MATRIX POWERS OVER FINITE FIELDS

MARIA T. ACOSTA-DE-OROZCO

Department of Mathematics
Penn State University
Beaver Campus
Monaca, Pennsylvania 15061

and

JAVIER GOMEZ-CALDERON

Department of Mathematics
Penn State University
New Kensington Campus
New Kensington, Pennsylvania 15068

(Received May 23, 1991 and in revised form May 18, 1992)

ABSTRACT. Let GF(q) denote the finite field of order q = p^n with p odd. Let M denote the ring of 2x2 matrices with entries in GF(q). Let n denote a divisor of q - 1 and assume 2 ≤ n and 4 does not divide n. In this paper, we consider the problem of determining the number of n-th roots in M of a matrix B ∈ M. Also, as a related problem, we consider the problem of lifting the solutions of X^2 = B over Galois rings.

KEY WORDS AND PHRASES. Finite fields and matrix powers.

1991 AMS SUBJECT CLASSIFICATION CODES. Primary 15A33.

1. INTRODUCTION.

Let GF(q) denote the finite field of order q = p^n with p odd. Let M denote the ring of 2x2 matrices with entries in GF(q). Let n denote a positive divisor of q - 1. In this paper, we consider the problem of determining the number N = N(n,B) of n-th roots in M of a matrix B ∈ M; i.e., the number of solutions in M of the equation

\[ X^n = B \] (1.1)

Our present work generalizes a recent paper of Donovan [1] in which the quadratic equation \( X^2 = B \) is solved over the ring M.

As a related problem, we also consider the problem of lifting solutions of equation (1.1), for n = 2, over Galois rings. The Galois ring of order \( p^{rm} \), denoted by GR(\( p^r \), m), can be obtained as a Galois extension of \( Z_{p^r} \) of degree m. The reader can find further details about Galois rings in the reference [4].

If B denotes a scalar matrix, a multiple of the identity matrix, then equation (1.1) is called "scalar equation ". Scalar equations have been already studied by Hodges in [2]. In particular, if
n = 2 and B denotes the identity matrix, then the solutions of (1.1) are called "involutory matrices". Involutory matrices over either a finite field or a quotient ring of the rational integers have been extensively researched, with a detailed extension to all finite commutative rings given by McDonald in [5].

2. OVER FINITE FIELDS.

Let GF(q) denote the finite field of order q = p^e with p odd. Let M denote the ring of 2x2 matrices with entries in GF(q) and let GL denote its group of units. For each B in M let S(B) and [B] denote, respectively, the stabilizer and the conjugate class of B defined by

\[ S(B) = \{ A \in GL : AB = BA \} \] (2.1)

and

\[ [B] = \{ ABA^{-1} : A \in GL \}. \] (2.2)

Thus

\[ |[B]| = |GL : S(B)|. \] (2.3)

Now for the purpose of the present work we will need the following stabilizers:

(i) \[ S \left( \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \right) = \text{GL}(q) \]

(ii) \[ S \left( \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right) = \left\{ \begin{pmatrix} x & 0 \\ y & z \end{pmatrix} : x, y \in GF(q), x \neq 0 \right\} \]

(iii) \[ S \left( \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \right) = \left\{ \begin{pmatrix} x & 0 \\ y & z \end{pmatrix} : x, y \in GF(q), xy \neq 0 \right\}, \quad (a - b)ab \neq 0 \]

(iv) \[ S \left( \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix} \right) = \left\{ \begin{pmatrix} x & y \cdot GF(q) \\ z & y \end{pmatrix} : z^2 - ay^2 \neq 0 \right\}, \quad a \neq 0 \]

We now give a series of lemmas from which our main result, Theorem 6, will follow.

