SOME ASPECTS OF THE CONSTRUCTION AND IMPLEMENTATION OF ERROR-CORRECTING LINEAR CODES

Geoffrey L. Booth

A dissertation in partial fulfillment of the requirements for the degree of Master of Science in Pure Mathematics of Rhodes University.

Grahamstown, July, 1977.

CONTENTS

ACKNOWLEDGEMENTS

I am indebted to Dr. Denis Riordan and Professor H.J. Schutte for their advice and guidance, and to Daisy Turner for her hard work in typing this manuscript.

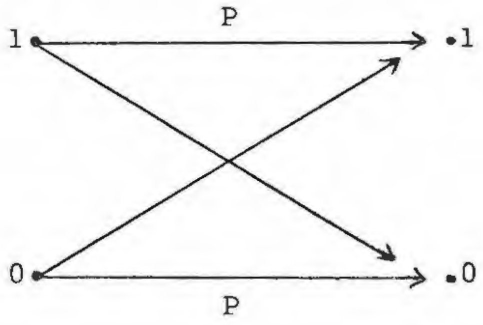
CHAPTER 1

INTRODUCTION TO BINARY ERROR-CORRECTING CODES

In many modern communications systems, such as computer links, it is required to transmit a message consisting of a string of elements of a finite "alphabet" across a channel which is corrupted by "noise". For our purposes, the "alphabet" may be considered to be a finite field containing q elements, GF(q). We will be concerned mainly with binary codes, in which the alphabet contains only two elements, and may be considered to be the field GF(2) = {0,1}.

The Binary Symmetric Channel

Suppose we wish to transmit a message consisting of symbols of two types (o's and 1's) across a noisy channel, and that either type of symbol has a probability p ($0<p<1$) of being correctly received, and a probability q (where $q=1-p$) of being incorrectly received. Such a channel is known as a Binary Symmetric Channel. The binary symmetric channel may be represented diagramatically as follows:-



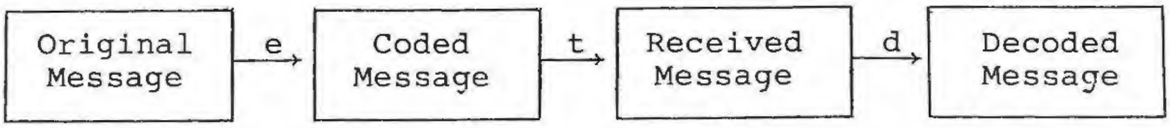
A Binary Block Code is a set V of binary n-tuples of the form $(a_1 \ldots a_n)$, where each a_i is either a 0 or a 1. A block code may be thought of as a subset of the n-dimensional vector space over the field GF(2). Addition of binary n-tuples is defined in the obvious way:-

$$(a_1 \ldots a_n) + (b_1 \ldots b_n) = (c_1 \ldots c_n)$$

where $c_j = 0$ if $a_j = b_j$ and $c_j = 1$ if $a_j \neq b_j$.

Error correction and detection

Let E^m denote the m-dimensional vector space over GF(2). an(m,n) binary code (where $m \leq n$) is a block code $V \subset E^n$, together with an encoding function e: $E^m \rightarrow E^n$, and a decoding function d: $E^n \rightarrow E^m$. Thus we encode binary m-tuples into binary n-tuples, which are our code words, transmit these over a binary symmetric channel, and decode the received codewords. This may be represented diagramatically as follows:-



t is an "error function" caused by channel noise. Clearly, we must have $d \circ e = i$, the identity function on E^m, in order to ensure correct decoding of correctly transmitted code words. e and d will be constructed so that $d \circ e \circ t$ is as close to i as possible, to ensure maximum efficiency in correcting transmission errors.

The case n=m is of limited interest, as codes of this type have no error correcting abilities. They are of some use as cipher codes, but we will not be concerned with this type of code. In the case n>m, the code contains more digits than are actually required to transmit the information. We may say that a code word contains m information symbols, and d=n-m parity-check symbols.

