Pattern Distributions of Legendre Sequences

Cunsheng Ding

Abstract—Legendre sequences have a number of interesting randomness properties and are closely related with quadratic residue codes. In this correspondence we give lower and upper bounds on the number of patterns distributed in a cycle of the Legendre sequences and establish the relationship between the weight distribution of quadratic residue codes and the pattern distribution of Legendre sequences. Our result shows that Legendre sequences have an ideal distribution of patterns of small length and the pattern distribution of Legendre sequences. We prove that every binary periodic sequence with known pattern distribution gives a cyclic code with known weight distribution; by known pattern distribution we mean that the numbers of all patterns present in a cycle of the sequence are known.

Index Terms—Linear codes, pattern distribution, sequence, weight distribution.

I. INTRODUCTION

Pseudorandom sequences have wide applications in ranging systems, global positioning systems, code-division multiple-access systems, spread-spectrum communication systems, radar systems, and stream ciphers [2]. One randomness aspect of sequences is the pattern distribution. One of the three postulates for the randomness of sequences made by Golomb is about the distribution of some special patterns [4].

Let \( N \) be a prime. The Legendre sequence \( l^\infty \) with respect to the prime \( N \) is defined by

\[
l_i = \left\lfloor \frac{i + \left( \frac{i}{N} \right)}{2} \right\rfloor, \quad \text{if } i \not\equiv 0 \pmod{N}
\]

\[
0, \quad \text{otherwise}
\]

for each \( i \geq 0 \), where \( \left( \frac{i}{N} \right) \) is the Legendre symbol. Here and hereafter we are dealing with the \( \{0, 1\} \) version of Legendre sequences.

Legendre sequences have optimal balance between 1’s and 0’s. It is well known that the periodic autocorrelation function of the Legendre sequence is two-valued if \( N \equiv 3 \pmod{4} \) and three-valued if \( N \equiv 1 \pmod{4} \). They have thus optimal autocorrelation properties. Ding, Helleseth, and Shan [3] proved that the linear span of the Legendre sequence takes on one of \( \left\{ \frac{N-1}{2}, \frac{N+1}{2} \right\} \), depending on the value of \( N \pmod{8} \). In this correspondence, we develop lower and upper bounds on the number of patterns distributed in a cycle of a Legendre sequence, which is much superior to known bounds. Our result shows that Legendre sequences have an ideal distribution of patterns of small length. We also establish the relationship between the weight distribution of quadratic residue (Q.R.) codes and the pattern distribution of Legendre sequences.

II. PATTERN DISTRIBUTIONS

Let \( w^\infty \) be a binary sequence of period \( N \). Here \( N \) is not necessarily the least period. Let \( Z_N = \{0, 1, \ldots, N-1\} \) be the residue ring with respect to integer addition and multiplication modulo \( N \). Define the set

\[
C_i = \{ j \in Z_N: w_j = i \}
\]

for each \( i = 0, 1 \).

Let \( i_0, i_2, \ldots, i_{s-1} \in \{0, 1\} \) and \( r_0, r_1, \ldots, r_{s-1} \) be \( s \) pairwise distinct elements of \( Z_N \). Define

\[
D_{i_0, \ldots, i_{s-1}}(r_0, \ldots, r_{s-1}) = \bigcap_{k=0}^{s-1} \left( C_{i_k} + r_k \right)
\]

\[
d_{i_0, \ldots, i_{s-1}}(r_0, \ldots, r_{s-1}) = \bigcap_{k=0}^{s-1} \left( C_{i_k} + r_k \right)
\]

Then for a fixed \( s \) and a set of fixed \( r_0, \ldots, r_{s-1} \), the set

\[
\{ D_{i_0, \ldots, i_{s-1}}(r_0, \ldots, r_{s-1}): i_0, \ldots, i_{s-1} \in \{0, 1\} \}
\]

forms a partition of \( Z_N \). The above parameters

\[
d_{i_0, \ldots, i_{s-1}}(r_0, \ldots, r_{s-1})
\]

measure the number of patterns distributed in a cycle of the sequence. A pattern is defined to be a string

\[
i_0 \ast \cdots \ast i_1 \ast \cdots \ast i_{s-1}
\]

where \( i_0, \ldots, i_{s-1} \) are fixed bits, those \( \ast \)'s are “do-not-care” bits that could be either 0 and 1, and the distances among \( i_0, i_1, \ldots, i_{s-1} \) are fixed. The distribution of special patterns of maximum-length sequences, i.e., blocks and gaps, was studied by Golomb [4].

