General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)
ON THE BINARY WEIGHT DISTRIBUTION OF SOME REED-SOLOMON CODES

Technical Report
to

NASA
Goddard Space Flight Center
Greenbelt, Maryland

Grant Number NAG 5-407

Shu Lin
Principal Investigator
Department of Electrical Engineering
University of Hawaii at Manoa
Honolulu, Hawaii 96822

February 22, 1985
ON THE BINARY WEIGHT DISTRIBUTION OF SOME REED-SOLOMON CODES

Tadao Kasami
Osaka University
Toyonaka, Osaka 560, Japan

Shu Lin
University of Hawaii at Manoa
Honolulu, Hawaii 96822, U.S.A.

ABSTRACT

Consider an (n,k) linear code with symbols from GF(2^m). If each code symbol is represented by a m-tuple over GF(2) using certain basis for GF(2^m), we obtain a binary (nm,km) linear code. In this paper, we investigate the weight distribution of a binary linear code obtained in this manner. Weight enumerators for binary linear codes obtained from Reed-Solomon codes over GF(2^m) generated by polynomials, (X-α), (X-1)(X-α), (X-α)(X-α^2) and (X-1)(X-α)(X-α^2) and their extended codes are presented, where α is a primitive element of GF(2^m). Binary codes derived from Reed-Solomon codes are often used for correcting multiple bursts of errors.
1. Introduction

Let \( \{B_1, B_2, \ldots, B_m\} \) be a basis of the Galois field \( GF(2^m) \). Then each element \( z \) in \( GF(2^m) \) can be expressed as a linear sum of \( B_1, B_2, \ldots, B_m \) as follows:

\[
Z = c_1 B_1 + c_2 B_2 + \ldots + c_m B_m,
\]

where \( c_i \in GF(2) \) for \( 1 \leq i \leq m \). There is a one-to-one correspondence between the element \( z \) and the \( m \)-tuple \( (c_1, c_2, \ldots, c_m) \) over \( GF(2) \). Thus \( z \) can be represented by the \( m \)-tuple \( (c_1, c_2, \ldots, c_m) \) over \( GF(2) \).

Let \( C \) be an \((n,k)\) linear block code with symbols from the Galois field \( GF(2^m) \). If each code symbol of \( C \) is represented by a \( m \)-tuple over the binary field \( GF(2) \) using the basis \( \{B_1, B_2, \ldots, B_m\} \) for \( GF(2^m) \), we obtain a binary \((mn, mk)\) linear block code \( C^b \). If code \( C \) is capable of correcting \( t \) or fewer random symbol errors, then \( C^b \) is capable of correcting any combination of

\[
\lambda = \frac{t}{1 + \left(\frac{2}{m-1}\right)}
\]

or fewer bursts of errors of length \( \ell \) [1].

In this paper, we investigate the weight distributions of binary codes derived from codes with symbols from \( GF(2^m) \). Weight enumerators for binary codes obtained from Reed-Solomon codes over \( GF(2^m) \) generated by polynomials, \((X-\alpha), (X-1)(X-\alpha), (X-\alpha)(X-\alpha^2)\) and \((X-1)(X-\alpha)(X-\alpha^2)\) and their extended codes are presented, where \( \alpha \) is a primitive element of \( GF(2^m) \).

2. Binary Weight Distributions of Linear Block Codes over \( GF(2^m) \)

Let \( C \) be an \((n,k)\) linear code with symbols from \( GF(2^m) \). Let \( C^b \) denote the binary \((nm, km)\) linear code obtained from \( C \) by representing each code symbol by a \( m \)-tuple over \( GF(2) \) using the basis \( \{B_1, B_2, \ldots, B_m\} \) for \( GF(2^m) \). Let \( H \) be an \((n-k) \times n\) parity-check matrix of \( C \). By rearranging the
bit positions, a parity-check matrix for the binary code $C^b$ can be represented in the following form:

$$H^b = [\beta_1 H: \beta_2 H: \ldots : \beta_m H]$$  \hspace{1cm} \text{(1)}$$

which is an $(n-k) \times mn$ matrix over $GF(2^m)$. For convenience, we will use the order of bit positions given by (1). Let $\bar{v} = (\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m)$ be a binary vector of $mn$ components, where $\bar{v}_i = (v_{i1}, v_{i2}, \ldots, v_{in})$ is a binary $n$-tuple for $1 \leq i \leq m$. Then, $\bar{v}$ is a codeword in $C^b$ if and only if

$$\sum_{i=1}^{m} \beta_i H^T v_i = 0 \hspace{1cm} \text{(2)}$$

Let $C^\perp$ denote the dual code of $C$. We assume that $C^\perp$ does not contain the all-one vector $(1,1,\ldots,1)$. Let $C_e$ denote the linear code over $GF(2^m)$ whose parity-check matrix is of the following form:

$$H_e = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ H & & & \end{bmatrix}$$  \hspace{1cm} \text{(3)}$$

Clearly $C_e$ is a subcode of $C$. Let $C_b$ and $C_{e,b}$ denote the binary subfield subcodes of $C$ and $C_e$ respectively. Then $C_{e,b}$ is the even-weight subcode of $C_b$.

