the use of permutation groups acting on the field itself. In [9] the author defined two matrices \( A \) and \( B \) to be equivalent if \( b_{ij} = \phi(a_{ij}) \) for some permutation \( \phi \) of the field while in [10] \( B \) was said to be equivalent to \( A \) if \( B = \phi(A) \) where \( \phi(A) \) was computed by substitution. In the present paper we study another form of matrix equivalence over a finite field through the use of permutation matrices and the Pólya-deBruijn theory of enumeration.

Let \( F = \text{GF}(q) \) denote the finite field of order \( q = p^b \), \( p \) is prime and \( b \geq 1 \) and let \( F_{mn}^{mxn} \) denote the ring of \( m \times n \) matrices over \( F \) so that \( |F_{mn}^{mxn}| = q^{mn} \). Let
$P_n$ be the set of all $n \times n$ matrices over $F$ consisting entirely of zeros and ones with the property that there is exactly one 1 in each row and column. In the literature, such matrices have been called permutation matrices. It is not hard to show that $P_n$ is a group under matrix multiplication which is isomorphic to $S_n$, the symmetric group on $n$ letters and consequently has order $n!$. If $P \in P_n$ the isomorphism can be defined as follows. If

$$P = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

then define $\phi_P \in S_n$ by $\phi_P(i) = \alpha_i (i = 1, \ldots, n)$. Then $\psi : P_n \rightarrow S_n$ defined by $\psi(P) = \phi_P$ is an isomorphism.

2. GENERAL THEORY.

If $\Omega$ is a subgroup of $P_n$ we may make

**DEFINITION 1.** If $A, B \in F_{mxn}$ then $B$ is equivalent to $A$ relative to $\Omega$ if there exists $P \in \Omega$ such that $AP = B$.

This is an equivalence relation on $F_{mxn}$ so we let $\mu(A, \Omega)$ denote the order of the class of $A$ relative to $\Omega$ and let $\lambda(\Omega)$ be the total number of classes induced by $\Omega$.

**THEOREM 2.1.** If $A, B \in F_{mxn}$ then $B$ is equivalent to $A$ relative to $\Omega$ if and only if the columns of $B$ are a permutation of the columns of $A$.

**PROOF.** Suppose $AP = B$ where $A = (a_{ij})$. In $P$ suppose that for $j = 1, \ldots, n$ the 1 in column $j$ occurs in row $i_j$. Then $AP = (a_{ij})P = (a_{ij})$ so that column $j$ of $A$ becomes column $i_j$ of $AP$.

Conversely, suppose column $j$ of $A$ is column $i_j$ of $B$. Define $P$ so that in column $j$ we have a 1 in row $i_j$ and zeros elsewhere. Then $P \in P_n$ and $AP = B$ so that $A$ is equivalent to $B$.

**COROLLARY 2.2.** If $A, B \in F_{nxn}$ and $B$ is equivalent to $A$ relative to $\Omega$ then $\det(B) = \pm \det(A)$.

In fact, if $AP = B$ and $P$ corresponds to $\phi_P \in S_n$ where $\phi_P$ is an even permutation then $\det(B) = \det(A)$ while if $\phi_P$ is an odd permutation then $\det(B) = -\det(A)$.
DEFINITION 2. If $A \in P_{mxn}$ then $P$ is an automorphism of $A$ relative to $\Omega$ if $P \in \Omega$ and $AP = A$.

If $\text{Aut}(A,\Omega)$ denotes the set of all automorphisms of $A$ relative to $\Omega$, then it is easy to check that $\text{Aut}(A,\Omega)$ is a group under matrix multiplication whose order will be denoted by $\nu(A,\Omega)$. It is easy to prove

THEOREM 2.3. If $A \in P_{mxn}$ then for any subgroup $\Omega$ of $P_n$

$$\mu(A,\Omega)\nu(A,\Omega) = |\Omega|,$$

(2.1)

where $|\Omega|$ denotes the order of $\Omega$.

If $P \in P_n$ let $N(P,m,n,q)$ denote the number of $m \times n$ matrices $A$ over $\text{GF}(q)$ such that $AP = A$.

THEOREM 2.4. If $P$ corresponds to $\phi \in S_n$ and $\phi$ has $\ell(P)$ distinct cycles then

$$N(P,m,n,q) = q^m\ell(P).$$

PROOF. Suppose the distinct cycles of $\phi$ are $C_1,\ldots,C_{\ell(P)}$. Using Theorem 2.1 it is clear that $AP = A$ if and only if within a given cycle of $\phi$ the columns of $A$ are identical. The theorem then follows from the fact that a given column can be constructed in $q^m$ ways.