In what follows we develop lower and upper bounds on the parameters \( d_{i_0, \ldots, i_{s-1}}(r_0, \ldots, r_{s-1}) \) for Legendre sequences. By the definition of Legendre sequences, \( C_i \) is the set of quadratic residues modulo \( N \), and \( C_0 = Z_N \setminus C_1 \). Thus pattern distributions of Legendre sequences are actually the distribution of patterns of quadratic residues and nonresidues.

For the distribution of patterns of length two in Legendre sequences we have the following exact result, which means that Legendre sequences have the best possible distribution of patterns of length two.
Proposition 1: If \( N \equiv 3 \pmod{4} \), then

\[
d_{ij}(r_0, r_1) = \begin{cases} 
\frac{N - 3}{4}, & \text{for } r = (i, j) = (1, 1) \\
\frac{-1}{2}, & \text{for the other half elements} \\
\frac{N + 3}{4}, & \text{elements } r \text{ of } Z_N \\
\frac{-1}{4}, & \text{for the other half elements} \\
\frac{N - 1}{4}, & \text{elements } r \text{ of } Z_N \\
\frac{N + 3}{4}, & \text{for the other half elements} \\
\frac{-1}{4}, & \text{for the other half elements}
\end{cases}
\]

where \( r = r_1 - r_0 \neq 0 \).

If \( N \equiv 1 \pmod{4} \), the cyclotomic numbers of order two are given by \([11]\)

\[
(0, 0) = \frac{N - 5}{4} \quad (0, 1) = (1, 0) = (1, 1) = \frac{N - 1}{4}.
\]

If \( N \equiv 3 \pmod{4} \), they are given by

\[
(0, 0) = (1, 1) = \frac{N - 3}{4} \quad (0, 1) = \frac{N + 1}{4}.
\]

Note that \(-1 \in D_0\) if and only if \( N \equiv 1 \pmod{4} \) and \( r \in D_j \) if and only if \( r^{-1} \in D_j \). The conclusions of this proposition then follow from the cyclotomic numbers of order 2 and the above four formulae.

It is noted that the above proposition can also be proved with the following Jacobsthal formulae \([1]\):

\[
\sum_{i=0}^{N-1} \frac{ax + b}{N} = N \left( \frac{b}{N} \right) \left[ 1 - \left( \frac{a}{N} \right)^2 \right]
\]

\[
\sum_{i=0}^{N-1} \frac{(ax^2 + bx + c)}{N} = \begin{cases} 
(N - 1) \left( \frac{a}{N} \right) & \text{if } b^2 - 4ac \equiv 0 \pmod{N} \\
- \left( \frac{a}{N} \right) & \text{if } b^2 - 4ac \not\equiv 0 \pmod{N}.
\end{cases}
\]

There have been some papers on the estimation of

\[
d_{i_0 \cdots i_{t-1}}(r_0, \cdots, r_{t-1})
\]

but no tight bound is known. The general lower and upper bound \( N/2^t + t(3 + \sqrt{N}) \) on

\[
d_{i_0 \cdots i_{t-1}}(r_0, \cdots, r_{t-1})
\]

was developed by Peralta \([8]\), which will be referred to as Peralta bounds. Peralta treated zero as a quadratic residue, while we take it as a nonresidue. These two kinds of treatment make little difference. We now derive a better bound on

\[
d_{i_0 \cdots i_{t-1}}(r_0, r_1, \cdots, r_{t-1}).
\]

To this end, we need the calculation and estimation of the character sum

\[
\phi_r(a_1, \cdots, a_s) = \sum_{x=0}^{N-1} \frac{(x + a_1) \cdots (x + a_s)}{N}
\]

where \( a_1, \cdots, a_s \) are pairwise distinct elements of \( Z_N \).

The Peralta bounds are based mainly on the following estimate.

Lemma 1: Let \( r \geq 2 \). Then

\[
|\phi_r(a_1, \cdots, a_s)| < (r - 1)\sqrt{N}
\]

where the \( a_i \)'s are pairwise distinct elements of \( Z_N \).

