RESEARCH PAPER | Mathematics | Volume : 4 | Issue : 12 | Dec 2014 | ISSN - 2249-555X

A NOTE ON EXTENTED AND PUNCTURED CODES

KEYWORDS: Extented codes,Punctured Codes,Encoding,Decoding,generator Matrix

Dr.M.Mary Jansi Rani	Mrs.S.Tamil Mani	Mrs.A.Sathiya
Head and Assistant Professor, Department of Mathematics, Thanthai Hans Roever College, Perambalur.	Assistant Professor, Department of Mathematics,Thanthai Hans Roever college, Perambalur	Assistant Professor, Department of Mathematics,Thanthai Hans Roever college,Perambalur

ABSTRACT The transmission of a message through a 'noisy' channel is done by choosing efficient encoding and decoding function. This note studies the role of linear transformations $L:F_q{}^k \to F_q{}^n$ where $k<n$ and $F_q{}^k$ & $F_q{}^k$ are vector spaces of dimension k and n respectively over a finite field $F_q(q=p^m, p$ is a prime, $m\geq 1)$ the mathematical background of extended and punctured linear codes is highlighted in the following theorems:
Let $L:F_q{}^k \to F_2{}^n (k<n)$ be an encoding function giving a linear code Im $L=C$. If $E:F_2{}^n \to F_2{}^{n+1}$ is the $[n+1,\ n]$ parity check code, the composite $E\circ L$ gives a linear code. Further, if the minimum distance for c is $2l+1$, then $E\circ L$ gives a linear code with minimum distance $2l+2$.
If $L:F_2{}^k \to F_2{}^h$, $E^1:F_2{}^n \to F_2{}^r$ give linear codes and if the generator matrices associated with them are $G\&G'$ repetitively, then the generator matrix associated with $E'\circ L$ is GG'.
The extended Hamming codes and Reed-Muller codes are shown as illustrations.

INTRODUCTION Let F_q denotes a field of q elements where $q=p^m\,(p$ a prime $m\geq 1)$, $F_q{}^n$ stands for vector space of n - tuples $a_0,a_1,..........a_{n-1}$, where $a_i\in F_q\,(i=0,1,2............n-1)$ over F_q. Let $v(k,q)$ be a vector space of dimension k over F_q. A function E : V(k,q) → $F_q{}^n$ is called an encoding function. We take $k<n$ Image of $V\ (k,\ q)$ under E written I m E = C is a sub space $F_q{}^n$. C = I m E is called a linear **code**. An [n, k] - linear code consists of the encoding function E:V(k,q) → $F_q{}^n$ and a decoding function $D:F_q{}^n \to V(n,q)$, $k<n$ indicates that the function E will be adding "Check digits" to the original message. Given a long message, we treat it into blocks of length n. We assume that E is 1-1 so that no two message blocks have the same code word. A channel T transmits each digit with probability of error p and D decodes received blocks into blocks of length k. We seek to choose $E\&D$ in such a way that the probability that a decoded block will equal the original message block will be high. There are two additional requirements.

First, we seek an efficient code that does not transmit Γ_{00} many extra digits.(which are elements of F_q possibly repeated). $R=k/n$ is called the rate of the code. If R is close to 1 the code will be efficient.

Secondly,the code is useless the functions E&D can be implemented in practice say by digital electronic is unity.

Next usually p is small so that a code word will be transmitted without error and most received words Containing an error will contain only one error. They are called single error – correcting codes which decode all received words containing at most one error multiple error – correcting codes are to be considered when p is not small is called a binary linear code when q = 2. $F_2=\{0,1\}$. If we define $E:F_2{}^n \to F_2{}^{n+1}$ by

$$E(a_0,a_1..........a_{n-1})=a_0\,a_1.....a_{n-1}\,a_n \longrightarrow \quad (0.1)$$

Volume : 4 | Issue : 12 | Dec 2014 | ISSN - 2249-555X

where $a_n = a_0 + a_1 + ...a_{n-1}\left(a_i = 0\ or, i = 0,1,......n\right) \longrightarrow$ (0.2)

We notice that $a_n = 0\,or\,1$ according as the number of 1's in $a_0, a_1.......a_{n-1}$ is error or odd.

Definition 0.1 The weight of a code word $\vec{c} = c_0c_1.......c_{n-1}$ is the number of non zero digits occurring among $c_0, c_1...........c_{n-1}$. It is denoted by $\omega t\left(\vec{c}\right)$.