3. CYCLIC GROUPS.

If $\Omega = <P>$ is a cyclic group of permutation matrices where $|\Omega| = s$, let $H(t)$ denote the subgroup of $\Omega$ of order $t$ where $t|s$ so that $H(t) = <P^{s/t}>$. If $P$ corresponds to $\phi \in S_n$ let $\ell_t(P)$ denote the number of cycles of $\phi_{P^{s/t}}$ and suppose $M(t,m,n,q)$ denotes the number of $m \times n$ matrices $A$ over $\text{GF}(q)$ such that $\text{Aut}(A,\Omega) = H(t)$.

THEOREM 3.1. For each divisor $t$ of $s$

$$M(t,m,n,q) = \sum_{a|\frac{s}{t}} \mu(a)q^{m\ell_t(P)},$$

(3.1)

where $\mu(a)$ is the Mobius function.

PROOF. By Theorem 2.4 $q^m\ell_t(P)$ counts the number of $m \times n$ matrices $A$ over $\text{GF}(q)$ such that $\text{Aut}(A,\Omega) \leq H(t)$. From this we subtract those for which the containment is proper. This number is given by

$$M(t,m,n,q) = q^m\ell_t(P) - \sum_{u|m,n,q} M(u,m,n,q),$$

(3.2)

where the sum is over all $u|s$, $t|u$ and $t \neq u$. After applying Mobius inversion
to (3.2) we obtain (3.1).

**COROLLARY 3.2.** For each divisor $t$ of $s$ there are $tM(t, m, n, q)/s$ classes of $s/t$ and

$$
\lambda(\Omega) = \frac{1}{s} \sum_{t|s} tM(t, m, n, q).
$$

(3.3)

As an illustration, suppose $q = 2$, $m = n = 3$, and

$$
P = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0
\end{bmatrix}
$$

so that if $\Omega = <P>$ then $|\Omega| = 3$. One can easily check that $M(3, 3, 3, 2) = 8$ and $M(1, 3, 3, 2) = 504$ so that there are 168 classes of order 3, 8 classes of order 1 and thus from (3.3), $\lambda(\Omega) = 176$.

4. **THE CASE $\Omega = P_n$.**

In this section we consider the group $P_n$ of all permutation matrices of order $n$ so that, as noted in the introduction, $P_n$ is isomorphic to $S_n$, the symmetric group on $n$ letters. We will employ the Pólya theory of enumeration to determine the number of classes induced by $P_n$. Suppose the permutation group $K$ acts on a set of $r$ elements. If $\pi \in K$ consider the monomial $b_1 x_1 x_2 \cdots x_r$ where for $t = 1, \ldots, r$, $b_t$ denotes the number of cycles of $\pi$ of length $t$. The polynomial

$$
P_K(x_1, \ldots, x_r) = |K|^{-1} \sum_{\pi \in K} b_1^{x_1} b_2^{x_2} \cdots b_r^{x_r}
$$

(4.1)

is called the **cycle index** of $K$. It is well known [5] that

$$
P_S(x_1, \ldots, x_n) = \Sigma(k_1^{k_1} k_2^{k_2 - 1} \cdots k_n^{k_n - 1} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n})
$$

where the sum is over all $k_1 + 2k_2 + \cdots + nk_n = n$.

In the Pólya theory of enumeration, let the domain $D$ be the set of $n$ columns and let the range $R$ be the set of $q^m$ possible column vectors so that $|D|^m = q^m = |R|$. If $K$ is a permutation group acting on $D$ then Pólya's theorem [5, p. 157] states that the number of distinct classes is given by $P_K(|R|, \ldots, |R|)$ so that $\lambda(P_n) = P_S(q^m, \ldots, q^m)$. It follows directly from Theorem 2.1 that $\lambda(P_n)$ is also the number of distributions of $n$ indistinguishable objects into $q^m$ labelled cells, or $\left(\begin{array}{c}
n + q^m - 1 \\
q^m - 1
\end{array}\right)$ so that we have proven
Theorem 4.1. If $\lambda(P_n)$ is the number of classes induced by $P_n$ then