Let $A_i(X) = A_0 + A_1 X + A_2 X^2 + \ldots + A_n X^n$ be the weight enumerator of $C_b$. Then, $A_{0,i}$ is the number of codewords of weight $i$ in $C_b$. Note that $A_{00} = 1$. Assume that there are $\ell$ types of cosets modulo $C_b$ including $C_b$ itself, and cosets of type-$j$ have the same weight enumerator $A_j(X)$ for $0 \leq j \leq \ell$. Let $\gamma$ be a $(n-k)$-tuple over $GF(2^m)$. Then $\gamma$ is said of "type-$j$" if and only if $\gamma$ is the syndrome of a coset of type-$j$. Since $C_{e,b}$ is the even-weight subcode of $C_b$, each coset of $C_b$ can be partitioned into two cosets of $C_{e,b}$, an even-weight coset and an odd-weight coset. Hence there are $2\ell$
types of cosets modulo $C_{e,b}$. Let $A_{j,e}(X)$ and $A_{j,o}(X)$ denote the even part and odd part of $A_j(X)$ respectively, for $0 \leq j \leq \ell$.

For nonnegative integers $s_1, s_2, \ldots, s_{\ell-1}$ such that $\sum_{j=1}^{\ell-1} s_j \leq m$, let $N_{s_1, s_2, \ldots, s_{\ell-1}}$ denote the number of $(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_m)$'s such that

(i) $\bar{y}_i$ is an $(n-k)$-tuple over $\mathbb{F}_2^m$ for $1 \leq i \leq m$;

(ii) the number of components $\bar{y}_i$ of type $j$ is $s_j$ for $1 \leq j \leq \ell$; and

(iii) the following equality holds

$$\sum_{i=1}^{m} \beta_i \bar{y}_i = 0.$$  \hspace{1cm} (4)

Then, it follows from (2), (4) and the definition of $N_{s_1, s_2, \ldots, s_{\ell-1}}$ that we have Theorem 1.

**Theorem 1:** The weight enumerator of $C_{b}$, denoted $A_{b}(X)$, is given by

$$A_{b}(X) = \sum_{s_1, s_2, \ldots, s_{\ell-1}} N_{s_1, s_2, \ldots, s_{\ell-1}} [A_0(X)]^{m-\lambda} \sum_{j=1}^{\ell-1} s_j \cdot A_j(X),$$  \hspace{1cm} (5)

where $S_{\ell,m} = \{(s_1, s_2, \ldots, s_{\ell-1}) : s_j \geq 0 (1 \leq j \leq \ell)$ and $\sum_{j=1}^{\ell-1} s_j = m \}$ and $\lambda = \sum_{j=1}^{\ell-1} s_j$.

Let $\bar{v} = (\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m)$ be a binary vector of mn components where $\bar{v} = (v_{i1}, v_{i2}, \ldots, v_{in})$ is a binary $n$-tuple for $1 \leq i \leq m$. Let $C_{e}$ be the binary code of length $mn$ derived from $C_{e}$ by representing each code symbol of $C_{e}$ by a binary $m$-tuple using the basis $\{\beta_1, \beta_2, \ldots, \beta_m\}$. Then $\bar{v}$ is a codeword in $C_{b}$ if and only if

$$\sum_{i=1}^{m} \beta_i \cdot \sum_{j=1}^{n} v_{ij} = 0,$$  \hspace{1cm} (7)

$$\sum_{i=1}^{m} \beta_i \cdot v^T = 0.$$  \hspace{1cm} (8)

Since $\beta_1, \beta_2, \ldots, \beta_m$ are linearly independent over $\mathbb{F}_2$, we have that

$$\sum_{j=1}^{n} v_{ij} = 0, \text{ for } 1 \leq i \leq m.$$  \hspace{1cm} (9)

Hence we have Theorem 2.
Theorem 2: The binary code $C^b_e$ is an even-weight code and its weight enumerator $A^b_e(X)$ is given by

$$A^b_e(X) = \sum_{S_{k,m}} N_{s_1, s_2, \ldots, s_{k-1}} [A_0^e(X)]^{m-\lambda} \prod_{j=1}^{k-1} A_{j, e}^b(X),$$

(10)

where $S_{k,m} = \{(s_1, s_2, \ldots, s_{k-1}) : s_j \geq n$ for $1 \leq j \leq k$ and $\sum_{j=1}^{k-1} s_j < m\}$ and $\lambda = \sum_{j=1}^{k-1} s_j$.