**LEMMA 1.** Assume \( T^n = B \) for some \( T \) and some non-scalar \( B \) in \( M \). Then \( S(T) = S(B) \).

**PROOF.** Since \( B \) is non-scalar, the minimal polynomial of \( T \) is a quadratic polynomial \( f_T(x) = x^2 + ax + b \). Therefore, \( B = T^n = dT + eI \) for some constants \( e \) and \( 0 \neq d \) in \( GF(q) \). Thus, \( S(T) = S(B) \).

**LEMMA 2.** If \( n \geq 2 \) then the number of matrices \( T \) in \( M \) so that \( T^n = 0 \) is \( q^2 \).

**PROOF.** \( T^n = 0 \) if and only if the minimal polynomial of \( T \) is either \( x \) or \( x^2 \). Hence, \( T^n = 0 \) if and only if \( T \) is similar to either \( A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \) or \( B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \). Therefore,

\[ |\{ T \in M : T^n = 0 \}| = |[A]| + |[B]| = |GL : S(A)| + |GL : S(B)| = 1 + q(q - 1)(q^2 - 1)/(q^2 - q) = q^2. \]

**LEMMA 3.** Let \( 2 \leq n \) denote a divisor of \( q - 1 \) and assume that 4 does not divide \( n \). For each \( r \) in \( GF(q)^* \) the number of distinct matrices \( T \) in \( M \) such that \( T^n = \text{diag}(r, r) \) is given by
a) \[ n + (q^2 - q)(n-1)n/2 \] if \( r \in GF(q)^n = \{ y^n : y \in GF(q) \} \)

b) \[ (q^2 - q)n/2 \] if \( r \notin GF(q)^n \) but \( r^2 \in GF(q)^n \)

c) \[ 0 \] if \( r^2 \notin GF(q)^n \)

**PROOF.** Let \( w \) denote a primitive element of \( GF(q) \) and write \( r = w^m \) for some integer \( 1 \leq m \leq q-1 \). Then \( T^n = \text{diag}(r,r) \) if and only if the minimal polynomial of \( T \) divides \( f(z) = z^n - w^m \). Now, if \( D = (n,m) \) denotes the greatest common divisor of \( n \) and \( m \), then we obtain

\[
f(z) = \left(\frac{z^n}{D}\right)^D - \left(\frac{w^m}{D}\right)^D
\]

\[
= D^{-1} \sum_{i=0}^{D-1} \left(\frac{z^n}{D} - \frac{w(q-1)i/D}{D} \right)^{D-i}
\]

We also see that \( w^{(q-1)i/D + m/D} \) does not belong to \( GF(q)^s \) for every odd prime factor \( s \) of \( n/D \). Therefore, by [3, ch. VIII, Th. 16], \( h_i(z) \) is irreducible over \( GF(q) \) for all \( i \). Thus, \( n/D = 1 \), \( n/D = 2 \) or there are no matrices \( T \) so that \( T^n = \text{diag}(r,r) \).

**CASE 1:** \( n/D = 1 \). Then \( n \) divides \( m \) and \( T^n = \text{diag}(r,r) \) if and only if the minimal polynomial of \( T \) is either \( z - a \) or \( (z - a)(z - b) \) where \( a \) and \( b \) denote two distinct roots in \( GF(q) \) of the equation \( z^n = r \). Hence, \( T^n = \text{diag}(r,r) \) if and only if \( T \) is similar to either \( A = \text{diag}(a,a) \) or \( B = \text{diag}(a,b) \). Therefore,

\[
\left| \{ T \in M : T^n = \text{diag}(r,r) \} \right| = n \left( \left| [A] \right| + \binom{n}{2} \left| [B] \right| \right)
\]

\[
= n + \binom{n}{2} \frac{q(q-1)(q^2-1)}{(q-1)^2}
\]

\[
= n + (q^2 + q)(n-1)n/2
\]

**CASE 2:** \( n/D = 2 \). Then \( n/2 \) divides \( m \) and \( T^n = \text{diag}(r,r) \) if and only if the minimal polynomial of \( T \) is a quadratic irreducible polynomial of the form \( z^2 - c \) where \( c \) denotes a root of the equation \( z^{n/2} = r \). Thus, \( T^n = \text{diag}(r,r) \) if and only if \( T \) is similar to \( A = \begin{pmatrix} 0 & c \\ 1 & 0 \end{pmatrix} \). Therefore,

\[
\left| \{ T \in M : T^n = \text{diag}(r,r) \} \right| = \frac{q(q-1)(q^2-1)n}{(q^2-1)(2)}
\]

if \( r \notin GF(q)^n \) but \( r^2 \in GF(q)^n \).

