Linear permutation polynomials with coefficients in a subfield

by

J. V. Brawley (Clemson, S. C.), L. Carlitz* (Durham, N. C.),
and Theresa Vaughan (Durham, N. C.)

To Professor Carl Ludwig Siegel

1. Introduction. Let GF($q^n$) denote the finite field of order $q^n$, where $q = p^r$ for some $r > 0$ and some prime $p$, and let GF($q^n$) be a subfield of GF($q^n$) so that $q = n^s$ for some integer $s$, $1 \leq s \leq n$. If $f$ is any function from GF($q^n$) to GF($q^n$) it is well-known that $f$ has a unique polynomial representation

$$f(x) = \sum_{i=0}^{q^n-1} a_i x^i,$$

where the coefficients $a_i \in$ GF($q^n$). In case $f$ is a permutation of GF($q^n$) the corresponding polynomial $f(x)$ is called a permutation polynomial. The set of all such permutation polynomials under composition modulo $x^{q^n} - x$ forms a group which is isomorphic to the symmetric group $S_{q^n}$. Those permutation polynomials of the form (1.1) whose coefficients $a_i$ are in GF($q^n$) constitute a subgroup, the structure of which has been determined by Carlitz and Hayes [2] as a semi-direct product of certain symmetric groups and cyclic groups. In this paper we consider an analogous situation for polynomials of the form

$$f(x) = \sum_{i=0}^{n-1} a_i x^{i^n}$$

with coefficients $a_i$ in GF($q^n$). Such polynomials (which represent a subalgebra of the algebra of linear transformations of GF($q^n$) over GF($q$) are generalizations of the Ore polynomials [6], [7], where the coefficients $a_i$ are assumed to lie in GF($q$). The set of all Ore polynomials under the modulo $x^{q^n} - x$ operations of addition and composition of functions, and scalar multiplication by elements of GF($q$), forms a commutative algebra over GF($q$) which is isomorphic to GF($q$)[x]($x^{q^n} - 1$) (see [7]).

* Supported in part by NSF grant GP-17031.
In order to generalize Ore's work, put

\[(1.3) \quad R_m = \{ f(x) = \sum_{i=0}^{n-1} a_i x^i; \; a_i \in \text{GF}(p^m) \}. \]

Then \( R_m \) under the above mentioned operations is an algebra over \( \text{GF}(q) \).

The case \( m = 1 \) is that treated by Ore. In § 2 of the present paper we show that \( R_m \) is isomorphic to the ring of \( n \times n \) matrices over \( \text{GF}(q) \) from which it follows that the group of units of \( R_m \), the so-called Betti-Mathieu group, is isomorphic to the general linear group \( \text{GL}(n, q) \). (See [1], [3].) In § 3 we prove that \( R_m \) is isomorphic to the ring of \( m \times m \) matrices with entries from the residue class ring \( \text{GF}(q)[x]/(x^m-1) \). This includes Ore's result as well as that given in § 2 as special cases. Using this isomorphism it is easy to describe the group of units of \( R_m \) as a direct product of subgroups in contrast to the Carlitz-Hayes result. This description and several interesting combinatorial results are contained in § 4.

2. Preliminaries. The ring of polynomials with coefficients in \( \text{GF}(q) \) will be denoted by \( \text{GF}(q)[x] \). If \( f(x) \in \text{GF}(q)[x] \), the principal ideal generated by \( f(x) \) is denoted by \( (f(x)) \), and the residue class ring consisting of the elements of \( \text{GF}(q)[x] \) reduced modulo \( f(x) \) is written \( \text{GF}(q)[x]/(f(x)) \).

Also if \( S \) is any ring with identity and \( k \) is any positive integer, the ring of \( k \times k \) matrices with elements from \( S \) will be written as \( (S)_k \), and \( \text{GL}(k, S) \) will denote the group of nonsingular \( k \times k \) matrices over \( S \). In case \( S = \text{GF}(q^m) \) the notation \( \text{GL}(k, q^m) \) is used for \( \text{GL}(k, S) \).