Let $C^e_{ex}$ denote the extended code obtained from $C$ by adding an overall parity-check symbol. Hence $C^e_{ex}$ is a code of length $n+1$ with symbols from $GF(2^m)$ and parity-check matrix

$$H^e_{ex} = \begin{bmatrix} 1 & 1 & \ldots & 1 & 1 & 0 \\ & & & & & \vdots \\ & & & & & 0 \\ & & & & & \end{bmatrix}$$

(11)

Let $C^e_{ex, b}$ be the subfield subcode of $C^e_{ex}$. Then $C^e_{ex, b}$ is the extended code of $C^b_b$. It follows from Theorem 1 that we have Theorem 3.

Theorem 3: The weight enumerator $A^e_{ex}(X)$ of $C^e_{ex}$ is given by

$$A^e_{ex}(X) = \sum_{S_{k,m}} N_{s_1, s_2, \ldots, s_{k-1}} [A_0^e(X)]^{m-\lambda} \prod_{j=1}^{k-1} A_{j, e}^e(X),$$

(12)

where $S_{k,m} = \{(s_1, s_2, \ldots, s_{k-1}) : s_j \geq 0$ for $1 \leq j \leq k$ and $\sum_{j=1}^{k-1} s_j < m\}$, $\lambda = \sum_{j=1}^{k-1} s_j$, and

$$A_{j, e}^e(X) = A_{j, e}^e(X) + XA_{j, e}^e(X)$$

(13)

for $0 \leq j < k$.

From Theorems 1, 2 and 3, we see that, if we know the weight enumerators of cosets of the binary subfield subcode $C^b_b$ and coefficients $N_{s_1, s_2, \ldots, s_{k-1}'}$, we can obtain the binary weight enumerators $A^b_e(X), A^b_e(X)$ and $A^e_{ex}(X)$. Weight enumerators of cosets for some classes of codes are known, e.g., the Hamming codes [2]. Let $A^e_H(X)$ denote the weight enumerator of a Hamming code which is known [1-4]. Let $C^b_{b}$ be a Hamming code of length $n=2^m-1$. Then the weight enumerator $A^e_{CH} (C^b_{CH})$ of a coset of $C^b_b$ (other than $C^b_b$) is given by

$$A^e_{CH}(X) = \frac{1}{n} ((X+1)^n - A^e_H(X)).$$

(14)
If $C_b$ has minimum weight at least $2t+1$ and all cosets of $C_b$ with minimum weight $t$ have the same weight enumerator $A_t(X)$, then it follows from MacWilliams equation [2,5] that

$$A_t(X) = \binom{n}{t}^{-1} 2^{-(n-k)} \sum_{j=0}^{n} A_j^t P_t(j) (1+X)^{n-j} (1-X)^j,$$

(15)

where $A_j^t$ is the number of codewords of weight $j$ in the dual of $C_b$ and $P_t(j)$ is a Krawtchouk polynomial. Theorem 4 provides a sufficient condition for all cosets with the same minimum weight to have the same weight enumerator.

**Theorem 4:** If $C_b$ has minimum weight at least $2t+1$ and the number of non-zero weight $w$'s such that there exists a codeword of weight $w$ in the dual code of $C_b$ is not greater than $t+1$, then the minimum weight of a coset other than $C_b$ is at most $t$ and all cosets of $C_b$ with the same minimum weight have the same weight enumerator.

**Proof:** In a coset of $C_b$, there is at most one vector whose weight is not greater than $t$. Hence this theorem follows immediately from Theorem 20 in [p. 169;2].

For example, the condition of Theorem 4 holds for primitive BCH codes of minimum distance 5 and code length $2^m-1$ with odd $m \geq 3$.

### 3. Binary Weight Enumerators for Some Reed-Solomon Codes

In this section we will derive the weight enumerators for the binary codes obtained from some Reed-Solomon codes with symbols from $GF(2^m)$. Let $C$ be a Reed-Solomon code of length $n=2^m-1$ with generator polynomial $\bar{g}(X)$. Let $\alpha$ be a primitive element of $GF(2^m)$.

