On Quinary Hamming Code for r=2

Farzana Karim Elora
Department of ECE
North South University
Dhaka, Bangladesh

A.K.M.Toyarak Rian
Department of ECE
North South University
Dhaka, Bangladesh

Partha Pratim Dey
Department of ECE
North South University
Dhaka, Bangladesh
Email: ppd {at} northsouth.edu

Abstract— **In this paper we investigate the quinary Hamming code for $r=2$. We construct the code and establish its perfectness. Using the MDS property of the code, we find its weight distribution. We also investigate the dual code and other codes having relation with the Hamming code. Finally we illustrate how this Hamming code can be decoded.**

Key-Words:- **Linear code, generator matrix, parity-check matrix.**

I. INTRODUCTION

Let $GF(q)$ be the Galois field with q elements. An $[n,k]$ linear code over $GF(q)$ is a $k-$ dimensional subspace of $GF(q)^n$, the space of all $n-$ tuples with components from $GF(q)$. Since a linear code is a vector sub-space it can be given by a basis. The matrix whose rows are the basis vectors is called a generator matrix. For an acquaintance with coding theory at a basic level the reader may consult $[1,2,3]$.

A very important concept in coding is the weight of a vector v. By definition, this is the number of non-zero components v has and is denoted by $wt(v)$. The minimum weight of a code, denoted by d, is the weight of a non-zero vector of smallest weight in the code. A well-known theorem [1] says that if d is the minimum weight of a code C, then C can correct $t=\left\lfloor \frac{d-1}{2} \right\rfloor$ or fewer errors, and conversely.

An $[n,k]$ linear code with minimum weight d is often called an $[n,k,d]$ code. A quinary code is a $[n,k,d]$ code over $GF(5)$. Recall that $GF(5)$ denotes the Galois field of order 5 comprising of $0,1,2,3$ and 4 with the following addition and multiplication tables:

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

×	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

II. CONSTRUCTION OF THE CODE

Throughout this paper $GF(5)$ will denote the Galois field of order 5. Then the Cartesian product $GF(5)\times GF(5)$ comprises of the following 25 pairs:

$(0,0),(0,1),(0,2),(0,3),(0,4),$
$(1,0),(1,1),(1,2),(1,3),(1,4),$
$(2,0),(2,1),(2,2),(2,3),(2,4),$
$(3,0),(3,1),(3,2),(3,3),(3,4),$
$(4,0),(4,1),(4,2),(4,3),(4,4).$

We split the 24 nonzero elements of $GF(5)\times GF(5)$ into 6 disjoint sets:

$S_1=\{(1,1),(2,2),(3,3),(4,4)\},$
$S_2=\{(1,2),(2,4),(3,1),(4,3)\},$
$S_3=\{(1,3),(2,1),(3,4),(4,2)\},$
$S_4=\{(1,4),(2,3),(3,2),(4,1)\}$
$S_5=\{(1,0),(2,0),(3,0),(4,0)\}$
$S_6=\{(0,1),(0,2),(0,3),(0,4)\},$

where any two pairs of the same set are multiples of each other over $GF(5)$.

We take 6 pairs, one from each set, namely $(1,1),(1,2),(1,3),(1,4),(1,0),(0,1)$ from $S_1,S_2,...,S_6$ and use their transposes to form the following 2×6 matrix:

$$H=\begin{bmatrix}1&1&1&1&1&0\\1&2&3&4&0&1\end{bmatrix}$$

where each column is a transpose of a pair, like $\begin{bmatrix}1\\1\end{bmatrix}$ is the transpose of $(1,1)$, $\begin{bmatrix}1\\2\end{bmatrix}$ is the transpose of $(1,2)$ and thereafter. The matrix H is called Hamming or parity check matrix in literature.

Let $C=\{x=(x_1,...,x_6)\in GF(5)^6 \mid Hx^T=0\}$. Then C is a subspace of $GF(5)^6$, and therefore a linear code over $GF(5)$. Notice that $Hx^T=0$ implies the following system:

$$\begin{cases}x_1+x_2+x_3+x_4+x_5=0\\x_1+2x_2+3x_3+4x_4+x_6=0\end{cases}\qquad(1)$$

which then yields:

$x_5=4x_1+4x_2+4x_3+4x_4$ and

$x_6=4x_1+3x_2+2x_3+x_4$.

