(NASA-CR-174375) ON THE BINARY WEIGHT DISTRIBUTION OF SOME REED-SOLOMON CODES (Hawaii Univ., Manoa.) 13 p HC A02/MF A01 CSCL 12A
N85-16628
Unclas
G3/64 14163

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, $(X-\alpha)$, $(X-1)(X-\alpha)$, $(X-\alpha)(X-\alpha^2)$ and $(X-1)(X-\alpha)(X-\alpha^2)$ and their extended codes are presented, where 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 $\{\beta_1,\beta_2,\ldots,\beta_m\}$ 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 $\beta_1,\beta_2,\ldots,\beta_m$ as follows:

$$z = c_1\beta_1 + c_2\beta_2 + \ldots + c_m\beta_m,$$

where $c_i \in GF(2)$ for $1 \le i \le m$. 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 $\{\beta_1,\beta_2,\ldots,\beta_m\}$ 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

$$\lambda = \frac{t}{1+\lfloor(\ell+m-2)/m\rfloor}$$

or fewer bursts of errors of length ℓ [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 α 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 $\{\beta_1,\beta_2,\ldots,\beta_m\}$ 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:

$$H^b = [\beta_1 H : \beta_2 H : \ldots : \beta_m H] \,, \tag{1}$$

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

$$\sum_{i=1}^{m} \beta_i H \bar{v}_i^T = 0 \,. \tag{2}$$

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

$$H_e = \begin{bmatrix} 1 & 1 & . & . & . & 1 \\ & & & H & & \end{bmatrix} . \tag{3}$$

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 <u>even-weight</u> subcode of C_b.

Let $A_0(X) = A_{00}+A_{01}X+A_{02}X^2+\ldots+A_{0,n}X^n$ 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 ℓ types of cosets modulo C_b including C_b itself, and cosets of type-j have the same weight enumerator $A_j(X)$ for $0\leq j<\ell$. Let $\bar{\gamma}$ be a (n-k)-tuple over $GF(2^m)$. Then $\bar{\gamma}$ is said of "type-j" if and only if $\bar{\gamma}$ 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 2ℓ

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 $0 \le j < \ell$.

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

(i) $\bar{\gamma}_i$ is an (n-k)-tuple over GF(2^m) for $1 \le i \le m$;

(ii) the number of components $\bar{\gamma}_i$ of type-j is s_j for $1 \le j < \ell$; and

(iii) the following equality holds

$$\sum_{i=1}^{m} \beta_i \bar{\gamma}_i = 0 . \tag{4}$$

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

<u>Theorem 1</u>: The weight enumerator of C^b, denoted $A^b(X)$, is given by

$$A^b(X) = \sum_{S_{\ell,m}} N_{s_1,s_2,\dots,s_{\ell-1}} [A_0(X)]^{m-\lambda} \sum_{j=1}^{\ell-1} A_j^{s_j}(X) , \tag{5}$$

where $S_{\ell,m} = \{(s_1, s_2, \dots, s_{\ell-1}): \ s_j \ge 0 (1 \le j < \ell) \text{ and } \sum_{j=1}^{\ell-1} s_j \le m\}$ and $\lambda = \sum_{j=1}^{\ell-1} s_j$.

ΔΔ

Let $\bar{v} = (\bar{v}_1, \bar{v}_2, \dots, \bar{v}_m)$ be a binary vector of mn components where $\bar{v} = (v_{i1}, v_{i2}, \dots, v_{i,n})$ is a binary n-tuple for $1 \le i \le m$. 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 $\{\beta_1, \beta_2, \dots, \beta_m\}$. Then $\bar{v}$ is a codeword in C_e^b if and only if

$$\sum_{i=1}^{m} \beta_i \sum_{j=1}^{n} v_{ij} = 0 , \tag{7}$$

$$\sum_{i=1}^{m} \beta_i H \bar{v}_i^T = 0 . \tag{8}$$

Since $\beta_1, \beta_2, \dots, \beta_m$ are linearly independent over GF(2), we have that

$$\sum_{j=1}^{n} v_{ij} = 0, \quad \text{for } 1 \le i \le m . \tag{9}$$

Hence we have Theorem 2.

<u>Theorem 2</u>: The binary Code C_e^b is an even-weight code and its weight enumerator $A_e^b(X)$ is given by

$$A_e^b(X) = \sum_{S_{\ell,m}} N_{s_1,s_2,\ldots,s_{\ell-1}} [A_{0,e}(X)]^{m-\lambda} \prod_{j=1}^{\ell-1} A_{j,e}^{s_j}(X) , \tag{10}$$

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

ΔΔ

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

$$H_{ex} = \begin{bmatrix} 1 & 1 & . & . & . & 1 & 1 \\ & & & & & & 0 \\ & & & & & & . \\ & & & H & & & . \\ & & & & & & . \\ & & & & & & 0 \end{bmatrix} . \tag{11}$$

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

<u>Theorem 3</u>: The weight enumerator $A_{ex}(X)$ of C_{ex} is given by

$$A_{ex}^b(X) = \sum_{S_{\ell,m}} N_{s ,s ,\ldots,s_{\ell-1}} [A_{0,ex}(X)]^{m-\lambda} \prod_{j=1}^{\ell-1} A_{j,ex}^{s_j}(X) \tag{12}$$

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

$$A_{j,ex}(X) = A_{j,e}(X) + XA_{j,o}(X) \tag{13}$$

for $0 \leq j < \ell$.