Consider the finite field \( \text{GF}(q^m) \) as a vector space of dimension \( n \) over \( \text{GF}(q) \). Let \( L \) be the algebra of linear transformations of \( \text{GF}(q^m) \) over \( \text{GF}(q) \). The set \( R_m \) of all polynomials of the form

\[(2.1) \quad f(x) = \sum_{i=0}^{n-1} a_i x^i \]

with the coefficients \( a_i \in \text{GF}(q^m) \), equipped with the modulo \( x^m - x \) operations of addition and composition of functions and scalar multiplication by elements of \( \text{GF}(q) \), is an algebra over \( \text{GF}(q) \), which is in fact isomorphic to the algebra \( L \).

Theorem 2.1. The algebras \( R_m \) and \( L \) are isomorphic.

Proof. For each \( f(x) \in R_m \) let \( f \) be the function from \( \text{GF}(q^m) \) to \( \text{GF}(q^m) \) defined by substitution and let \( \psi \) denote the mapping that takes \( f(x) \) to \( f \). Then for each \( f(x) \in R_m \) \( \psi(f(x)) \in L \) as \( (f(x)) = \sum_{i=0}^{n-1} a_i x^i \) and \( (\lambda f(x)) = \sum_{i=0}^{n-1} \lambda a_i x^i \) for all integers \( \lambda \geq 0 \) and \( \xi, \eta \in \text{GF}(q) \), \( \lambda \in \text{GF}(q) \), i.e., \( \psi: R_m \rightarrow L \). Moreover, it is immediate that

\[\psi(f(x) + g(x)) = f + g,\]

\[\psi(\lambda f(x)) = \lambda f,\]

and

\[\psi(f(x)) = f \circ g,\]

so that \( \psi \) is an algebra homomorphism. \( \psi \) is one-one by the remark in the first paragraph of § 1, and since \( |R_m| = (q^m)^n = q^{nm} = |L| \), the proof is complete.

It follows immediately that

Corollary 2.2. The algebra \( R_m \) is isomorphic to \( (\text{GF}(q))^n \).

The Betti-Mathieu group is by definition the group of units of \( R_m \); hence we have reproved (see [1], [3]).

Corollary 2.3. The Betti-Mathieu group is isomorphic to \( \text{GL}(n, q) \).

3. The algebra \( R_m \). If \( m = ns \), where \( m \) and \( s \) are positive integers, we define the algebra \( R_m \) to be the set of all polynomials of the form

\[(3.1) \quad f(x) = \sum_{i=0}^{n-1} a_i x^i \]

with coefficients \( a_i \in \text{GF}(q^m) \subseteq \text{GF}(q^s) \), equipped with addition and composition of functions and scalar multiplication by elements of \( \text{GF}(q) \). When \( m = n \), the algebra \( R_n \) is, as we have seen, isomorphic to the algebra of all linear transformations of \( \text{GF}(q)^m \) over \( \text{GF}(q) \). When \( m = 1 \), the algebra \( R_1 \) is the algebra of polynomials studied by Ore in [1], [7] who has shown that \( R_1 \) is isomorphic to the residue class ring \( \text{GF}(q)[x]/(x^{n-1}) \). The next theorem is a generalization of these results.

Theorem 3.1. If \( m = ns \), where \( m \) and \( s \) are positive integers, then the algebra \( R_m \) is isomorphic to the algebra \( \{(\text{GF}(q)[x]/(x^{n-1}))_m\} \) of \( m \times m \) matrices with elements from the residue class ring \( \text{GF}(q)[x]/(x^{n-1}) \).