A code is called error correcting if the decoding algorithm is capable of correcting certain errors in the received word, and error detecting is it is capable of detecting that an error has occurred. Our main problem will be to devise codes which correct as many errors as possible, while keeping the ratio of parity check symbols to information symbols to a minimum, in order to ensure maximum efficiency in transmission.

Distance in block codes

The Hamming Weight of a vector $\bar{a}$ in E^n, written $W(\bar{a})$, is the number of symbols equal to 1 in it. The Hamming distance between two elements $\bar{a}$ and $\bar{b}$ of E^n, $d(\bar{a},b)$, is the Hamming weight of $\bar{a}+\bar{b}$. Clearly, this defines a metric on E^n. The minimum distance of a code $V \subset E^n$ is defined as

$d = \min_{\substack{\bar{a},\bar{b} \in V \\ a \neq b}} \{d(\bar{a},\bar{b})\}$. The minimum distance is closely related to the error correcting and detecting abilities of a code. Two trivial, but important results follow immediately from the definition of minimum distance.

Theorem 1.1: An (m,n) binary code will detect all sets of k or fewer errors in the digits of the received word if and only if its minimum distance is at least k+1.

Theorem 1.2: If a code V is capable of correcting all possible combinations of k or fewer errors in the digits of the received code word, then its minimum distance is 2k+1. Conversely, if the code has minimum distance of 2k+1, it is possible to construct a decoding function d which corrects all combinations of k or fewer errors.

The proofs of these theorems are easy. For the converse part of Theorem 1.2 we decode a received vector $\bar{x}$ to code vector $\bar{a}$ such that $d(\bar{x},\bar{a})$ is a minimum.

Group or linear codes. An (m,n) binary code V is called a group code or linear code if the elements of V form a group with respect to the addition operation already defined.

Proposition 1.3: A binary (m,n) code V is a group code if and only if it is a vector subspace of E^n.

Proof: $\Rightarrow$: Suppose V is a group code. Then we need only show that if $\alpha \in GF(2)$, then for every element $\bar{x}$ of V, $\alpha x \in V$. But α can only be 0 or 1. $0\,\bar{x} = \bar{0} \in V$, since V is a group.

$1.\overline{x} = x \in V$.

$\Leftarrow$: Follows directly from the definition of a subspace of E^n. Q.E.D.

Proposition 1.4: Let V be a group code. Then the minimum distance of the code is equal to the Hamming weight of some non zero code vector.

Proof: Let $\overline{a}, \overline{b} \in V$. Then $d(a,b) = W(\overline{a} - \overline{b})$. But since V is a group, $\overline{a} - \overline{b} \in V$. Hence the minimum distance of V is equal to

$$\min_{\substack{\overline{a} \in V \\ a \neq 0}} \{W(a)\}.$$

Q.E.D.

Consider an (m,n) group code V. Now V is a subgroup of E^n. By Lagrange's Theorem, E^n may be decomposed into 2^{n-m} disjoint cosets of V, i.e.

$$V = (\overline{e}_1 + V) \cup (\overline{e}_2 + V) \quad \ldots \quad \cup (\overline{e}_{2^{n-m}} + V),$$

where $(\overline{e}_i + V) \cap (\overline{e}_j + V) = \emptyset$, for $i \neq j$.

We choose the coset leader $\overline{e}_i$ $1 \leq i \leq 2^{n-m}$ to be a code vector of minimal weight in that coset. We construct a decoding table as follows:

Let $V = \{\overline{c}_1 \ldots \overline{c}_{2^m}\}$. We write the elements of the cosets $e_j + V$ $(1 \leq j \leq 2^{n-m})$ as row vectors of a $2^{n-m} \times 2^m$ matrix, with V itself as the top row

$$\begin{array}{ccc}
\overline{c}_1 & \ldots & \overline{c}_{2^m} \\
\overline{c}_1 + \overline{e}_1 & \ldots & \overline{c}_{2^m} + \overline{e}_1 \\
\vdots & & \vdots \\
\overline{c}_1 + \overline{e}_{2^{n-m}} & \ldots & \overline{c}_{2^m} + \overline{e}_{2^{n-m}}
\end{array}$$

Let $V \in E^n$ be a received code vector. Then $\overline{V} = \overline{c} + \overline{e}$, where

$\overline{c}$ is the code vector which was transmitted. $\overline{e}$ is known as the error vector. We decode using our group decoding table as follows.