Since x_1,x_2,x_3 and x_4 are free variables we can assign Them conveniently chosen values. Thus setting $x_1=1,x_2=0,x_3=0$ and $x_4=0$, we obtain $x_5=4$ and $x_6=4$. Thus (1,0,0,0,4,4) is a solution of (1). Similarly $(0,1,0,0,4,3)$, $(0,0,1,0,4,2)$ and $(0,0,0,1,4,1)$ are three other solutions of (1) .Thus (1,0,0,0,4,4), $(0,1,0,0,4,3)$, $(0,0,1,0,4,2)$ and $(0,0,0,1,4,1)$ are four code-words of C. Since they are independent, we can use them to build a generator matrix G of C with

$$G=\begin{bmatrix}1&0&0&0&4&4\\0&1&0&0&4&3\\0&0&1&0&4&2\\0&0&0&1&4&1\end{bmatrix}.$$

III. THE WEIGHT DISTRIBUTION

We now show that the minimum weight of the code C is 3.

Theorem(3.1). The minimum weight of the Hamming code C over $GF(5)$ is 3.

Proof. Notice that the weight of each row of G is three.

Now let $\alpha,\beta\neq 0, i\neq j, 1\le i,j\le 4$. Then

$$\alpha(4,i)+\beta(4,j)=(4\alpha,\alpha i)+(4\beta,\beta j)$$
$$=(4(\alpha+\beta),\alpha i+\beta j)$$

Two cases may arise here.

Case 1: $\alpha+\beta=0$.

Then $\beta=-\alpha$ and $\alpha i+\beta j=\alpha i-\alpha j=\alpha(i-j)\neq 0$ as $i\neq j$.

Case 2: $\alpha+\beta\neq 0$.

Then $(\alpha+\beta)\cdot 4\neq 0$.

Hence for $\alpha,\beta\neq 0, i\neq j, 1\le i,j\le 4, \alpha(4,i)+\beta(4,j)\neq(0,0)$.

Hence a linear combination of two rows of G with nonzero scalars from $GF(5)$ has weight at least 3. On the other hand a linear combination of 3 or 4 rows of G has obviously wt at least 3. Hence the minimum weight of C is 3. ■

Corollary. The $[6,4]$ Hamming code over $GF(5)$ corrects 1 error.

Proof. Follows from the fact that if d is the minimum weight of a code C, then C can correct $t=\left\lfloor\frac{d-1}{2}\right\rfloor$ or fewer errors. Since the d of C is 3, the t is 1. ■

By singleton bound [1], for an $[n,k,d]$ code, $d\le n-k+1$ and when $d=n-k+1$, the code is called an MDS or maximum distance separable code. Hence our $[6,4,3]$ code C is an MDS code. We then can apply the following theorem [3] on C.

Theorem(3.2). Let C be an $[n,k,d]$ MDS code over $GF(q)$ and A_i be the number of code-words of weight i. Then $A_0=1$, $A_i=0, 1\le i<d$ and

$$A_i=\binom{n}{i}\sum_{j=0}^{i-d}(-1)^j\binom{i}{j}(q^{i+1-d-j}-1),\ d\le i\le n.$$

By the theorem above, there is exactly one code-word of weight 0 in C and no code-word of weight 1 and 2.

Moreover

$$A_3=\binom{6}{3}\sum_{j=0}^{0}(-1)^j\binom{3}{j}(5^{1-j}-1)$$
$$=\binom{6}{3}(5-1)=\frac{6!}{3!3!}=80,$$

$$A_4=\binom{6}{4}\sum_{j=0}^{1}(-1)^j\binom{4}{j}(5^{2-j}-1)$$
$$=\binom{6}{4}\left[\binom{4}{0}(5^2-1)-\binom{4}{1}(5^1-1)\right]=120,$$

$$A_5 = \binom{6}{5}\sum_{j=0}^{2}(-1)^j\binom{5}{j}(5^{3-j}-1) =$$

$$\binom{6}{5}\left[\binom{5}{0}(5^3-1)-\binom{5}{1}(5^2-1)+\binom{5}{2}(5-1)\right] = 264$$ and

Thus we have the following theorem.

$$A_6 = \binom{6}{6}\sum_{j=0}^{3}(-1)^j\binom{6}{j}(5^{3-j}-1) = 160$$

Theorem (3.3). The code C has the following weight distribution.

Weight	Number of Words
0	1
3	80
4	120
5	264
6	160

IV. A PERFECT CODE

Recall from Section1 above that an $[n,k]$ linear code C over $GF(q)$ is a $k-$dimensional subspace of $GF(q)^n$. A linear code C of minimum weight d is called perfect if all the vectors in $GF(q)^n$ are contained in the spheres of radius $t = \left\lfloor \frac{d-1}{2} \right\rfloor$ about the code-words. In this case the spheres are said to cover the space $GF(q)^n$. It is well known [2] that if there exists a perfect $[n,k,d]$ code over $GF(q)$ then the polynomial $L_t(x)$ defined by

$$L_t(x) = \sum_{j=0}^{t}(-1)^j(q-1)^{t-j}\binom{x-1}{j}\binom{n-x}{t-j}$$

has t distinct integer roots in the interval $[1,n]$. Hence if a $[6,3,3]$ code is perfect the polynomial $L_1(x) = 25-5x$ is sure to have an integer root in the interval $[1,6]$, which indeed it does, namely 5. Hence we expect a $[6,3,3]$ code to be perfect and indeed it turns out to be perfect.