Proof. For convenience, let \( R_m = \{(\text{GF}(q)[x]/(x^{n-1}))_m\} \). Fix any ordered basis \( B = \{ \beta_1, \beta_2, \ldots, \beta_m \} \) for \( \text{GF}(q^m) \) over \( \text{GF}(q) \). If \( f(x) = \sum_{i=0}^{n-1} a_i x^i \) with coefficients \( a_i \in \text{GF}(q^m) \), let \( [f]_B \) denote the matrix in \( \{(\text{GF}(q)[x]/(x^{n-1}))_m\} \) which represents the linear transformation \( f(x) \) in the ordered basis \( B \).

We first note that any element of \( R_m \), say

\[(3.2) \quad g(x) = \sum_{i=0}^{n-1} a_i x^i; \quad a_i \in \text{GF}(q^m) \]

may be rewritten as

\[(3.3) \quad g(x) = \sum_{i=0}^{n-1} \sum_{k=0}^{m-1} a_{ik} x^{i+k} \]

If we let \( g_i(x) = \sum_{k=0}^{m-1} a_{ik} x^{i+k} \) for \( i = 0, 1, \ldots, s - 1 \), then we may write

\[(3.4) \quad g(x) = \sum_{i=0}^{m-1} g_i(x^{i+s}) \].
On the other hand, any element $F$ of $S_m$ has the form $F = \{f_i(x)\}$ (for $i, j = 0, 1, \ldots, m-1$), where each $f_i(x)$ is a polynomial over $GF(q)$ of degree less than $s$, and we may rewrite the matrix $F$ as follows:

\begin{equation}
F = F_0 + F_1 x + F_2 x^2 + \cdots + F_{s-1} x^{s-1}
\end{equation}

where each $F_k$ ($k = 0, 1, \ldots, s-1$) is an $m \times m$ matrix over $GF(q)$, and the $(i,j)$th entry of $F_k$ is the coefficient of $x^k$ in the polynomial $f_i(x)$.

Now for each $F_k$, there exists a polynomial

\begin{equation}
g_k(x) = \sum_{i=0}^{m-1} b_{ik} x^i
\end{equation}

such that $F_k$ is the matrix representing $g_k(x)$ in the ordered basis $B$, that is, $F_k = [g_k]_B$.

It is now fairly obvious how to define an isomorphism between $S_m$ and $R_m$. If $F$ is given by (3.5) and the corresponding $g_k(x)$ are given by (3.6), define a mapping $\varphi: S_m \to R_m$ by

\begin{equation}
\varphi(F_{ab}) = g_k(x^{ab})
\end{equation}

where it is understood that map $\varphi$ is to be extended linearly to all of $S_m$. It is clear that this is indeed a map from $S_m$ to $R_m$, since every element of $R_m$ may be written in the form (3.4). Evidently addition and scalar multiplication by elements of $GF(q)$ are preserved, and it follows from Theorem 2.1 that $\varphi$ is bijective. It remains only to show that the map $\varphi$ preserves composition. Suppose that $F = [g]_B$ and $H = [h]_B$ are any two matrices in $GF(q)_m$. Then if $i$ and $k$ are positive integers less than $s$, and $i + k = j (\text{mod } s)$, we have

\begin{equation}
\varphi(Gx^i) \circ \varphi(Hx^k) = g(x^{im}) \circ h(x^{km}) = g(h(x^{km})x^{im}) = g(h(x^{km}))
\end{equation}

since the coefficients of $h(x)$ are elements of $GF(q)$. Since $Gx^i Hx^k = Gx^j$, and $\varphi(Gx^i) = g(h(x^{km}))$, it follows that

\begin{equation}
\varphi(Gx^i) \circ \varphi(Hx^k) = \varphi(GHx^j)
\end{equation}

and so $\varphi$ preserves composition. Thus $\varphi$ is an isomorphism. This completes the proof.

4. The group of units of $R_m$ and related results. In order to characterize the group of units of $R_m$ we will use the following known facts.