Suppose $\overline{x} \epsilon E^n$ is received. Then $\overline{x}$ is located in the decoding table and is decoded as the code vector $\overline{c}_j$ at the top of the column in which $\overline{x}$ appears. The following results are immediately obvious:-

Theorem 1.5: Decoding by group decoding table corrects precisely those errors whose error vectors are the coset leaders $e_1 \dots e_{2^{n-m}}$.

Proof: Let $\overline{r} \epsilon E^n$ be a received code vector. Then $\overline{r} \epsilon \overline{e}_j + V$ for some j $(1 \leq j \leq 2^{n-m})$. Thus $\overline{r} = \overline{e}_j + \overline{c}_k$ for some $\overline{c}_k \epsilon V$. Then F will be decoded as $\overline{c}_k$. This decoding will be correct if and only if the error vector of $\overline{r}$ is some coset leader $\overline{e}_j$. Q.E.D.

Theorem 1.6: If $\overline{r}$ is a received code vector, and F is decoded to the code vector $\overline{c}_j$, then $d(\overline{c}_k,\overline{r}) \geq d(\overline{c}_j,\overline{r}) (1 \leq k \leq 2^m)$. In this sense, decoding by group decoding table is an optimal decoding process.

Proof: Suppose $\overline{r}$ is decoded to $\overline{c}_j$. Then

$$\overline{r} = \overline{c}_j + \overline{e}_\ell \text{ for some coset leader } e_\ell$$

$$\begin{aligned}
d(\overline{r},\overline{c}_j) &= W(\overline{r} + \overline{c}_j) \\
&= W(\overline{e}_\ell + \overline{c}_j + \overline{c}_j) - c_j \\
&= W(\overline{e}_\ell) \quad (\text{since } \overline{x} + \overline{x} = 0,\ \forall \overline{x} \epsilon E^n) \\
d(\overline{r},\overline{c}_k) &= W(\overline{c}_j + \overline{e}_\ell + \overline{c}_k) \\
&= W(\overline{c}_j + \overline{c}_k + \overline{e}_\ell)
\end{aligned}$$

Now $\overline{c}_j + \overline{c}_k + \overline{e}_\ell$ is in the same coset as e_ℓ, and by the way in which the coset leaders were chosen, $w(\overline{e}_\ell) \leq w(\overline{c}_m+\overline{c}_k+\overline{e}_\ell)$

i.e. $d(\overline{r},\overline{c}_j) \leq d(\overline{r},\overline{c}_k)$

Q.E.D.

Matrix representation of group codes

Let E be non-singular m × n matrix with entries from GF(2). If $\overline{c} \in E^m$, then $\overline{c}\ E \in E^n$

$$\text{e.g.}\quad (1\ 0\ 1)\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} = (1\ \ 0\ \ 1\ \ 1\ \ 1)$$

Where m = 3, n = 5.

The image V of E^m under E will be an m-dimensional subspace of E^n, and we can thus construct an (m,n) group code in this way. E is known as an encoding matrix of V.

Theorem 1.7: If E is an encoding matrix of an (m,n) code $V \subset E^n$, then V is the row space of E^n.

Proof: Let $E = (e_{ij})$ $(e_{ij} \in GF(2),\ 1 \leq i \leq m,\ 1 \leq j \leq n)$

Let $\overline{a} = (a_1 \ldots a_m) \in E^m$.

$$\overline{a}E = (a_1 \ldots a_m)\begin{bmatrix} e_{11} & \ldots & e_{1n} \\ e_{m1} & \ldots & e_{mm} \end{bmatrix}$$

$$= (b_1 \ldots b_n) = \overline{b}$$

$$\text{where } b_j = \sum_{i=1}^{m} a_i e_{ij} \qquad (1 \leq j \leq n)$$

$$\therefore\quad \overline{b} = a_1(e_{11} \ldots e_{1n}) + a_2(e_{21} \ldots e_{2n}) \ldots$$

$$+ a_m(e_{m1} \ldots e_{mn}),$$ which is a linear combination of the row vectors of E. Therefore $\overline{b}$ is in the row space of H.