**LEMMA 4.** If \( T^n = \text{diag}(h,k) \) with \( h \neq k \), then \( T = \text{diag}(r,s) \) for some \( r \) and \( s \) in \( GF(q) \).

**PROOF.** Let \( f(z) = z^2 + az + b \) denote the minimal polynomial of \( T \). So, \( T^2 = -aT - bI \) and \( cT + eI = \text{diag}(h,k) \) for some \( c \) and \( e \) in \( GF(q) \). Therefore, \( T = \text{diag}(r,s) \) for some \( r \) and \( s \) in \( GF(q) \).

**LEMMA 5.** A non-scalar \( 2 \times 2 \) diagonalizable matrix over \( GF(q) \) is a \( n \)-th power in \( M \) if and only if its eigenvalues, necessarily distinct, are \( n \)-th powers in \( GF(q) \).

**PROOF.** Assume \( T \) to be non-scalar and diagonalizable so that for some matrix \( P \) in \( GL \), \( P T P^{-1} = \text{diag}(h,k) \) where \( h \neq k \) are the eigenvalues of \( T \). If \( h \) and \( k \) are \( n \)-th powers, say \( h = r^n \) and \( k = s^n \), then
Conversely, suppose \( T = N^n \) and \( T \) is diagonalizable. Say \( P^{-1}TP = \text{diag}(h,k) \) where \( h \neq k \) are the eigenvalues of \( T \). Hence
\[
\text{diag}(h,k) = P^{-1}TP = P^{-1}N^nP = (P^{-1}NP)^n.
\]
Therefore, by Lemma 4, \( P^{-1}NP = \text{diag}(r,s) \) with \( r^n = h \) and \( k^n = s \).

**THEOREM 6.** Let \( B \) denote an element of \( M \). Let \( n \) denote a divisor of \( q - 1 \). Assume \( 2 < n \) and \( 4 \) does not divide \( n \). Then \( B \) has

(a) more than \( n^2 \) \( n \)-th roots in \( M \) if and only if \( B = rI \) for some \( r \) in \( GF(q) \) so that \( r^2 \in GF(q)^n \).

(b) exactly \( n^2 \) distinct \( n \)-th roots in \( M \) if and only if \( B \) has unequal nonzero eigenvalues which are \( n \)-th powers in \( GL(q) \).

(c) at most \( n \) distinct roots in \( M \), otherwise.

**PROOF.** If \( B = rI \) for some \( r \) in \( GF(q) \), then, by Lemma 3, \( T \) has

(i) more than \( n^2 \) \( n \)-th roots if and only if \( r^2 \in GF(q)^n \) and

(ii) zero \( n \)-th roots if and only if \( r^2 \notin GF(q)^n \).

We now assume that \( T \) is non-scalar.

**CASE 1:** \( B \) diagonalizable. Then by Lemma 5, \( B \) is a \( n \)-th power in \( M \) if and only if its eigenvalues, necessarily distinct, are \( n \)-th powers in \( GF(q) \). Therefore, \( B \) has exactly

(iii) \( n^2 \) distinct \( n \)-th roots in \( M \) if and only if \( B \) has unequal nonzero eigenvalues which are \( n \)-th powers in \( GF(q) \) and

(iv) zero \( n \)-th roots otherwise.