$$\lambda(P_n) = (n + q^m - 1).$$

Suppose $A \in \mathbb{F}_{mxn}^m$ has $t$ distinct columns so that we have a partition of $n$ with $t$ parts say $n = m_1 + \ldots + m_t$ where each distinct column occurs $m_i$ times. By Theorem 2.1 for each such $A$ we have $\nu(A, P_n) = \prod_{i=1}^{t} m_i!$ so that by (2.1) $\mu(A, P_n) = (m_1, \ldots, m_t)$. The number of such $A$ is the same as the number of functions from $D$ into $R$ whose range is of size $q^m$, whose domain is of size $n$ and whose preimage partition has type $m_1 + \ldots + m_t = n$. We may rewrite this with distinct $m$'s say $j_{m_1} m_1 m_1 + \ldots + j_{m_s} m_s = n$ where $j_{m_1} + \ldots + j_{m_s} = t$. Then the number of such functions is $(q^m)^t h(j_{m_1}, \ldots, j_{m_s})$ where $h(j_{m_1}, \ldots, j_{m_s})$ is the number of partitions of $n$ of type $j_{m_1} m_1 + \ldots + j_{m_s} m_s = n$ and is given by Cauchy's formula

$$h(j_{m_1}, \ldots, j_{m_s}) = n'/(m_1')! \cdot (j_{m_1'})! \cdot \ldots \cdot (m_s')! h(j_{m_s')!})$$

and $(q^m)^t = q^m (q^m - 1) \ldots (q^m - t + 1)$ is the falling factorial which assigns image values to the partition blocks. Hence we have proven

**COROLLARY 4.2.** The number of classes induced by $P_n$ of order $(m_1, \ldots, m_s)$ is

$$\binom{m}{t} (j_{m_1}, \ldots, j_{m_s}).$$

As an illustration of the above theory suppose $q = 2$ and $m = n = 3$ so that we are considering the $512$ $3 \times 3$ matrices over GF(2) under the action of the symmetric group $S_3$. Thus from Corollary 4.2 when $t = 1$ we have $n = 3$ so that there are $8(1) = 8$ classes of order 1, when $t = 2$ we have $n = 1+2$ so that there are $8(2) = 56$ classes of order 3 and when $t = 3$ we have $n = 1 + 1 + 1$ so that there are $8(3) = 56$ classes of order 6 so that $\lambda(P_3) = 120$. Moreover, from Theorem 4.1 we also see that $\lambda(P_3) = \binom{10}{3} = 120$ classes.

5. A GENERALIZATION

In this section we generalize Definition 1 by considering a notion of matrix equivalence which is similar to the idea of weak equivalence of functions over a finite field considered by Cavior and the author in [3] and [8]. Let $P_m$ be the group of $m \times m$ permutation matrices over GF(q). If $\Omega_1$ is a subgroup of $P_m$ and $\Omega_2$ is a subgroup of $P_n$ we may make
DEFINITION 3. If \( A, B \in \mathbb{F}_{mn} \) then \( B \) is equivalent to \( A \) relative to \( \Omega_1 \) and \( \Omega_2 \) if there exist \( Q \in \Omega_1 \) and \( P \in \Omega_2 \) such that \( QAP = B \).

Thus \( P \in \mathbb{P}_n \) permutes the columns of \( A \) while \( Q \in \mathbb{P}_m \) permutes the rows of \( A \) so that \( \Omega_1 \) acts as a permutation group on the range \( R \) and \( \Omega_2 \) is a permutation group acting on the domain \( D \). Clearly if \( \Omega_1 = \{\text{id.}\} \) we obtain the previous cases considered in sections 3 and 4. In this more general setting we will make use of the extended Pólya theory of enumeration.

THEOREM 5.1. (Polya-deBruijn) The number of classes induced by permutation groups \( \Omega_2 \) of \( D \) and \( \Omega_1 \) of \( R \) is

\[
P_{\Omega_2}(\sum_{q=1}^{\infty} q^{z_1}z_2^2 \ldots)P_{\Omega_1}(e^{z_1}e^{2z_2} \ldots) = z_1^2z_2^2 \ldots = 0
\]