Conversely, if $\overline{c}$ is in the row space of E, then

$\exists \lambda_1 \ldots \lambda_m \in GF(2)$ such that

$$\overline{c} = \sum_{i=1}^{m} \lambda_i (e_{i1} \ldots e_{in}).$$

Put $\overline{a} = (\lambda_1 \ldots \lambda_m)$. Then $\overline{a}E = \overline{c}$, therefore $\overline{c} \in V$. Q.E.D.

Now let V be an (m,n) code with encoding matrix E. Then V is an m-dimensional subspace of E^n, and hence has a basis consisting of m elements, $\{\overline{V}_1 \ldots \overline{V}_m\}$. We extend this to a basis for E^n, $\{\overline{V}_1 \ldots \overline{V}_m, \overline{V}_{m+1}, \ldots \overline{V}_n\}$. Now let $\overline{x} \in E^n$. Then $\overline{x}$ can be written

$$\overline{x} = \sum_{i=1}^{n} \lambda_i \overline{V}_i \qquad (\lambda_i \in GF(2) \quad 1 \leq i \leq n)$$

We define $H : E^n \to E^{n-m}$ by $H(\overline{x}) = (\alpha_{m+1} \ldots \alpha_1)$.

Clearly, H defines a linear map, and V is the nullspace of H. The matrix representation of H is known as a parity-check matrix of V. Note that the parity-check matrix for a given code V is not unique, since there are, in general, many ways of extending the basis for V to a basis for E^n.

Let $\overline{x} \in E^n$ be a received vector. Then if the code V is the nullspace of an(n-m) × n matrix H, then H $\overline{x}$ is known as the syndrome of $\overline{x}$ with respect to H.

Theorem 1.8: Let V be a group code which is the nullspace of H. Then two elements $\overline{x}$ and $\overline{y}$ are in the same coset of V in E^n if and only if $H\overline{x}^t = H\overline{y}^t$.

Proof: Suppose that $\overline{x}$ and $\overline{y}$ are in the coset $\overline{e}$ + V. Then there exist $\overline{V}_1, \overline{V}_2 \in$ V such that $\overline{x} = \overline{e} + \overline{V}_1$, $y = \overline{e} + \overline{v}_2$. Then $H(\overline{x}^t, = H(\overline{e}^t + \overline{V}_1{}^t) = H(\overline{e}^t) + H(\overline{V}_1^t) = H(\overline{e}^t)$ (since V is the nullspace of H). Similarly, $H(\overline{y}^t) = H(\overline{e}^t)$. Thus $\overline{x}$ and $\overline{y}$ have the same syndrome.

Conversely, suppose $H(x^{-t}) = H(\overline{y}^t)$. Then $H(\overline{x}^t - \overline{Y}^t) = \overline{0}$, whence $\overline{x} - \overline{y} \in V$. Therefore, $\overline{x} + V = \overline{y} + V$, thus $\overline{x}$ and $\overline{y}$ are in the same coset of V. Q.E.D.

The following result is of importance in determining the error correcting capabilities of a code.

Theorem 1.9: Let V be an (m,n) code which is the nullspace of an $(n-m)\times n$ matrix H. Then there exists a code word $\overline{x}\in V$ with weight w if and only if there is a linear dependence relation between w column vectors of H.

Proof: Let $H\in(h_{ij})$. Suppose $\overline{x}=(x_1\ \ldots\ x_n)\in V$ is a code word of weight w. Then w of the x_i's $(1\leq i\leq n)$ are 1's and n-w are zeroes.

Now $H\overline{x}^t=\overline{0}\Rightarrow\ \sum_{j=1}^{h}h_{ij}x_j=0\quad(1\leq i\leq n-m)$.