**CASE 2:** \( B \) non-diagonalizable. Then the minimal polynomials of both \( B \) and \( T \) are either: quadratic irreducible or quadratic perfect square polynomials. We also see that if \( T^n = B \) then the minimal polynomial of \( T \) is a factor of \( \phi_B(z^n) \) where \( \phi_B(z) \) denotes the minimal polynomial of \( B \).

Therefore, there are at most \( n \) possible minimal polynomial \( f_T(z) \). Further, \( (P^{-1}TP)^n = B \) if and only if \( P \in S(B) \). Therefore, since \( [S(B):S(T)] = 1 \) by Lemma 1, \( B \) has at most \( n \) distinct \( n \)-th roots in \( M \).

3. **LIFTING SOLUTIONS.**

Let \( GR(p^r,m) \) denote the Galois ring of order \( p^{rm} \) with \( p \) odd. For purposes of construction and ease of implementation of Galois rings, one can construct \( GR(p^r,m) \) by considering \( (z_p^r)[x]/(f) \) where \( f \) is a monic irreducible polynomial of degree \( m \geq 1 \) over the finite field \( GF(p^m) = GF(q) \) with \( p \) prime. Further details concerning properties of Galois rings can be found in the reference [4].

In this section, we will consider a special case, \( n = 2 \), of lifting solutions over Galois rings. More specifically, we will prove the following

**THEOREM 7.** Let \( M(p^r+1,m) \) denote the ring of all \( 2 \times 2 \) matrices with entries in \( GR(p^r+1,m) \). Let \( A \) denote an element of \( M \). Assume that \( \overline{A} \), the reduction of \( A \) modulo \( p \), is a non-scalar invertible matrix in \( M(p,m) \). Let \( X_o \) be \( \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) \in M(p^r,m) \) denote a solution of \( X^2 = A \) (mod \( p^r \)). Then \( X_o \) can be lifted from \( M(p^r,m) \) to \( M(p^{r+1},m) \) in

(a) a unique way if \( \overline{cd} \neq 0 \).

(b) \( q^m \) different ways if either \( \overline{d} = 0 \) or \( \overline{cd} \neq 0 \) and \( \overline{b} = 0 \).

(c) \( q^2 \) different ways if \( \overline{d} \neq 0 \) and \( \overline{c} = 0 \).

**PROOF.** Let \( X = \left( \begin{array}{cc} x & y \\ z & w \end{array} \right) \) where \( x,y,z \) and \( w \) are elements of the field \( GR(p,m) \) to be specified presently, then
Now, since $X_0^2 = A$ over $GR(p^r, m)$, we can write $X_0^2 = A - C p^r$ for some $2 \times 2$ matrix $C$ over the ring $GR(p,m)$. Hence,

$$ (X_0 + X p^r)^2 = A + (X_0 X + X X_0 - C) p^r \mod p^r + 1 $$

Therefore, $(X_0 + X p^r)^2 = A$ over the ring $GR(p^r+1, m)$, if and only if

$$ X_0 X + X X_0 = C $$

over the field $GR(p,m)$; i.e., if and only if

$$ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} \pmod{p} $$

where $C = \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}$

Hence, we have to count the number of solutions, in $GR(p,m)$, of the linear system

$$ \begin{pmatrix} y \\ z \\ w \\ x \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} \pmod{p} $$

or

$$ \begin{pmatrix} y \\ z \\ w \\ x \end{pmatrix} = \begin{pmatrix} c_1 c \\ c_3 \\ (c_4 - c_1) b c \\ E_2 \end{pmatrix} \pmod{p} $$

where $E_1 = 2(a + d)(ad - bc)$ and $E_2 = c_1 ad + c_1 d^2 - c_2 cd - b c_3 d + c_4 bc - c_1 bc$. So, since $A$ is non-scalar and invertible, $E_1 \neq 0$. Therefore, a straightforward inspection of the above last augmented matrix will complete the proof of the theorem.

**REFERENCES**

1. DONOVAN, T.P., Matrix squares over $GF(q)$, q odd, Congressus Numerantium 66 (1988), 113-122.