Definition 0.2 Let C be an $\left[n,k\right]$ binary linear code. For $\vec{a} = a_0a_1......a_{n-1}$, $\vec{b} = b_0b_1.......b_{n-1}$ (elements of C) the distance $d\left(\vec{a},\vec{b}\right)$ is $wt\left(\vec{a}+\vec{b}\right)$, the number of locations i with $a_i \neq b_i$ $\left(i = 0,1,2,.......n-1\right)$

For $\vec{a} \in c$, if $\vec{r}$ is the received word $\vec{r} = r_0r_1.......r_{n-1}$ the error-pattern $\vec{e} = e_0, e_1,...........e_{n-1}$ is such that

$$e_i = \begin{cases} 0 & if\ a_i = r_i \\ 1 & if\ a_i = r_i \end{cases} \qquad \left(i = 0,..........n-1\right) \longrightarrow \qquad \left(0.3\right)$$

We note that $\vec{a} = \vec{r} + \vec{e} \longrightarrow$ (0.4)

Next, we state two theorems without proof, They have been drawn from Dornhoff and Hohn [2].

Theorem I A code C can detect all error pattern of weight ≤ 1, if and only if, the minimum distance between code words is at coast t+1.

Theorem II If the minimum distance between code words is at least $2t+1$, we can choose a decoding function D that will correct all error -patterns of weight $\leq t$.

Definition 0.3 Let $k < n$, A $k \times n$ matrix with entries from $\mathsf{F}_2 = \left\{0,1\right\}$ is called a generator matrix G if its first k columns form I_k (the $k \times k$ unit matrix) given such a matrix G we can define an encoding function $E : F_2^k \rightarrow F_2^n$ by $E\left(\overrightarrow{X}\right) = \overrightarrow{X}G \longrightarrow$ (0.5)

Where $\overrightarrow{X}$ a vector is expressed as a row vector F_2^k Im (0.5) $\overrightarrow{X}G$ means $[X_0X_1........X_k]\left[I_k/A\right]$ where A is a $k\times\ n-k$ matrix. So that G is a $k \times n$ matrix. Clearly $\overrightarrow{X}G$ is a $1\times n$ matrix representing a row vector $\in \mathsf{F}_q^{\ n}$.

Definition0.4 Let $k < n$, An $\left(n-k\right)\times n$ matrix H whose last $\left(\ n-k\right)$ columns are I_{n-k} $\left[the\left(n-k\right)\times\left(n-k\right)unit\ matrix\right]$ is called a Parity check matrix .

The parity check matrix provided an encoding function $E : \mathsf{F}_2^{\ k} \rightarrow \mathsf{F}_2^{\ w}$. For any message word $\vec{w} \in \mathsf{F}_2^{\ k}$. The codeword in the unique word $E\left(\vec{w}\right) \in \mathsf{F}_2^{\ W}$ whose 1st k digits are the digits of $\vec{w}$ and whose remaining digits are determined by the equation.

Volume : 4 | Issue : 12 | Dec 2014 | ISSN - 2249-555X

Proof: Suppose that some column (say i^{th}) of H is $\vec{0}$. Then if $\vec{e} = 000...10...0$ (i^{th} *digits*) and $\vec{C}$ is any code word . Then $H\left(\vec{C}+\vec{e}\right)^T = \vec{0}$ so $\vec{C}+\vec{e}$ appears to be a code word and any error in the i^{th} digits will not be detected at all . Let $\vec{C}$ be a code word, then $H\vec{C}^T = \vec{0}$

Since the received word $\vec{r} = \vec{c}+\vec{e}$

$$H\vec{r}^T = H\left(\vec{c}+\vec{e}\right)^T = H\vec{c}^T + H\vec{e}^T$$

$$= \vec{0} + H\vec{e}^T$$

$$= H\vec{e}^T$$

= the i^{th} column of H.

So, any error pattern of weight 1 will be decoded correctly . We write $H\vec{r} = \vec{A}$ and call $\vec{A}$, the syndrome.

If $\vec{A} = \vec{0}$ transmission was probably correct.

If $\vec{A}$ is the i^{th} column of H, these was probably a single error in the i^{th} digits

If $\vec{A}$ is neither $\vec{0}$ nor the i^{th} column at least two errors must have occurred with transmission.

Also, if i^{th} column $= j^{th}$ column= $\vec{A}$ we cannot tell if the error is in the i^{th} or j^{th} digit. So H will decode all single errors correctly if the columns of H are non zero and distinct.

conversely, if the columns of H are non zero and distinct , then H will decode all single errors correctly , by the property of the syndrome $\vec{A}$.