Combining these n-m equations, we obtain

$$x_1\begin{pmatrix}h_{11}\\ \vdots\\ \vdots\\ h_{n-m1}\end{pmatrix}+x_2\begin{pmatrix}h_{12}\\ \vdots\\ \vdots\\ h_{n-m2}\end{pmatrix}\quad\cdots\quad+x_h\begin{pmatrix}h_{1n}\\ \vdots\\ \vdots\\ h_{n-mn}\end{pmatrix}=\overline{0}$$

Since w of the x_i's are non-zero, this is a linear dependence relation involving w column vectors of H.

Conversely, suppose the is a linear dependence relation involving w column vectors of H, i.e. the exist $\alpha_{j1}\ \ldots\ \alpha_{jw}$ such that

$$\alpha_{j_1}\begin{pmatrix}h_{1j_1}\\ \\ \\ h_{n-m\ j_1}\end{pmatrix}\qquad\cdots\ +\ \alpha_{j_w}\begin{pmatrix}h_{1j_w}\\ \vdots\\ \vdots\\ h_{n-m\ j_w}\end{pmatrix}=\overline{0}$$

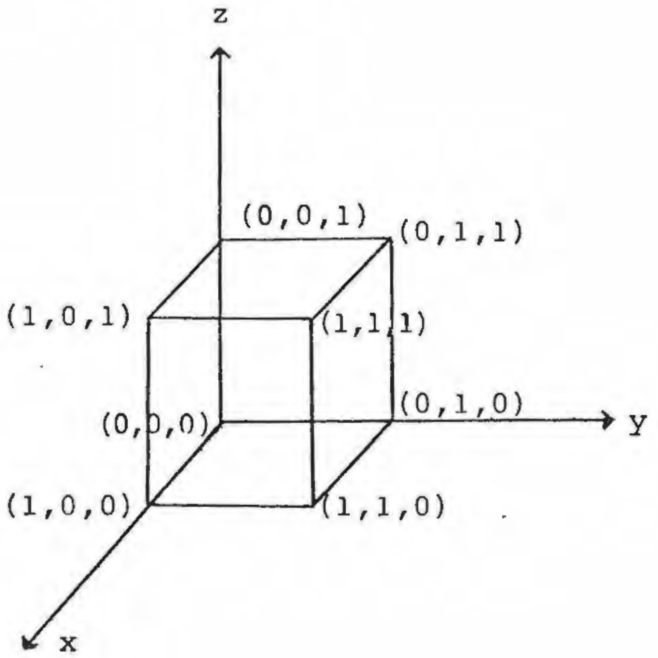
Since none of the αj_k's is zero $(1\leq k\leq w)$, they are all 1's. Complete the set $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ by setting all the remaining α_k's to zero. Then if $\overline{x}=(\alpha_1\ \ldots\ \alpha_n)$, then clearly $\overline{x}$ has weight w, and $H\overline{x}^t=\overline{0}$, i.e. $\overline{x}\in V$. Q.E.D.

An immediate result of Theorem 1.9 is

Corollary 1.10: Let V be an (m,n) code, which is the null space of an (n-m)×n matrix H. Then V has minimum distance w if and only if every set of w-1 or fewer column vectors of H is linearly dependent.

The "Packing Problem": An (m,n) code is t-error correcting if it corrects every error whose error vector has weight t or less. A t-error correcting code is perfect if it corrects no errors whose error vectors have weight greater than t. In other words, if an (m,n) code is perfect, there is a positive integer t such that the decoding function will always correctly decode any received vector $\overline{x}$ which has t or fewer errors in its digits, but will always fail to decode correctly any received vector $\overline{y}$ with more than t errors in its digits. Perfect codes are, in general, extremely difficult to construct. In the next chapter we will describe the Hamming codes, which are perfect single-error correcting codes.

We may conveniently interpret E^n geometrically as the set of all vertices of an n-dimensional hypercube. (For example E^3 is the set of all vertices of a normal three dimensional cube).