**Case 1:** $\bar{g}(X) = X-\alpha$.

In this case, the parity-check matrix for $C$ is

$$H = [1 \ \alpha \ \alpha^2 \ldots \alpha^{n-1}] .$$
The binary subfield subcode $C_b$ of $C$ is the Hamming code of length $2^m-1$.

There are two types of cosets of $C_b$ with weight enumerators $A_H(X)$ and $A_{CH}(X)$ respectively. $A_H(X)$ is the weight enumerator of $C_b$. $A_{CH}(X)$ is the weight enumerator for the cosets with minimum weight equal to 1, and is given by (14).

For $\bar{v}=(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m) \in C_b$, $\bar{v}_i$ belongs to a coset with weight enumerator $A_{CH}$ if and only if $\bar{v}_i = H \bar{v}_i^T \neq 0$. Then $N_s$ with $0 \leq s \leq m$ is equal to the number of $(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_m)$'s with $s$ nonzero components for which

$$\sum_{i=1}^{m} \bar{y}_i = 0.$$ 

Hence, $N_s$ is the same as the number of codewords of weight $s$ in a maximum distance separable code of length $m$ and minimum distance 2 with symbols from $GF(2^m)$. Consequently, we have [1,2]

$$N_s = \binom{n}{s} \sum_{j=0}^{s-2} (-1)^j \binom{s-j}{j} (2^{m(s-j-1)})^{-1}.$$ (16)

**Case 2:** $\bar{g}(X) = (X-1)(X-a)$.

In this case, $C_e$ has minimum distance 3. It follows from Theorem 2 that

$$A_e^b(X) = \sum_{s=0}^{m} N_s [A_H,X(X)]^{m-s}[A_{CH},X(X)]^s,$$ (17)

where $N_s$ is given by (16), $A_{H,e}$ and $A_{CH,e}$ are the even parts of $A_H$ and $A_{CH}$ respectively. From Theorem 3, $A_e^b$ can be obtained.

**Case 3:** $\bar{g}(X) = (X-a)(X-a^2)$.

In this case, $C$ has minimum distance 3 and

$$H = \begin{bmatrix} 1 & \alpha^2 & \ldots & \alpha^{n-1} \\ 1 & \alpha^2 & 4 & \ldots & \alpha^{2(n-1)} \end{bmatrix}.$$ (18)

The binary subfield subcode $C_b$ is the Hamming code of length $2^m-1$. For $\bar{v}=(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m)$, let
Since \( \tilde{v}_i \) is binary, we have
\[
\gamma_{12} = \gamma_{11}^2.
\] (20)

Then \( \tilde{v} \) is a codeword in \( c^b \) if and only if
\[
\sum_{i=1}^{m} \beta_i \gamma_{1i} = 0,
\] (21)
\[
\sum_{i=1}^{m} \beta_i \gamma_{1i}^2 = 0.
\]

Note that \( \tilde{v}_i \) is in a coset with weight enumerator \( A_{CH}(X) \) if and only if \( \gamma_{11} \neq 0 \). Since
\[
\sum_{i=1}^{m} \beta_i \gamma_{1i} = 0,
\]
if and only if
\[
\sum_{i=1}^{m} \beta_i \gamma_{1i}^2 = 0.
\]

\( N_s \) is equal to the number of \( m \)-tuples, \( (\delta_1, \delta_2, \ldots, \delta_m) \), over \( GF(2^m) \) with \( s \) nonzero components for which
\[
\sum_{i=1}^{m} \beta_i \delta_i = 0,
\] (22)
\[
\sum_{i=1}^{m} \beta_i^2 \delta_i = 0.
\]

Since, for \( 1 \leq i < j \leq m \),
\[
\begin{vmatrix}
\beta_i & \beta_j \\
\beta_i^2 & \beta_j^2 \\
\end{vmatrix} \neq 0,
\]
$N_s$ is equal to the number of codewords of weight $s$ in a maximum distance separable code of length $m$ and minimum weight 3, and is given by (1,2),

\[ N_s = \binom{m}{s} \sum_{j=0}^{s-3} (-1)^{j} \binom{s}{j} (2^m s-j-2) -1 \]  

(23)

Then it follows from Theorem 1 that

\[ A^b(X) = \sum_{s=0}^{m} N_s [A_{H}(X)]^{m-s} [A_{CH}(X)]^{s} \]  

(24)

where $N_s$ is given by (23).