ΔΔ

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

$$A_{CH}(X) = \frac{1}{n}\{(X+1)^n - A_H(X)\} . \tag{14}$$

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

$$A_t(X) = \binom{n}{t}^{-1} 2^{-(n-k)} \sum_{j=0}^{n} A'_j P_t(j)(1+X)^{n-j}(1-X)^j , \qquad (15)$$

where A'_j 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 nonzero 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]. ΔΔ

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

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

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]

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

<u>Case 2</u>: $\bar{g}(X) = (X-1)(X-\alpha)$.

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

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

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_{ex}^b can be obtained.

<u>Case 3</u>: $\bar{g}(X) = (X-\alpha)(X-\alpha^2)$.

In this case, C has minimum distance 3 and

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

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

$$\begin{bmatrix} \gamma_{i1} \\ \gamma_{i2} \end{bmatrix} \overset{\Delta}{=} H\bar{v}_i^T , \qquad 1 \le i \le m \tag{19}$$

Since $\bar{v}_i$ is binary, we have

$$\gamma_{i2} = \gamma_{i1}^2 . \tag{20}$$

Then $\bar{v}$ is a codeword in C^b if and only if

$$\begin{aligned} \sum_{i=1}^{m} \beta_i \gamma_{i1} &= 0 , \\ \sum_{i=1}^{m} \beta_i \gamma_{i1}^2 &= 0 , \end{aligned} \tag{21}$$

Note that $\bar{v}_i$ is in a coset with weight enumerator $A_{CH}(X)$ if and only if $\gamma_{i1} \neq 0$. Since

$$\sum_{i=1}^{m} \beta_i \gamma_{i1} = 0 ,$$

if and only if

$$\sum_{i=1}^{m} \beta_i^2 \gamma_{i1}^2 = 0 ,$$

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

$$\begin{aligned} \sum_{i=1}^{m} \beta_i \delta_i &= 0 , \\ \sum_{i=1}^{m} \beta_i^2 \delta_i &= 0 . \end{aligned} \tag{22}$$

Since, for $1 \le i < j \le m$,

$$\begin{vmatrix} \beta_i & \beta_j \\ \beta_i^2 & \beta_j^2 \end{vmatrix} \neq 0 ,$$

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],

$$N_s = \binom{m}{s} \sum_{j=0}^{s-3} (-1)^j \binom{s}{j} (2^{m(s-j-2)} - 1) \ . \tag{23}$$

Then it follows from Theorem 1 that

$$A^b(X) = \sum_{s=0}^{m} N_s [A_H(X)]^{m-s} [A_{CH}(X)]^s \ , \tag{24}$$

where N_s is given by (23).

<u>Case 4</u>: $\bar{g}(X) = (X-1)(X-\alpha)(X-\alpha^2)$

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

$$A_e^{(b)}(X) = \sum_{s=0}^{m} N_s [A_{H,e}(X)]^{m-s} [A_{CH,e}(X)]^s \ , \tag{25}$$

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

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

<u>Case 5</u>: $\bar{g}(X) = (X-\alpha)(X-\alpha^3)$, 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 $A_\ell(X)$ for $1 \le \ell \le 2$ can be obtained by MacWilliam's equation given by (15), and $A_3(X)$ is given by the following equation:

$$A_3(X) = [2^n - 2^k(1+n+\binom{n}{2})]^{-1}\{(X+1)^n - A_0(X) - nA_1(X) - \binom{n}{2}A_2(X)\} \ . \tag{26}$$

Consider the case for which $\bar{g}(X)=(X-\alpha)(X-\alpha^3)$. 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

$$\begin{bmatrix}\gamma_{i1}\\ \gamma_{i3}\end{bmatrix} \triangleq H\bar{v}_i^T .$$

Then, $\bar{v}$ is a codeword in C^b if and only if

$$\sum_{i=1}^{m} \beta_i\gamma_{i1} = 0 \quad \text{and} \quad \sum_{i=1}^{m} \beta_i\gamma_{i3} = 0 .$$

For $1\leq i\leq m$, $\bar{v}_i$ is a codeword in C_b if and only if $\gamma_{i1}=\gamma_{i3}=0$; $\bar{v}_i$ is in a coset with minimum weight 1 if and only if $\gamma_{i3}=\gamma_{i1}\neq 0$; $\bar{v}_i$ is in a coset with minimum weight 2 if and only if $\gamma_{i1}\neq 0$ and trace $(1+\gamma_{i3}/\gamma_{i1})=0$; 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: $\bar{g}(X)=(X-\alpha)(X-\alpha^{-1})$ or $(X-\alpha)(X-\alpha^2)(X-\alpha^{-1})(X-\alpha^{-2})$. 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 $\bar{g}(X) = (X-\alpha)(X-\alpha^{-1})$, 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 GF(2^m). Weight enumerators for binary codes derived from some Reed-Solomon codes over GF(2^m) have been obtained.

Reed-Solomon codes with symbols from 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 $\bar{g}(X) = X^{16}+X^{12}+X^5+1$ 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 outercode 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

1. S. Lin and D.J. Costello, Jr., Error Control Coding: Fundamentals and Applications, Prentice-Hall, New Jersey, 1983.

2. F.J. Macwilliams and N.J.A. Sloane, Theory of Error-Correcting Codes, North Holland, Amsterdam, 1977.

3. E.R. Berlekamp, Algebraic Coding Theory, McGraw-Hill, New York, 1968.

4. W.W. Peterson and E.J. Weldon, Jr., Error-Correcting Codes, 2nd. ed., MIT Press, Cambridge, Mass., 1972.

5. F.J. MacWilliams, "A Theorem on the Distribution of Weights in a Systematic Code, Bell System Technical Journal, Vol. 42, pp. 79-94, 1963

6. L.H. Zetterberg, "Cyclic Codes from Irreducible Polynomials for Correction of Multiple Errors," IEEE Transactions on Information Theory, Vol. 8, pp. 13-20, 1962.