Suppose $V \subset E^n$ is an (m,n) code. Then the decoding function for V decodes each r element of E^n onto an element of V, $\bar{x}$. Thus we may partition E^n into equivalence classes, each containing one element. We write the equivalence class containing the code vector $\bar{c}$ as $[\bar{c}]$. Now suppose a code word $\bar{c}$ is corrupted in transmission to $\bar{r}$, such that t errors occur. Then $d(\bar{c},\bar{r}) = t$. Thus if V is t-error correcting, then each equivalence class $[\bar{c}]$ $(\bar{c} \in V$, contains the closed sphere

$S_t(\bar{c}) = \{\bar{x} \in E^n : d(x,c) \leq t\}$. We note that s errors can occur in a transmitted code vector $\bar{c}$ in C_s^n ways. Hence the number of elements of $S_t(\bar{c})$ is equal to

$$\sum_{s=0}^{t} C_s^n \quad . \qquad *$$

We note further that, if V is a perfect t-error correcting code, then for each $\bar{c} \in V$, $[\bar{c}] = S_t(\bar{c})$.

If V is an (m,n) group code, it contains 2^m elements, while E^n contains 2^n elements.

Hence $2^m \sum_{s=0}^{t} C_s^n = 2^n$ i.e

$$\sum_{s=0}^{t} C_s^n = 2^{n-m}$$ is a necessary condition for an

(m,n) code to be perfect. Thus we need to find those integers w for which

$$\sum_{t=0}^{w} C_s^n$$ is a power of 2, and for which

$1 < w < \frac{h-1}{2}$. Clearly, this is not an easy problem.

Actually, no general method is known of constructing a binary code which is optimal in the sense that it maximises the probability of correcting random errors. To do this requires the packing of as many non-overlapping spheres as possible into the hypercube E^n.

* In this text, C_r^n is the binomial coefficient, also written $\binom{n}{r}$.

Generalization to q-ary codes: Most of the ideas described above can be generalised to q-ary codes, where q is any power of a prime integer. The code digits are taken from the field GF(q). We will generally let q be a prime, in which case GF(q) will be the residue class field modular q. The Hamming weight of a code vector $\bar{c}$, written $(\bar{c})$, will be defined as the number of non-zero digits that the code vector contains. The q-ary (m,n) code will be thought of as a subspace of the n-dimensional vector space over GF(q). With this definition, we may define the distance between two code vectors $\bar{a}$ and $\bar{b}$, as $w(\bar{a} - \bar{b})$. It is now easy to verify that results 1.1 to 1.10 are valid for q-ary codes. It is not so easy to describe the error correcting capabilities of q-ary codes as it is in the binary case. These will be examined for certain codes later.

CHAPTER 2

SOME IMPORTANT BINARY ERROR-CORRECTING CODES

Hamming Codes

Suppose we wish to construct a perfect single-error correcting binary code V. From the section on the "packing problem" in the previous chapter, if V is to be an (m,n) code, we require

$$1 + n = 2^{n-m}. \qquad (1)$$

Let us put n-m = k. Then, solving (1) for n and m, we have

$$n = 2^k - 1. \qquad (2)$$

$$m = 2^k - 1 - k. \qquad (3)$$

We shall show that, for any positive integer k, there exists a perfect single-error correcting binary (m,n) code, where n and m satisfy equations (2) and (3). These codes are known as the Hamming codes.

Hamming codes are best defined by their parity check matrices. Given k, we construct the parity-check matrix H as follows:

1. Calculate n and m from equations (2) and (3).
2. Construct an (n-m) × n matrix H by putting the jth column equal to j, expressed in binary $(1 \leq j \leq n)$.
3. In each code vector $\overline{c} = (c_1\ c_2\ \ldots\ c_n)$, use $c_1, c_2, c_{2^2} \ldots c_{2^{\ell-1}} \ldots C_{2^{k-1}}$ as check digits, and use the remaining n-k digits for the information to be transmitted.
4. Encode a vector $\overline{b} \in E^m$, as follows. Place the m digits of $\overline{b}$ in the appropriate positions in $\overline{c} \in E^n$, as determined in 3. Form the equation $H\overline{c}^t = \overline{0}$. This gives k equations in the k parity-check digits of $\overline{c}$, from which these may be found.
5. Decode a received vector F as follows. Calculate $H\overline{r}^t$.