**Case 4:** $g(X) = (X-1)(X-\alpha)(X-\alpha^2)$

In this case, $C_e$ has minimum distance 4. It follows from Theorem 2 that

\[ A^e_b(X) = \sum_{s=0}^{m} N_s [A_{H,e}(X)]^{m-s} [A_{CH,e}(X)]^{s} \]  

(25)

where $N_s$ is given by (23). Also, it follows from Theorem 3 that $A^e_{ex}(X)$ can be obtained.

For all the cases considered above, the binary weight distribution is independent of the choice of the basis $\{\beta_1, \beta_2, \ldots, \beta_m\}$.

**Case 5:** $g(X) = (X-\alpha)(X-\alpha^3)$, or $(X-\alpha)(X-\alpha^2)(X-\alpha^2)(X-\alpha^3)$ or $(X-\alpha)(X-\alpha^2)(X-\alpha^3)(X-\alpha^4)$

In either case, $C_b$ is the primitive BCH code of length $2^m-1$ and minimum distance 5. Hence $C_b$ is quasi-perfect [2-4]. For odd $m$, $C_b$ satisfies the conditions of Theorem 5, and there are three types of cosets of $C_b$ other than $C_b$ with minimum weights 1, 2, and 3 respectively. The weight enumerator $A_3(x)$ for $1 \leq \ell \leq 2$ can be obtained by MacWilliam's equation given by (15), and $A_3(x)$ is given by the following equation:

\[ A_3(X) = \left[ 2^n - 2^{\ell(k+1)} \right]^{-1} [(X+1)^n - A_0(X) \]  

\[ - nA_1(X) - (2^n)A_2(X) \]  

(26)
Consider the case for which \( g(X) = (X-a)(X-a^3) \). For \( \bar{v} = (\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m) \) with \( \bar{v}_i \) as a binary n-tuple for \( 1 \leq i \leq m \), let

\[
\begin{bmatrix}
Y_{i1} \\
Y_{i3}
\end{bmatrix} = H \bar{v}_i^T.
\]

Then, \( \bar{v} \) is a codeword in \( C_b \) if and only if

\[
\sum_{i=1}^{m} \beta_i Y_{i1} = 0 \quad \text{and} \quad \sum_{i=1}^{m} \beta_i Y_{i3} = 0.
\]

For \( 1 \leq i \leq m \), \( \bar{v}_i \) is a codeword in \( C_b \) if and only if \( Y_{i1} Y_{i3} = 0 \); \( \bar{v}_i \) is in a coset with minimum weight 1 if and only if \( Y_{i1} \neq Y_{i3} \); \( \bar{v}_i \) is in a coset with minimum weight 2 if and only if \( Y_{i1} = Y_{i3} \) and \( \text{trace}(1 + Y_{i3} / Y_{i1}) = 0 \); and otherwise \( \bar{v}_i \) is in a coset with minimum weight 3. A closed formula for

\[
N_{s_1, s_2, s_3}
\]

is under study.

Other interesting cases are: \( g(X) = (X-a)(X-a^{-1}) \) or \( (X-a)(X-a^2)(X-a^{-1})(X-a^{-2}) \).

There exists a cyclic code with the same \( n, k \) and the minimum distance as those of the extended code \( C_{ex} \). For the case with \( g(X) = (X-a)(X-a^{-1}) \), the binary subfield subcode \( C_{ex,b} \) of the cyclic version of \( C_{ex} \) is a Zetterberg's code \([2,6]\) for even \( m \). However, the weight distribution of a coset of \( C_{ex,b} \) is unknown.

4. Conclusion

In this paper, we have investigated the weight distribution of binary linear block codes derived from codes with symbols from \( \text{GF}(2^m) \). Weight enumerators for binary codes derived from some Reed-Solomon codes over \( \text{GF}(2^m) \) have been obtained.

Reed-Solomon codes with symbols from \( \text{GF}(2^m) \) are widely used as the outer codes in a concatenated coding scheme for error control in data communication. Recently, we are investigating a concatenated coding scheme for NASA's Telecommand System. Two possible outer codes are considered, one is the X.25 standard code with generator polynomial \( g(X) = x^{16} + x^{12} + x^5 + 1 \) and
the other is the Reed-Solomon code with symbols from GF(2^8) and generator polynomial \( \bar{g}(X) = (X-1)(X-\alpha) \). The case with X.25 standard code as the outer code has been analyzed. Now we are analyzing the case with the above Reed-Solomon code as the outer code. Knowing the binary weight distribution of the Reed-Solomon code, we should be able to analyze the performance of the proposed concatenated coding scheme for NASA's Telecommand System.
REFERENCES