Next, we denote the minimum distance between code words of a lode c by d. we specify C as an [n, k, d] Code. We emphasize the role of the parity check matrix in the following manner:

An encoding function $E: \mathsf{F}_2^{\,k} \to \mathsf{F}_2^{\,n}$ defined by a parity check matrix H can correct all single errors if and only if the columns of H are non-zero and distinct.

The $(n-k)\times n$ matrix H produces $(n-k)$ parity check equations via $H\left(E\left(\vec{w}\right)\right)^T = \vec{0}$. These equations determined an [n, k, d] code. The number of information digits is k. For a fixed number $(n-k)$ of parity check equations we want to send as much information as possible. So we make the number of columns n of H as large as possible. So we take $n = 2^k - 1$ the number of non– zero $(n-k)$ digits columns which are the binary representation of the numbers 1,2, 3,…….$2^{n-k} - 1$. Then we obtain an $\left[2^{n-k} - 1, k\right]$ code . As no two columns of H are multiples of are another, the code, so obtained will have minimum weight at least 3. It can be shown that the minimum weight of such a code is 3.

Volume : 4 | Issue : 12 | Dec 2014 | ISSN - 2249-555X

Proof: Suppose that some column (say i^{th}) of H is $\vec{0}$. Then if $\vec{e} = 000...10...0$ (i^{th} *digits*) and $\vec{C}$ is any code word . Then $H\left(\vec{C}+\vec{e}\right)^{T} = \vec{0}$ so $\vec{C}+\vec{e}$ appears to be a code word and any error in the i^{th} digits will not be detected at all . Let $\vec{C}$ be a code word, then $H\vec{C}^{T} = \vec{0}$

Since the received word $\vec{r} = \vec{c} + \vec{e}$

$$\begin{aligned} H\vec{r}^{T} &= H\left(\vec{c}+\vec{e}\right)^{T} = H\vec{c}^{T} + H\vec{e}^{T} \\ &= \vec{0} + H\vec{e}^{T} \\ &= H\vec{e}^{T} \\ &= \text{ the } i^{th} \text{ column of H.} \end{aligned}$$

So, any error pattern of weight 1 will be decoded correctly . We write $H\vec{r} = \vec{A}$ and call $\vec{A}$, the syndrome.

If $\vec{A} = \vec{0}$ transmission was probably correct.

If $\vec{A}$ is the i^{th} column of H, these was probably a single error in the i^{th} digits

If $\vec{A}$ is neither $\vec{0}$ nor the i^{th} column at least two errors must have occurred with transmission.

Also, if i^{th} column $= j^{th}$ column= $\vec{A}$ we cannot tell if the error is in the i^{th} or j^{th} digit. So H will decode all single errors correctly if the columns of H are non zero and distinct.

conversely, if the columns of H are non zero and distinct , then H will decode all single errors correctly , by the property of the syndrome $\vec{A}$.

Next, we denote the minimum distance between code words of a lode c by d. we specify C as an [n, k, d] Code. We emphasize the role of the parity check matrix in the following manner:

An encoding function $E : \mathsf{F}_2^{\,k} \to \mathsf{F}_2^{\,n}$ defined by a parity check matrix H can correct all single errors if and only if the columns of H are non-zero and distinct.

The $(n-k)\times n$ matrix H produces $(n-k)$ parity check equations via $H\left(E\left(\vec{w}\right)\right)^{T} = \vec{0}$. These equations determined an [n, k, d] code. The number of information digits is k. For a fixed number $(n-k)$ of parity check equations we want to send as much information as possible. So we make the number of columns n of H as large as possible. So we take $n = 2^{k} - 1$ the number of non– zero $(n-k)$ digits columns which are the binary representation of the numbers 1,2, 3,…….$2^{n-k} - 1$. Then we obtain an $\left[2^{n-k} - 1, k\right]$ code . As no two columns of H are multiples of are another, the code, so obtained will have minimum weight at least 3. It can be shown that the minimum weight of such a code is 3.

Volume : 4 | Issue : 12 | Dec 2014 | ISSN - 2249-555X

If $n = 2^r - 1, r = n - k\ (o\ )k = n - r = 2^r - 1 - r$.So, we get $a\left[2^r - 1, 2^r - 1 - r, 3\right]$ code called the Hamming codes.

1. EXTENDED CODES

Before we go to extended codes, we need the basic ideas about elementary row transformations of $(m \times n)$ matrices with entries from the set $\mathbb{R}$ of real numbers.