Theorem(4.1) A $[6,4,3]$ code over $GF(5)$ is perfect.

Proof. Recall that a $[6,4,3]$ code C over $GF(5)$ is a subspace of $GF(5)^6$. Hence a code-word $c \in C$ has the following form $c = (s_1, s_2, s_3, s_4, s_5, s_6)$ where $s_i \in GF(5)$. We would like to find out the number of elements in $S_1(c)$, the sphere of radius 1 with center in c. Notice that $S_1(c)$ contains the vectors in $GF(5)^6$ which do not differ with c at all or differ with c in just one coordinate. The only vector in $GF(5)^6$ that do not differ with c is c itself. We now go after the vectors of $GF(5)^6$ which differ with c in just 1 coordinate. Let $v \in GF(5)^6$ and the distance $d(v,c) = 1$. Without loss let v differ from c in 1^{st} coordinate. Then $v = (*, s_2, s_3, s_4, s_5, s_6)$ where $* \in GF(5)$, but $* \neq s_1$. Hence there are four choices for $*$. Similarly for each i, which is one of the five remaining coordinates, there exist 4 vectors in $GF(5)^6$ that differ with c in the i^{th} coordinate. Thus there are in total $6\times 4 = 24$ vectors in $GF(5)^6$ that differ with c in just 1 coordinate. Hence taking c into account there are $1+24 = 25$ vectors that differ with c in at most 1 coordinate. Thus there are in total 25 vectors in $S_1(c)$. Since it is well known [1] that $S_1(c_1) \cap S_1(c_2) = \phi$ for $c_1 \neq c_2$, $c_1, c_2 \in C$, the spheres $S_1(c)$, $c \in C$ together contain $|C|\cdot 25 = 5^4 \cdot 5^2 = 5^6$ vectors of $GF(5)^6$. Thus $S_1(c)$ s cover the whole space $GF(5)^6$ and a $[6,4,3]$ code over $GF(5)$ is perfect. ■

V. DUAL AND OTHER RELATED CODES

Let $C^{\perp} = \{u \in GF(5)^6 \mid u\cdot c = 0 \text{ for all } c \in C\}$. Then $C^{\perp}$ is called the dual or orthogonal code of C. It is well known [1] that if C is an MDS code then $C^{\perp}$ is also an MDS code. Since $C^{\perp}$ is generated by H in Section 2, $C^{\perp}$ is a $[6,2]$ code with minimum distance $6-2+1=5$. Thus $C^{\perp}$ can correct 2 errors.

Next we use Theorem (3.2) to compute the weight distribution of $C^{\perp}$. By this theorem, there is exactly one code-word of weight 0 in C and no code-word of weight 1 through 4. Moreover

$$A_5 = \binom{6}{5}\sum_{j}^{0}(-1)^j\binom{5}{j}(5^{5+1-5-j}-1)$$

$$= 6\binom{5}{0}(5-4) = 24$$

$$A_6 = \binom{6}{6}\sum_{j}^{0}(-1)^j\binom{6}{j}(5^{6+1-5-j}-1)$$

$$= \sum_{j=0}^{1}(-1)^j\binom{6}{j}(5^{2-j}-1) = 0$$

Thus we have the following theorem.

Theorem(5.1) Each non-zero code-word of $C^{\perp}$ has weight 5.

Let us now check if C is cyclic. Notice that

$$H\begin{bmatrix}0\\1\\0\\0\\4\\3\end{bmatrix} = \begin{bmatrix}1&1&1&1&1&0\\1&2&3&4&0&1\end{bmatrix}\begin{bmatrix}0\\1\\0\\0\\4\\3\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

but $H\begin{bmatrix}3\\0\\1\\0\\0\\4\end{bmatrix} = \begin{bmatrix}1&1&1&1&1&0\\1&2&3&4&0&1\end{bmatrix}\begin{bmatrix}3\\0\\1\\0\\0\\4\end{bmatrix} = \begin{bmatrix}4\\0\end{bmatrix} \neq \begin{bmatrix}0\\0\end{bmatrix}$.