This result is a special case of a more general result about multiplicative characters due to Weil \([9, \text{Theorem 2C, p. 43}]\). Following Peralta \([8]\), we call this inequality the Weil bound.
$\phi_2(a_1, a_2)$ can be evaluated exactly, since it can be reduced to the calculation of $\phi_1(a)$. This can be generalized into the fact that

$$
\phi_r(a_1, \cdots, a_r)
$$

reduces to

$$
\phi_{r-1}(0, 1, b_1, \cdots, b_{r-3})
$$

if $r$ is even, and to

$$
\phi_r(0, 1, b_1, \cdots, b_{r-2})
$$

if $r$ is odd. The proof of the case $r = 4$ was given by Davenport [1]. We now generalize Davenport’s proof for the case $r = 4$ into the general case of $r$ being even.

**Lemma 2:** Let $r \geq 2$ be even, and let $a_1, \cdots, a_r$ be $r$ pairwise distinct elements of $Z_N$. Then

$$
\phi_r(a_1, a_2, \cdots, a_r) = -1 + \left(\frac{c(a_1, \cdots, a_r)}{N}\right) \phi_{r-1}(d_2, \cdots, d_r)
$$

where these $d_i$ are pairwise distinct and

$$
d_i = \frac{(a_i - a_2)(a_i - a_3)}{(a_i - 1)(a_2 - a_3)} \quad i \geq 2
$$

$$
c(a_1, \cdots, a_r) = (a_1 - a_2)(a_1 - a_3) \prod_{i=2}^{r-1} [(a_i - 1)(a_2 - a_3)].
$$

**Proof:** The key to such a reduction is the use of the transformation

$$
x = \frac{uy + v}{gy + h}
$$

with $uh - gv \neq 0 \pmod{N}$, which gives

$$
\sum_{x \neq \phi_{r-1}} F(x) = \sum_{x \neq \phi_{r-1}} F\left(\frac{uy + v}{gy + h}\right)
$$

where $F(x)$ is a polynomial of $Z_N[x]$. We choose now

$$
u = -a_1(a_2 - a_3) \quad v = -a_2(a_1 - a_3)$$

$$
g = a_2 - a_3 \quad h = a_1 - a_3.
$$

In our case, define

$$
F(x) = \left(\frac{x + a_1}{N}\right) \cdots \left(\frac{x + a_r}{N}\right).
$$

Note that $(x^2/N) = 1$ for all $x \neq 0 \pmod{N}$. We have

$$
\left(\frac{c}{N}\right) \frac{\prod_{i=2}^{r}(N - y + d_i)}{N} = \left(\frac{1}{N}\right) = 1.
$$

Then applying the above transformation yields

$$
\phi_r(a_1, \cdots, a_r)
$$

$$
= \sum_{x \neq \phi_{r-1}} F(x)
$$

$$
= \sum_{x \neq \phi_{r-1}} F\left(\frac{uy + v}{gy + h}\right)
$$

$$
= \sum_{x \neq \phi_{r-1}} \left(\frac{(gy + h)^{r-1} \prod_{i=1}^{r} [(a_i, g + u)g + a_i h + v]}{N}\right)
$$

$$
= \sum_{x \neq \phi_{r-1}} \left(\frac{\prod_{i=1}^{r} [(a_i, g + u)g + a_i h + v]}{N}\right)
$$

$$
= \left(\frac{c}{N}\right) \sum_{x \neq \phi_{r-1}} \left(\frac{\prod_{i=2}^{r} (y + d_i)}{N}\right)
$$

$$
= \left(\frac{c}{N}\right) \frac{\prod_{i=2}^{r} (y + d_i)}{N} - 1
$$

$$
= -1 + \left(\frac{c}{N}\right) \phi_{r-1}(d_2, \cdots, d_r).
$$

It can be easily proven that $d_2, \cdots, d_r$ are pairwise distinct.

We will refer to this result as the Davenport reduction theorem, which will play an important role in developing the new bounds.

The following combinatorial results are needed in the sequel. Their proofs are standard and can be found in many textbooks on combinatorics.

**Lemma 3:** Let $n \geq 2$. Then

1) $\sum_{i=0}^{n} \binom{n}{i} = \sum_{i=0}^{n} \binom{n}{i} = 2^{n-1}$.