If $H\overline{r}^t = \overline{0}$, then the received vector is correct. If $H\overline{t}^t \neq 0$, then it is equal to one of the column vectors, say the jth one, of H. We then assume that there is a single error in the jth digit of $\overline{r}$, which is changed (from a 0 to a 1 or vice versa).

Example: Put k=3. Then $n=2^3-1=7$ and $m=2^3-1-3=4$. The parity-check matrix for the (4,7) Hamming code is

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

Suppose we wish to transmit $\overline{b} = (1\ 0\ 1\ 1)$. This is encoded to $\overline{c} = (c_1 c_2\, 1\, c_4\, 0\, 1\, 1)$. Putting $H\overline{c}^t = 0$, we obtain the equations

$$c_1 = 0$$
$$c_2 = 1$$
$$c_4 = 0$$

Thus $\overline{c} = (0\ 1\ 1\ 0\ 0\ 1\ 1)$.

Now suppose the vector $\overline{r} = (1\ 1\ 0\ 1\ 0\ 1\ 1)$ is received.

$$H\overline{r}^t = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

This is the 6th column vector of H. We thus assume that an error has occurred in the 6th digit and we decode $\overline{r}$ to (1 1 0 1 0 0 1).

We claim that the procedure outlined above corrects all single errors. Clearly, any correctly transmitted code vector $\overline{c}$ will be left unchanged by the error correction procedure. Suppose a single error occurs in the received word $\overline{r}$, in the jth digit. Then let $\overline{r} = \overline{c} + \overline{e}$, where $\overline{c}$ is the originally transmitted code word, and $\overline{e}$ is the error vector. Then

$$
\begin{aligned}
H\bar{r}^t &= H\bar{c}^t + he^t \\
&= H\bar{e}^t.
\end{aligned}
$$

Now all the digits of $\bar{e}$ are zero except for the jth one, which is a 1. Thus $H\bar{e}^t$ is equal to the jth column vector of H, and consequently our error correcting procedure identifies the correct position of the error. Now suppose 2 or more errors occur. Since our error correcting procedure corrects one digit or leaves the received word unchanged, it will always fail in this case. Thus the Hamming codes are perfect single-error correcting codes.

Cyclic Codes

An (m,n) code V is a cyclic code if, for every code vector $(a_o \ldots a_{n-1}) \epsilon V$, $(a_{n-1}, a_o \ldots a_{n-2})$ is also an element of V. As many important classes of error correcting codes are cyclic, cyclic codes are studied in detail. To study these codes, it is convenient to define an algebra structure on E^n. We do this as follows:-

Let F be a field, and f(x) a polynomial in $f[x]$. It is well known that the quotient ring $F[x]/(f(x))$ is a commutative vector algebra over F. Furthermore, if f is of degree n, it may be shown that the residue classes $\{1\}$, $\{x\} \ldots \{x^{n-1}\}$ (modulo f(x)) form a base for $F[x]/(f(x))$ which is thus of dimension n.

We will identify E^n with the algebra $F[x]/(x^n - 1)$ (where $F = GF(q)$ by the relationship

$$(a_o a_1 \ldots a_{n-1}) \leftrightarrow g(x) \text{ where}$$

$$g(x) = \sum_{k=o}^{n-1} a_k x^k$$

Since each polynomial residue class modulo $x^n - 1$ contains exactly one polynomial of degree < n, this relationship describes a 1 - 1 correspondence between elements of E^n and elements of $F[x]/(x^n - 1)$.

Group codes can now be described as vector subspaces of the algebra $F[x]/(x^n - 1)$. Cyclic codes can now be described by the following result:

Theorem 2.1: An (m,n) code V is cyclic if and only if it is an ideal of the algebra $F[x]/(x^n-1)$.