Consider the \( q^m \) possible column vectors of \( R \) in an \( m \times q^m \) array so that in row \( i \), we have \( q^{m-i+1} \) sets where in each set one element of \( \text{GF}(q) \) is repeated \( q^{i-1} \) times. For example, if \( q = 2 \) and \( m = 3 \) we list the 8 column vectors as

\[
\begin{align*}
0 & \ 0 & \ 0 & \ 1 & \ 1 & \ 1 & \ 0 & \ 1 \\
0 & \ 0 & \ 0 & \ 1 & \ 1 & \ 1 & \ 0 & \ 1 \\
0 & \ 0 & \ 0 & \ 0 & \ 1 & \ 1 & \ 1 & \ 1 \\
0 & \ 0 & \ 0 & \ 1 & \ 1 & \ 1 & \ 1 & \ 1 \end{align*}
\]

Suppose now that \( \Omega \) is the cyclic group of order \( m \) generated by the permutation \( \phi = (12 \ldots m) \). By letting \( \Omega \) permute the rows of the \( m \times q^m \) array, we induce a permutation group \( \Omega_1 \) on the column vectors of the range \( R \). For example, if \( \phi = (123) \) then the column vectors \( (C_1, \ldots, C_8) \) of (5.2) are permuted to \((C_1, C_3, C_5, C_7, C_2, C_4, C_6, C_8) \). If

\[
A = \begin{bmatrix}
1 & 0 & 1 \\
1 & 1 & 0 \\
0 & 1 & 1
\end{bmatrix} = [C_4, C_7, C_6]
\]

and \( Q \) is the permutation matrix

\[
Q = \begin{bmatrix}
0 & 0 & 1 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{bmatrix}
\]
then

\[
QA = \begin{bmatrix}
0 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 0
\end{bmatrix} = [C_7 C_6 C_4].
\]

By the isomorphism defined in section 1, \(\phi_Q \in S_3\) corresponds to the permutation matrix \(Q\). Thus by applying \(\phi_Q\) to the rows of the \(m \times q^m\) array, we induce a permutation on the column vectors of the range \(R\). This in turn induces a permutation of the rows of \(A\) which is equivalent to just permuting the rows of \(A\) by using the permutation matrix \(Q\). Hence we can permute the rows of any matrix by simply permuting the rows of the \(m \times q^m\) array.

If \(\Omega_1\) is the cyclic group of prime order \(m\) acting on the \(q^m\) column vectors induced by a cyclic group of prime order \(m\) acting on the rows of the \(m \times q^m\) array, it is not difficult to prove that

\[
P_{\Omega_1}(x_1, \ldots, x_m) = \frac{1}{m} (x_1^{q^m} + (m-1)x_1^{q^m} (q^m-q)/m).
\]

We are now ready to prove

**Theorem 5.2.** If \(\Omega_1\) is cyclic of prime order \(m\) and \(\Omega_2\) is cyclic of order \(n\), then if \(m \nmid n\)

\[
\lambda(\Omega_2, \Omega_1) = \frac{1}{mn} \sum_{t \mid n} \phi(t)(q^{mn/t} + (m-1)q^{n/t}),
\]

while if \(m \mid n\)

\[
\lambda(\Omega_2, \Omega_1) = \frac{1}{mn} \sum_{t \mid n} \phi(t)(q^{mn/t} + (m-1)q^{n/t}) + \frac{1}{n} \sum_{t \mid n} \phi(t)q^{mn/t}. \tag{5.5}
\]

**Proof.** We must evaluate (5.1) which becomes for fixed \(t \mid n\)

\[
\frac{\phi(t)}{mn} \frac{d^{n/t}}{d^{n/t}} e^{q(z_1 z_2 + \ldots)} + (m-1)e^{q(z_1 z_2 + \ldots)} e^{-q} (z_m + z_2 + \ldots) \tag{5.6}
\]

\(z_1 = 0\).

If \(t = 1\) (5.6) reduces to \(1/\text{mn}[q^{mn} + (m-1)q^n]\). If \(m \nmid n\) and \(t > 1\) is a divisor of \(n\) we have \(M = \phi(t)/\text{mn}(q^{mn/t} + (m-1)q^{n/t})\) which proves (5.4) upon summing over all \(t \mid n\). If \(m \mid n\) and \(1 < t \neq km\) for some positive integer \(k\) the (5.6) contributes \(M\) as before while if \(1 < t = km\) for some \(k\), (5.6) contributes \((\phi(t)q^{mn/t})/n\) from which (5.5) follows.