2) $\sum_{i=1}^{n} \binom{n}{i} (i - 1) = 2^{n-1}(n - 2) + 1$.

In the sequel we assume $s \geq 3$. To derive the bounds, we make use of the function

$$
G(x) = \frac{1}{2^{r}} \prod_{k=0}^{r-1} \left[ 1 + (-1)^{r+1} \left(\frac{x - T_k}{N} \right) \right].
$$

The following lemma is needed later.

**Lemma 4:** Define

$$
B(t, s, i_0, \cdots, i_{s-1}) = \sum_{0 \leq k_1 < \cdots < k_l \leq s-1} (-1)^l \sum_{j=1}^{l} \binom{l}{i_{k_j+1}}
$$

where $t$ and $s$ are positive integers with $t \leq s$ and $i_0, \cdots, i_{s-1} \in \{0, 1\}$. Then

$$
|B(t, s, i_0, \cdots, i_{s-1})| \leq \binom{s}{t}.
$$

Note that

$$
\sum_{x} \left(\frac{x}{N}\right) = 0.
$$

The following conclusion is obvious.
Lemma 5:
\[
\sum_{j=0}^{N-1} \sum_{t=0}^{s-1} (-1)^t \left( x - \frac{r_k}{N} \right) = 0.
\]

As we assume that \( r_0, r_1, \ldots, r_{s-1} \) are pairwise distinct, the following conclusion follows from the second Jacobsthal formula described earlier.

Lemma 6:
\[
\sum_{j=0}^{N-1} \sum_{0 \leq k_1 \cdots k_s \leq s-1} (-1)^{k_1 + \cdots + k_s} \left( \frac{x - r_{k_1}}{N} \right) = -B(2, s, i_0, \ldots, i_{s-1}).
\]

The following lemma derives directly from the Davenport reduction theorem.

Lemma 7: Let \( s' \) be even. Then
\[
\sum_{t \text{ even}}^{s'} \sum_{0 \leq k_1 < \cdots < k_t \leq s-1} (-1)^{k_1 + \cdots + k_t} \left( \frac{c_{k_1} \cdots c_{k_t}}{N} \right) \phi_t( -r_{k_1} \cdots -r_{k_t} )
\]
\[
= \sum_{t \text{ even}}^{s'} \sum_{0 \leq k_1 < \cdots < k_t \leq s-1} (-1)^{k_1 + \cdots + k_t} \left( \frac{c_{k_1} \cdots c_{k_t}}{N} \right) \phi_{t-1}( -r_{k_1} \cdots -r_{k_{t-1}} )
\]
\[
- \sum_{t \text{ even}}^{s'} B(t, s, i_0, \ldots, i_{s-1})
\]

Applying Lemma 7 and the Weil bound to the formula above yields
\[
2^s \sum_{x} G(x) = N + \sum_{t \text{ even}}^{s'} \sum_{0 \leq k_1 < \cdots < k_t \leq s-1} (-1)^{k_1 + \cdots + k_t} \phi_t( -r_{k_1} \cdots -r_{k_t} )
\]
\[
= N - B(2, s, i_0, \ldots, i_{s-1})
\]

By Lemma 4
\[
\sum_{t \text{ even}}^{s'} B(t, s, i_0, \ldots, i_{s-1}) \leq -1 + \sum_{0 \leq t \leq s} \binom{s}{t}
\]
\[
= 2^{s+1} - 1.
\]

Combining the above two formulas yields
\[
\frac{N}{2^s} - T \leq \sum_{x} G(x) \leq \frac{N}{2^s} + T
\]
(2)

where
\[
T = \sqrt{\frac{N(2^{s-1}(s-3) + 2) + 2^{s-1} - 1}{2^s}}.
\]

Now let \( WH(i_0 \cdots i_{s-1}) \) denote the Hamming weight of vector \( (i_0 \cdots i_{s-1}) \), i.e., the number of 1’s in the binary vector. Assume that \( WH(i_0 \cdots i_{s-1}) = s - t \). Note that \( r_j \in C_{i_j} + r_j \) if and only if \( i_j = 0 \), so there are at most \( t \) elements in \( \{ r_0, r_1, \ldots, r_{s-1} \} \) which belong to the set \( D_{i_0 \cdots i_{s-1}}(r_0, \ldots, r_{s-1}) \), where \( t \leq s \). We now assume that there are exactly \( u \) such elements, say \( r_0, r_1, \ldots, r_{u-1} \), where \( u \leq t \). Let
\[
B = \{ r_0, r_1, \ldots, r_u \}
\]

and
\[
A = D_{i_0 \cdots i_{s-1}}(r_0, \ldots, r_{s-1}) \setminus B.
\]