The matrix units are the $(n \times n)$ square matrices T_{ij} defined by

$$T_{ij} = \begin{array}{cc} & j \\ i & \left[\begin{array}{ccc} & \vdots & \\ \cdots\cdots & 1 & \cdots\cdots \\ & \vdots & \end{array}\right] \end{array} \longrightarrow (1.1)$$

$$= [t_{ij}]$$

$$\text{where } t_{ij} = \begin{cases} 1 & in \quad the\ (i,j)^{th}\ place \\ 0 & elsewhere \end{cases} \longrightarrow (1.2)$$

for a given matrix $A = [aij]$, for replacing the j^{th} column C_j by $C_j + C_k$ we have only to multiply A by $\left[I + T_{ij}\right]$ on the right, For instance to obtain from

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

The matrix $A_1 = \begin{pmatrix} a_{11} & a_{11} + a_{12} & a_{13} \\ a_{21} & a_{21} + a_{22} & a_{23} \\ a_{31} & a_{31} + a_{32} & a_{33} \end{pmatrix}$

We multiply A by $(I + T_{12})$ the right as

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{11} + a_{12} & a_{13} \\ a_{21} & a_{21} + a_{22} & a_{23} \\ a_{31} & a_{31} + a_{32} & a_{33} \end{pmatrix}$$

Further, $A\left[I+T_{13}+\ T_{23}\right]$ gives $\begin{pmatrix} a_{11} & a_{12} & a_{11}+a_{12}+a_{13} \\ a_{21} & a_{22} & a_{21}+a_{22}+a_{23} \\ a_{31} & a_{32} & a_{31}+a_{32}+a_{33} \end{pmatrix}$

Next, we consider binary linear codes defined by $L{:}\mathsf{F}_2^{\,k} \to \mathsf{F}_2^{\,n}$ where $k<n$. The $[n+1,n]$ parity check code $E{:}\mathsf{F}_2^{\,n} \to \mathsf{F}_2^{\,n+1}$ considered in (0, 1) can be obtained via its generator matrix , say $\mathbf{G}_1$ for $\vec{X} \in F_2^{\,n}$, $E\left(\vec{X}\right)=\vec{X}G_1$.

If $A=\begin{pmatrix} a_{11} & a_{12}................a_{1n} \\ a_{21} & a_{22}................a_{2n} \\ .... \\ .... \\ a_{n1} & a_{n2}................a_{nn} \end{pmatrix}$ is the matrix obtained from n basis vectors of the vector space $\mathsf{F}_2^{\,n}$

(1.2) $A\left(I^{'}+T_{1n+1}+T_{2n+1}..........T_{nn+1}\right)=G_1$ where $I^1 = The\ n\times\left(n+1\right)$ matrix in which $\left[I_n/0\right]^{(n+1)^{th}}$ column has Zero) $T_{i,n+1}$ is the $n\times n+1$ matrix in which the elements of $\left(i,n+1\right)^{th}$ place is 1 and zero at other entries. For n=3, we see that,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{11}+a_{12}+a_{13} \\ a_{21} & a_{22} & a_{23} & a_{21}+a_{22}+a_{23} \\ a_{31} & a_{32} & a_{33} & a_{31}+a_{32}+a_{33} \end{pmatrix}$$

So G_1 is determinable.

Theorem 1 Let $L{:}\mathsf{F}_2^{\,k} \to \mathsf{F}_2^{\,n}\ \left(k<n\right)$ be an encoding function giving a linear code $\operatorname{Im} L = C$. If $E{:}\mathsf{F}_2^{\,n} \to \mathsf{F}_2^{\,n+1}$ is the (n+1, n) parity check code, the composite $E\circ L$ is a linear code. Further, if the minimum distance for C in $2l+1$, then $E\circ L$ gives a linear code with minimum distance $2l+2$.

Proof: Let G_1 be the generator matrix for G_1. G_1 is an $\mathrm{n}\times\mathrm{n}{+}1$ matrix as shown in (1.2). We know that for $\vec{X} \in \mathsf{F}_2^{\,k}, L\left(\vec{X}\right)=\vec{X}G$ where G is the generator matrix for L. For $\vec{y} \in \mathsf{F}_2^{\,n}$, $\mathsf{F}\left(\vec{y}\right)=\vec{y}G_1$.

As $\vec{y}=L\left(\vec{X}\right), \vec{X} \in \mathsf{F}_2^{\,k}$

$E\left(\vec{y}\right)=E\left(\vec{X}G\right)=\left(\vec{X}G\right)G_1=\left(\vec{X}\right)GG_1$

$E\circ L$ is a linear code having the generator matrix GG_1

Next, Let $2l+1$ be the minimum distance for L. For $\vec{A}=\left(A_0\ A_1 ....A_{k-1}\right)\in \mathsf{F}_2^{\ k}$

$E\circ L\left(\vec{A}\right)=E\circ L\left(A_0\ A_1 ....A_{k-1}\right)=A_0\ A_1 .....A_{k-1}\ A_n$ where $A_n=A_0+A_1+......A_{n-1}$ by vertex of the property of G_1. $A_n=0$ or 1, So if minimum distance of C is $2l+1$, $E\circ L$ gives a code whose minimum distance is $2l+2$.