As an illustration, suppose \(q = m = n = 2\) so that we are considering the 16
2 x 2 matrices over GF(2). Let \( \Omega_1 \) be the cyclic group of order 2 acting on the
two rows of the 2 x 4 array

\[
\begin{pmatrix}
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1
\end{pmatrix}
\]

and let \( \Omega_2 \) be the cyclic group of order 2 acting on the 2 columns of \( D \). Then
from (5.5) we have \( \lambda(\Omega_2, \Omega_1) = 5 + 2 = 7 \) distinct classes which may also be easily
verified by direct calculation.

6. A FURTHER GENERALIZATION

In this section we consider a further generalization by allowing \( \Omega_1 \) to act
directly on the column vectors of \( R \) rather than on the rows of the \( m \times q \) array.
As before suppose \( \Omega_2 \) acts on the set of \( n \) columns of \( D \). Thus, after a matrix is
permuted by columns, it is then acted upon by a more general permutation of the
column vectors of \( R \) rather than just permuting the rows of the given matrix. For
example, using the example from section 5, suppose \( \Omega_1 \) is the cyclic group of order
8 generated by \( \phi = (12\ldots8) \). Then if \( \phi \) is applied to the matrix

\[
A = \begin{bmatrix}
1 & 0 & 1 \\
1 & 1 & 0 \\
0 & 1 & 1
\end{bmatrix} = [C_4, C_7, C_6]
\]

we obtain the matrix

\[
[C_5, C_8, C_7] = \begin{bmatrix}
0 & 1 & 0 \\
0 & 1 & 1 \\
1 & 1 & 1
\end{bmatrix}
\]

which cannot be obtained from \( A \) by just permuting the rows of \( A \). Hence we have a
more general setting than that considered in section 5 where equivalent matrices
were obtained by simply permuting the rows and columns of the given matrix.

Suppose \( \Omega_1 \) is cyclic of order \( q^m \) acting on the \( q^m \) column vectors of \( R \) while
\( \Omega_2 \) is cyclic of order \( n \) acting on the \( n \) columns of \( D \).

THEOREM 6.1. If \( p \nmid n \)

\[
\lambda(\Omega_2, \Omega_1) = \frac{1}{nq^m} \sum_{t \mid n} \phi(t) q^{mn/t}
\] (6.1)
while if \( p \mid n \)

\[
\lambda(\Omega_2, \Omega_1) = \frac{1}{nq^m} \left[ \sum_{t \mid n} \phi(t)q^{mn/t} + \sum_{t \mid n} \phi(t)(p^i - p^{i-1} + 1)q^{mn/t} \right]. \tag{6.2}
\]

**PROOF.** Since \( q = p^b \) where \( p \) is a prime and \( b \geq 1 \) we have

\[
P_{\Omega_1}(x_1, \ldots, x_m) = \frac{1}{q^m} \sum_{t \mid q^m} \phi(t)x_t^{q^{m/t}}
\]

\[
= \frac{1}{q^m} \left[ \sum_{i=1}^{b} x_1^{p^i} + \sum_{i=1}^{b} (p^i - p^{i-1})x_1^{p^{i-1}} \right].
\]

Substituting \( P_{\Omega_1} \) and \( P_{\Omega_2} \) into (5.1) we obtain for a general term with \( t \) fixed

\[
N = \frac{\phi(t)}{nq^m} \left\{ \frac{z_1^{b^m}}{z_1^{b^m}} + \sum_{i=1}^{b} (p^i - p^{i-1})z_1^{p^{i-1}} \right\}.
\]

If \( t = 1 \), \( N = q^{mn/(nq^m)} \) while if \( t > 1 \) and \( p \nmid n \) then \( t \neq kp^i \) so that

\( N = (1/nq^m)\phi(t)q^{mn/t} \)

from which (6.1) follows after summing over all \( t \mid n \). In the case where \( p \mid n \), if \( t \) is a divisor of \( n \) and \( t \neq kp^i \) for some \( k \) then \( N \) is the same as in the above case. If \( t = kp^i \) then \( N = (1/nq^m)\phi(t)(p^i - p^{i-1} + 1)q^{mn/t} \) so that summing over all \( t \mid n \) yields (6.2).

As an illustration, if \( q = p = m = n = 2 \) then using (6.2) we see that \( \lambda(\Omega_2, \Omega_1) = 3 \) so that the 16 2 x 2 matrices over GF(2) are decomposed into 3 disjoint equivalence classes.

**REFERENCES**


2. Carlitz, L. "Invariant theory of systems of equations in a finite field", J. Analyse Math. 3 (1953/54), 382-413.