Then we have
\[
\sum_{x \in A} G(x) = d_{i_0 \cdots i_{s-1}}(r_0, \ldots, r_{s-1}) - u
\]
\[
\sum_{x \in B} G(x) = u/2.
\]

It follows that
\[
d_{i_0 \cdots i_{s-1}}(r_0, \ldots, r_{s-1}) - \frac{t}{2} \leq \sum_{x \in D_{i_0 \cdots i_{s-1}}} G(x)
\]
\[
d_{i_0 \cdots i_{s-1}}(r_0, \ldots, r_{s-1}) \geq \sum_{x \in D_{i_0 \cdots i_{s-1}}} G(x)
\]
where \( D_{i_0 \ldots i_s} \) denotes \( D_{i_0 \ldots i_{s-1}}(r_0, \ldots, r_{s-1}) \). Hence, for each \((i_0 \cdots i_{s-1})\) we have

\[
d_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) = \frac{s}{2} \leq \sum_{x \in D_{i_0 \cdots i_{s-1}}} G(x)
\]

\[
d_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) \geq \sum_{x \notin D_{i_0 \cdots i_{s-1}}} G(x).
\]

If the Hamming distance between the vectors \((j_0 \cdots j_{s-1})\) and \((i_0 \cdots i_{s-1})\) is greater than or equal to two, it is not difficult to see that

\[
\sum_{x \in D_{i_0 \cdots j_{s-1}}} G(x) = 0.
\]

Furthermore, if the Hamming distance between the two vectors \((j_0 \cdots j_{s-1})\) and \((i_0 \cdots i_{s-1})\) is one, it is clear that

\[
0 \leq \sum_{x \in D_{i_0 \cdots j_{s-1}}} G(x) \leq \frac{1}{2}.
\]

Note that

\[
\bigcup_{i_0, \cdots, i_{s-1} \in \{0, 1\}} D_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) = Z_N.
\]

It follows that

\[
\begin{cases}
  d_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) - \frac{s}{2} \leq \sum_{x \in \mathbb{Z}_N} G(x) \\
  d_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) + \frac{s}{2} \geq \sum_{x \notin \mathbb{Z}_N} G(x).
\end{cases}
\]

Combining (2) and (4) proves the following result.

**Proposition 2:** Let the symbols and assumptions be as before and \( s \geq 3 \). We have

\[
-W \leq d_{i_0 \cdots i_{s-1}}(r_0, r_1, \cdots, r_{s-1}) - \frac{N}{2s} \leq W
\]

where \( r_0, r_1, \cdots, r_{s-1} \) are pairwise distinct elements of \( Z_N \) and

\[
W = \sqrt{N(2^{s-1}(s-3) + 2) + 2^{s-1}(s+1) - 1}.
\]

Note that \( W \) is essentially \( \sqrt{N\left[(s - 3)/2\right] + (s + 1)/2} \). Our new bounds are much superior to the Peralta bounds. The main contribution to the new bounds comes from the Davenport reduction theorem and several techniques.

For the case \( s = 3 \), the above proof shows that

\[
\begin{cases}
  d_{i_0 i_1}(r_0, r_1, r_2) - \frac{N}{2s} \leq \frac{2\sqrt{N + 15}}{2^3}.
\end{cases}
\]

Numerical computation shows that these lower and upper bound for \( s = 3 \) are quite tight.

It can be seen from the development of the bounds that the new bounds are usually tight for small \( s \). Proposition 2 shows that Legendre sequences have an ideal distribution of patterns of length \( s \) when \( s \) is small.

In summary, Legendre sequences have optimal balance between 0’s and 1’s, a large linear span, the optimal autocorrelation property, and an ideal distribution of patterns of small length.