**Lemma 4.1.** If $S$ is a commutative ring with 1 which has the direct sum decomposition $S = \bigoplus_{i=1}^{t} S_i$, then $(S)_m = \bigoplus_{i=1}^{t} (S_i)_m$ and moreover $GL(m, S) = \bigoplus_{i=1}^{t} GL(m, S_i)$ so that

\begin{equation}
|GL(m, S)| = \prod_{i=1}^{t} |GL(m, S_i)|.
\end{equation}

**Theorem 4.3.** Let $S = GF(q)x_{1}/(P(x)^t)$ where $P(x)$ is an irreducible in $GF(q)[x]$ of degree $d$, then

\begin{equation}
|GL(m, S)| = q^{md}\prod_{i=1}^{\infty} (1 - q^{-di}).
\end{equation}

The proof of Lemma 4.1 is easy. As for Lemma 4.2, one can use the formula of McDonald [5] once it is noted that $S$ is a finite local ring. Basically, the proof uses the correspondence theorem for rings together with the facts that (i) $M = P(x)$ is the unique maximal ideal of $S$, (ii) $M/M = GF(q^d)$ and (iii) $A \in (S)_m$ is nonsingular iff $\mu(A) = \mu(\alpha_0)^{\epsilon}/(S/M)_m$ is nonsingular where $\mu: S \to S/M$ is the natural homomorphism.

**Theorem 4.3.** Let $S = GF(q)x_{1}/(ax_1 - 1)$, and suppose that

\begin{equation}
A = P_1(x)^tP_2(x)^t \cdots P_t(x)^t
\end{equation}

where the $P_j(x)$ are distinct irreducible elements of $GF(q)[x]$, and the degree of $P_j(x)$ is $d_j$ for $j = 1, 2, \ldots, t$. Then $S_j = GF(q[x]/(P_j(x)^t))$. Then $GL(m, S)$ is isomorphic to the direct product of the set $\{GL(m, S_j): j = 1, 2, \ldots, t\}$ and moreover

\begin{equation}
|GL(m, S)| = q^{md}\prod_{i=1}^{\infty} (1 - q^{-di}).
\end{equation}

**Proof.** It is only necessary to note that $S = S_1 \oplus S_2 \oplus \cdots \oplus S_t$. Then by Lemma 4.1, $GL(m, S)$ is isomorphic to the direct product of the set $\{GL(m, S_j): j = 1, 2, \ldots, t\}$. To get the equality (4.4), use Lemma 4.2:

\begin{equation}
|GL(m, S)| = q^{md}\prod_{i=1}^{\infty} (1 - q^{-di}),
\end{equation}

and from Lemma 4.1,

\begin{equation}
|GL(m, S)| = \prod_{j=1}^{t} |GL(m, S_j)| = \prod_{j=1}^{t} q^{d_j}\prod_{i=1}^{\infty} (1 - q^{-d_j})
\end{equation}

**Corollary 4.4.** If $R_m = \{ \sum_{j=1}^{n} a_j x_j: a_j \in GF(q)^m \}$, then the group of units of $R_m$ is isomorphic to the direct product of the set $\{GL(m, S_j): j = 1, 2, \ldots, t\}$ of Theorem 4.3, and the order of the group of units of $R_m$ is given by (4.4).

**Corollary 4.5.** Under the hypothesis of Theorem 4.3, if also $(s, g) = 1$, then the group of units of $R_m$ is isomorphic to a direct product of general linear groups.
Proof. If \((r, q) = 1\), then each exponent \(e_k\) appearing in the factorization (4.3) of \(x^d - 1\) is equal to one, and each \(S_j\) is isomorphic to the field \(\text{GF}(q^d)\). Then \(\text{GL}(m, S_j) = \text{GL}(m, q^d)\) is a general linear group.

It should be noted that (4.4) can be derived directly from the result of Parahat [4] which gives the order of any finite ring \(S\) with 1 in terms of \(\text{rad}S\) and the structure of \(S/\text{rad}S\) as assured by the Wedderburn–Artin Theorem. This involves however computing \(\text{rad}S\) and knowing exactly how \(S/\text{rad}S\) decomposes into a direct sum of matrix rings over finite fields.