Corollary For the Hamming code $H_{2,r}\left[2^r-1,2^r-1-r,3\right]$ the extended Hamming code is $\left[2^r,2^r-1-r,4\right]$ which is the Reed- Muller code of length 2^r.

2. PUNCTURING OF CODES

We consider binary linear codes defined by $L:\mathsf{F}_2^{\ k}\to \mathsf{F}_2^{\ n}$ and $E_1:\mathsf{F}_2^{\ n}\to \mathsf{F}_2^{\ r}$ where $r<n$, As in section 1, the composite $E_1\circ L$ in also a linear code. If $C^1=\operatorname{Im} L$ has the generator matrix $G\left(a\quad k\times n\ \ matrix\right)$ and $C^1=\operatorname{Im} E_1$ has the generator matrix G^1 ($n\times r$ matrix), the generator matrix for $E_1\circ L$ is GG^1 which is a $k\times r$ matrix.

The effect of $E_1\circ L$ is to transform a code word $\vec{c}$ of length n to a code word, $\vec{c}^{\,1}$ of length r, The number of columns of G G^1 will be less than the number of columns of G . When $r=n-1$, it amounts to puncturing the code $\vec{C}$ represented by G, by deleting the same coordinate i from each code word. The resulting code c^1 is still linear and has length $\left(n-1\right)$ (we denote the punctured code by C*)

Theorem 2 If $L:\mathsf{F}_2^{\ k}\to \mathsf{F}_2^{\ n}, E^1:\mathsf{F}_2^{\ n}\to \mathsf{F}_2^{\ r}$ give linear codes and if the generator matrices associated with them as G& G' repetitively, the generator matrix associated with $E'\circ L$ is given by GG'.

Proof is similar to that of theorem 1.

Corollary The Reed -Muller code R(r, m) is a $\left[2^m,k,2^{m-r}\right]$ code where k is its dimension

$1+\binom{m}{1}+\binom{m}{2}............\binom{m}{r}$ and r < m. The puncturing of R (r, m) yields a binary code $\left[2^m-1,k,2^{m-r}-1\right]$,when m=3, puncturing of [8,4,4] code gives the binary Hamming code [7, 4, 3] .

Remark C [n, k ,d] denotes a binary linear code. To puncture C is to delete the same coordinate i from each code word. The punctured code is denoted by $C^*\left[n+1,k,d^*\right]$. If G denotes the generator matrix of C the

Volume : 4 | Issue : 12 | Dec 2014 | ISSN - 2249-555X

generator G* of C* is obtained from GG'. G' is the $n \times (n-1)$ matrix which is got from the $(n \times n)$ unit matrix by deleting the i^{th} column. In the case where C is a $[24, 12, 8]$ (Golay) code, by puncturing in any of the coordinates, we obtain C* = [23 , 12, 7] binary code.

REFERENCE

1. G.Cohen . Jr and J.B. Cain Error –Correction Coding for Digital Communication, Plenum Press, New York 1981. | 2. L.L.Dornhoff & F.E.Hoh Applied Modern Algebra, Macmillian Publishing &co., Ny (1978), Chapter 5 pages 211-235. | 3. W.C.Huffman & Vera Pless Fundamentals Of Error – Correcting Codes , Cambridge University Press, First South Asian Edition (2004),Chapter 12.1 and 12.2 | 4. Macwilliams ,R.J.Odlyzko,A.M, Sloane N.J.A and Ward H.N Self dual codes over GF(4),J.comb.Theory,Vol 28 (1978) 288-318. | 5. Michael Artin Algebra, prentice Hall of India (p) Ltd , New Delhi (1994) pages 1 to 18. | 6. F.J Macwilliams and N.J.A Sloane The Theory of Error Correcting code, New York , | North –Holland ,1997. | 7. San Ling & Chapping Xing A first course in coding theory , Cambridge University. | 8. Sivaramakrishnan R ., certain Number Theoretic Episodes in Algebra , CRC Press (2006) Florida, Champman & Hall /CRC | 9. Tom M. Apostol Introduction to Analytic Number Theory , Narosa Pub.House , New Delhi (1985), Reprint. | 10. J.H. Van Lint Introduction to coding theory, Volume 86 of Graduate Texts in Mathematics Springer -Verlag , Berlin ,Third edition ,1999. |