**III. PATTERN DISTRIBUTION AND WEIGHT DISTRIBUTION**

In this section we show a close relation between the pattern distribution of Legendre sequences and the weight distribution of Q.R. codes. This is also true for some other sequences and codes. Thus the importance of the pattern distribution of sequences in coding theory follows.

Let \( N \equiv \pm 1 \pmod{8} \) and let \( Q \) and \( P \) denote the set of quadratic residues and that of quadratic nonresidues modulo \( N \), respectively. The binary Q.R. codes are those with idempotents

\[
\sum_Q x^i \sum_P x^i 1 + \sum_Q x^i 1 + \sum_P x^i.
\]

Note that the idempotent \( \sum_Q x^i \) above is the polynomial form of the Legendre sequence defined in this correspondence. In the sequel we use \( C \) to denote the binary Q.R. code with this idempotent.

The linear span of a sequence is defined to be the length of the shortest linear-feedback shift-register that produces the sequence [2], [6], [7]. It is known that the Legendre sequence \( L^w \) has linear span \( (N+1)/2 \) if \( N \equiv -1 \pmod{8} \) and \( (N-1)/2 \) if \( N \equiv 1 \pmod{8} \) [3].

Let \( i_0, i_1, \cdots, i_{N-1} \) be the first periodic segment of the Legendre sequence, and let \( L_i \) denote the \( i \)-th left-shifted version of the vector \( (i_0, i_1, \cdots, i_{N-1}) \) for \( i = 0, 1, \cdots, N-1 \). Since \( L^w \) has linear span \( h \), by the definition of linear span the vectors \( L_0, \cdots, L_{h-1} \) of GF(2)\(^N\) form a basis of the Q.R. code \( C \), where \( h \) is either \( (N-1)/2 \) or \( (N+1)/2 \). Thus we have reached the following conclusion.

**Proposition 3:** The Q.R. code \( C \) has generator matrix

\[
G = \begin{bmatrix}
  L_0 \\
  L_1 \\
  \vdots \\
  L_{h-1}
\end{bmatrix}
\]

where \( L_i \) and \( h \) are the same as before.

This result, though straightforward, makes a connection between the Q.R. code and the Legendre sequence. It allows us to establish the relationship between the weight distribution of Q.R. codes and the pattern distribution of Legendre sequences.

Let \( h \) denote the linear span of the Legendre sequence. By Proposition 3 every codeword \( e \) of \( C \) is expressed as

\[
e = L_{r_0} + L_{r_1} + \cdots + L_{r_{s-1}}
\]

where \( 0 \leq r_0 < \cdots < r_{s-1} \leq h - 1 \). The Hamming weight of \( e \) and the pattern distribution of the Legendre sequence are related as follows.

**Proposition 4:**

\[
\text{WH}(e) = \sum_{i_0 + \cdots + i_{s-1} = 1} d_{i_0 \cdots i_{s-1}}(-r_0, -r_1, \cdots, -r_{s-1})
\]

where the addition in \( i_0 + \cdots + i_{s-1} = 1 \) is integer addition modulo 2 and \(-r_i = N - r_i\).

**Proof:** Let \( r_0, \cdots, r_{s-1} \) be fixed and the same as before. Consider the set of (1). It is easy to see that

\[
D_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) \cap D_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) = \emptyset
\]
where \((j_0, j_1, \ldots, j_{s-1}) \in \mathbb{GF}(2)^s\) and \((i_0, i_1, \ldots, i_{s-1})\) are different binary vectors. Then by (3) and definition

\[
\text{WH}(e) = \left\{ i \in \mathbb{Z}_n : \sum_{j=0}^{s-1} f_{j_0 + \cdots + j_{s-1}}(i) = 1 \right\}
\]

\[
= \bigcup_{i_0 + \cdots + i_{s-1} = 1} D_{i_0 \cdots i_{s-1}} \left( -r_0, \ldots, -r_{s-1} \right)
\]

\[
= \sum_{i_0 + \cdots + i_{s-1} = 1} d_{i_0 \cdots i_{s-1}} \left( -r_0, \ldots, -r_{s-1} \right).
\]

Note that these \(d_{i_0 \cdots i_{s-1}}(r_0, \ldots, r_{s-1})\) are actually the number of patterns that appear in a cycle of the Legendre sequence. By Proposition 4, the weight distribution of the Q.R. code is determined if the pattern distribution of the Legendre sequence is determined. However, the converse may not be true. In fact, (5) indicates that the determination of the pattern distribution of the Legendre sequence is much harder than the weight distribution problem of the Q.R. code.