Hence C is not cyclic. Also notice that

$$G\begin{bmatrix}1\\0\\4\\3\\2\\4\end{bmatrix} = \begin{bmatrix}1&0&0&0&4&4\\0&1&0&0&4&3\\0&0&1&0&4&2\\0&0&0&1&4&1\end{bmatrix}\begin{bmatrix}1\\0\\4\\3\\2\\4\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

but

$$G\begin{bmatrix}4\\1\\0\\4\\3\\2\end{bmatrix} = \begin{bmatrix}1&0&0&0&4&4\\0&1&0&0&4&3\\0&0&1&0&4&2\\0&0&0&1&4&1\end{bmatrix}\begin{bmatrix}4\\1\\0\\4\\3\\2\end{bmatrix} = \begin{bmatrix}4\\4\\1\\3\end{bmatrix} \neq \begin{bmatrix}0\\0\\0\\0\end{bmatrix}.$$

Hence $C^{\perp}$ is not cyclic.

Next we investigate $C + C^{\perp}$. The generator matrix M of $C + C^{\perp}$ is given by

$$M = \begin{bmatrix}G\\H\end{bmatrix} = \begin{bmatrix}1&0&0&0&4&4\\0&1&0&0&4&3\\0&0&1&0&4&2\\0&0&0&1&4&1\\1&1&1&1&1&0\\1&2&3&4&0&1\end{bmatrix}.$$

Using elementary row operations one can reduce it to the following form

$$\begin{bmatrix}1&0&0&0&4&4\\0&1&0&0&4&3\\0&0&1&0&4&2\\0&0&0&1&4&1\\0&0&0&0&0&1\\0&0&0&0&0&0\end{bmatrix}.$$

Hence $\dim(C + C^{\perp}) = 5$ and $wt(C + C^{\perp}) = 1$. Thus $C + C^{\perp}$ has no error-correction capacity.

Recall that $\dim(C + C^{\perp}) = \dim C + \dim C^{\perp} - \dim(C \cap C^{\perp})$.

Hence $\dim(C \cap C^{\perp}) = 1$. We notice that $(1,1,1,1,1,0)$ is the first row of H and it is also obtained when we add all the rows of G. Hence $C \cap C^{\perp}$ comprises of all multiples of $(1,1,1,1,1,0)$ over $GF(5)$ and has minimum weight 5. Thus we have the following theorem.

Theorem(5.2) The following hold.

i. Neither C nor $C^{\perp}$ is cyclic.

ii. $C + C^{\perp}$ has dimension 5 and minimum weight 1. Hence $C + C^{\perp}$ can not correct any error.

iii. $C \cap C^{\perp}$ has dimension 1 and minimum weight 5. Hence $C \cap C^{\perp}$ can correct 2 errors.

VI. DECODING ALGORITHM

Below we illustrate how decoding is done with a $[6,4,3]$ quinary code. By Section 2 above

$$H = \begin{bmatrix}1&1&1&1&1&0\\1&2&3&4&0&1\end{bmatrix}.$$

Suppose we send the code-word $c = (1,1,1,1,1,0)$, but due to noise in the transmission channel an error occurs and the received vector is $r = (1,1,4,1,1,0)$. Then $r - c = (0,0,3,0,0,0)$. This vector is called error vector and

is denoted by e. To recover the code-word c from the received vector r, we compute Hr^t as follows:

$$Hr^t$$
$$= H(c + (r - c))^t$$
$$= H(c + e)^t$$
$$= Hc^t + He^t.$$

Since $c \in KerH$, $Hc^t = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

$$Hr^t$$
$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 3 \end{bmatrix}$$
$= 3 \cdot 3^{rd}$ column of H.

This shows that the error vector e contains the field element 3 in the 3rd bit and error has occurred in the 3rd bit of the code-word c. Since $r - c = e$, we obtain c from $r - e$:

$$c = r - e = (1,1,4,1,1,0) - (0,0,3,0,0,0) = (1,1,1,1,1,0).$$

Thus we have the following decoding algorithm:

1. Form H using the basis vectors of $KerG$
2. Compute Hr^t
   a. If $Hr^t = \alpha \cdot j^{th}$ column of H where $j \in \{1,2,3,4,5,6\}$ and $\alpha \in GF(5)$ such that $\alpha \neq 0$, then the error has occurred in the j^{th} bit of the sent code-word and the error vector e has field element α in its j^{th} coordinate position and zeros in others
   i.e. $e = (0,0,..,\alpha,..,0,0)$ where α is the j^{th} bit of e.
   b. If Hr^t is zero vector, no error has occurred during transmission.

Compute $r - e$ to recover the code-word c.

REFERENCES

[1] V. Pless, Introduction to the Theory of Error Correcting Codes, Wiley Student Edition, John Wiley & Sons (Asia) Pte.Ltd., Singapore, 2003.
[2] R. Hill, A First Course in Coding Theory, The Oxford University Press, Oxford, UK, 1986.
[3] W. C. Huffman and V. Pless, Fundamentals of Error Correcting Codes, Cambridge University Press, New York,2003.