As final items we consider several interesting combinatorial questions. Suppose we are given the polynomial

\[
(4.5) \quad f(x) = \sum_{i=0}^{n-1} b_i x^i, \quad b_i \in \text{GF}(q^n),
\]

so that \(f(x)\) acting on \(\text{GF}(q^n)\) as a vector space over \(\text{GF}(q)\) is a linear transformation \(f\). The questions are (i) How many \(\psi(x) \in \mathbb{R}_n\) when acting on \(\text{GF}(q^n)\) equal \(f\) and (ii) How many of these \(\psi(x)\) are in the group of units of \(\mathbb{R}_n\); i.e., are permutations of \(\text{GF}(q^n)\). The answers to these questions are the content of our last theorem.

**Theorem 4.6. The number of polynomials**

\[
(4.6) \quad \psi(x) = \sum_{i=0}^{n-1} a_i x^i, \quad a_i \in \text{GF}(q^n),
\]

whose restriction to \(\text{GF}(q^n)\) defines the same functions as (4.3) is \(q^{n(n-m)}\). Of these, the number which are in the group of units of \(\mathbb{R}_n\) is zero if \(f(x)\) is not one-one on \(\text{GF}(q^n)\) and is \(q^{n(n-m)}|\text{GL}(n-m, q)|\) if \(f(x)\) is one-one on \(\text{GF}(q^n)\), where \(|\text{GL}(t, q)|\) is the well-known number \(\sum_{i=0}^{q-1} (q^t - q^i)\). Thus, in particular, the number of such extensions of \(f(x)\) is independent of the function \(f(x)\).

**Proof.** Any \(\psi(x)\) of the form (4.6) may be rewritten as

\[
\psi(x) = \sum_{i=0}^{n-1} \sum_{j=0}^{s-1} a_{i+j} x^{d+i+j}.
\]

If \(\xi \in \text{GF}(q^n)\) then

\[
\xi^{d+i+j} = \xi^i \xi^{d+j} = \xi^i,
\]

so that

\[
\psi(\xi) = \sum_{i=0}^{n-1} \left( \sum_{j=0}^{s-1} a_{i+j} \right) \xi^i.
\]

Hence

\[
\psi(\xi) = f(\xi) \quad \text{for all } \xi \in \text{GF}(q^n)
\]

if and only if

\[
\sum_{j=0}^{s-1} a_{i+j} = b_i \quad (i = 0, 1, \ldots, m-1).
\]

The number of solutions \((a_0, a_1, \ldots, a_{m-1})\) to this system of linear equations is independent of the particular \(b_i\) and is \(q^{(s-1)m} = q^{(n-m)}\) which completes the first part of the theorem.

As for the second part, clearly if \(f(x)\) is not one-one on \(\text{GF}(q^n)\) none of the \(\psi(x)\) maps on \(\mathbb{R}_n\) whose restriction to \(\text{GF}(q^n)\) equals \(f(x)\) can be one-one; thus, assume \(f(x)\) is one-one on \(\text{GF}(q^n)\). Any linear map is completely determined by its action on a basis. Thus if \(f(x)\) is given linear and one-one on \(\text{GF}(q^n)\), then the number of ways to extend \(f(x)\) to a one-one linear map on \(\text{GF}(q^n)\) is precisely the number of distinct ordered linearly independent sequences of \(n-m\) elements of \(\text{GF}(q^n)\) which are bases for complementary subspaces of \(\text{GF}(q^n)\). By a standard argument, this number is given by

\[
q^n q^{n-1} \cdots q^{n-1} = q^n(q^{n-m} - 1)q^n(q^{n-m} - 2) \cdots q^n(q^{n-m} - q^{n-m-1}) = q^{n(n-m)}\text{GL}(n-m, q)|.
\]

This completes the proof.

**References**


Received on 17.9.1972