Note that \(d_{i_0 i_1}(r_0, r_1)\) are given by Proposition 1. With the help of Proposition 4 these \(d_{i_0 i_1}(r_0, r_1)\) can be used to determine some weights of the Q.R. code \(C\). For \(s \geq 3\) there are no known formulas for \(d_{i_0 \cdots i_{s-1}}(r_0, \ldots, r_{s-1})\), though many well-known mathematicians have attacked this problem for hundreds of years.

While calculating \(d_{i_0 \cdots i_{s-1}}(r_0, \ldots, r_{s-1})\) is hard, we may hope to develop tight bounds on them. Such tight bounds can be used to say something about the weight distribution of Q.R. codes. To illustrate this, we take our bounds of Proposition 2 as an example.

**Proposition 5:** The Q.R. code \(C\) has at least \((\binom{s}{h})\) codewords \(e\) with

\[
\frac{N}{2} - 2s^{-1}W \leq \text{WH}(e) \leq \frac{N}{2} + 2s^{-1}W
\]

where \(W\) and \(h\) are the same as before.

**Proof:** There are \((\binom{s}{h})\) sets of integers

\[
0 \leq r_0 < \cdots < r_{s-1} \leq h - 1.
\]

Each set of such integers defines a codeword

\[
e = L_{r_0} + \cdots + L_{r_{s-1}}
\]

where \(L_i\) are the same as before, and different sets define different codewords by the definition of linear span. Note that \(x_0 + x_1 + \cdots + x_{s-1} = 1\) over \(\mathbb{GF}(2)[x_0, \ldots, x_{s-1}]\) has \(2s^{-1}\) solutions \((x_0, \ldots, x_{s-1})\). The conclusion then follows from Propositions 2 and 4.

The bounds of Proposition 5 are useful only when \(s\) is small. When \(s = 3\), it says that \(C\) has at least \((\binom{3}{2})\) codewords \(e\) with

\[
\frac{N - 2\sqrt{N} - 15}{2} \leq \text{WH}(e) \leq \frac{N + 2\sqrt{N} + 15}{2}.
\]

If we had better bounds

\[
\delta_2 \leq d_{i_0 \cdots i_{s-1}}(r_0, \ldots, r_{s-1}) \leq \delta_1
\]

the proof of Proposition 5 would also prove that there are at least \((\binom{s}{2})\) codewords \(e\) with

\[
2s^{-1}\delta_2 \leq \text{WH}(e) \leq 2s^{-1}\delta_1.
\]

Therefore, any improvement on the bounds on

\[
d_{i_0 \cdots i_{s-1}}(r_0, \ldots, r_{s-1})
\]

would enhance our knowledge about Q.R. codes.

**Remark:** Note that the above observation also applies to some other binary cyclic codes and binary sequences. Let \(w^n\) be a binary sequence of period \(n\) (not necessarily the least period). Let \(h\) denote the linear span of this sequence. Then all the shifted versions of the first periodic segment of this sequence generate an \([n, h]\) cyclic code. If the pattern distribution of this sequence is known, the weight distribution of the cyclic code is known, as Proposition 4 is also true in this general case. Hence, every binary periodic sequence with known pattern distributions gives a binary cyclic code with known weight distribution.

Helleseth has expressed the weights of codewords of Q.R. codes in terms of Legendre sums [5], while we do it in terms of the numbers of patterns of certain lengths present in a cycle of the Legendre sequence. It is possible to bridge the two approaches to get some interesting result.

**IV. CONCLUSIONS**

We have developed new bounds on the numbers of patterns appearing in a cycle of Legendre sequences, which is much superior to known bounds. We have also showed an application of such bounds in Q.R. codes. We have expressed the weights of codewords of a class of binary cyclic codes in terms of the numbers of patterns of certain lengths present in a cycle of the sequence. Thus the pattern distribution problem is also interesting in coding theory.

**ACKNOWLEDGMENT**

The author wishes to thank the two anonymous referees for pointing out some errors in the original version of this correspondence and making a number of suggestions and comments that greatly improved the presentation.

**REFERENCES**