Proof: Suppose V is an ideal of $F[x]/(x^n-1)$. Then if $\{a_o + a_1 x + \ldots + a_{n-1}x^{n-1}\}$ is an element of V,

$$
\begin{aligned}
& x(a_o + a_1x + \quad \dots \quad + a_{n-1}x^{n-1}) \\
&= a_ox + a_1x^2 \quad \dots \quad + a_{n-1}x^n \\
&= a_{n-1} + a_0x \quad \dots \quad + a_{n-1}(x^n - 1) \\
&\equiv a_{n-1} + a_ox + \quad \dots \quad + a_{n-2}x^{n-1} (\text{mod } x^n - 1)
\end{aligned}
$$

Hence $\{a_{n-1} + a_ox + \quad \dots \quad + a_{n-2}x^{n-1}\}$

$= \{x\} \{a_o + a_1x \quad \dots \quad + a_{n-1}x^{n-1}\}$

ϵ V, since V is an ideal. Since V is an ideal, it is a subspace of $F[x]/(x^n-1)$.

Conversely, let V be a cyclic code. Then, as before,

$$
\begin{aligned}
& \{x\} \{a_o + a_1x + a_2x^2 \dots \quad + a_{n-1}x^{n-1}\} \\
&= \{a_{n-1} + a_ox + a_1x^2 \quad \dots \quad + a_{n-2}x^{n-1}\}
\end{aligned}
$$

We have established that multiplication by $\{x\}$ constitutes a cyclic shift of the code digits to the right. Inductively, multiplication by $\{x^k\} = \{x\}^k$ constitutes k cyclic shifts to the right, hence if $\{a_o + a_1x \ \dots \ + a_{n-1}x^{n-1}\} \epsilon$ V, then so is $\{x^k\}\{a_o + a_1x \ \dots \ + a_{n-1}x^{n-1}\}$. Now if $\{f(x)\}$ is any code vector (where deg $f < n$), let $\{g(x)\} = \{b_o + b_1x \dots$ $\dots + b_{n-1}x^{n-1}\}$ be any other code vector. Then

$$
\begin{aligned}
& \{g(x)\}\{f(x)\} = \{b_o + b_1x \dots + b_{n-1}x^{n-1}\}\{f(x)\} \\
&= b_o\{f(x)\} + b_1\{x\}\{f(x)\} \quad \dots \quad + b_{n-1}\{x^{n-1}\}\{f(x)\}
\end{aligned}
$$

Now we have that $\{x^k\}\{f(x)\} \epsilon V$ $(o \leq k < n)$ and the above sum is a linear combination of terms of this form, and is hence in V, since V is a subspace.

Since V is a subspace it is a subgroup of the additive group of $F[x]/(x^{n-1})$. Hence the theorem is proved. Q.E.D.

We now proceed to describe the ideals of $F[x]/(x^n-1)$.

Theorem 2.2: Let F be any field, f(x) a polynomial over F. Let I be a non-trivial ideal of $F[x]/(f(x))$. Then, there exists a polynomial $g(x) \in F[x]$ with $0 \leq \deg g < n$, such that

(i) $\{g(x)\} \in I$

(ii) $g(x) | f(x)$

(iii) $\{h(x)\} \in I$, $0 \leq \deg h < n$ if and only if $g(x)|h(x)$.

Proof: Let g(x) be a non-zero polynomial of smallest degree such that $\{g(x)\} \in I$. Then $\deg g \neq 0$, since $I \neq F[x]/(f(x))$. By the polynomial division algorithm,there exist polynomials q(x) and r(x) such that

$$f(x) = g(x)q(x) + r(x),\ 0 \leq \deg r \leq \deg g.$$

Hence

$$\{f(x)\} = \{g(x)\}\{q(x)\} + \{r(x)\},\ \text{i.e.}$$
$$\{0\} = \{g(x)\}\{q(x)\} + \{r(x)\},\ \text{whence } \{r(x)\} \in I.$$

Now we had $\deg r < \deg g$, which contradicts the definition of g(x) unless $\deg r = 0$. This implies $r = 0$, since otherwise the ideal would contain a unit, and hence be equal to the entire ring. Hence

$$g(x)q(x) = f(x),\ \text{i.e. } g(x)|f(x).$$

Now suppose that $\{h(x)\}$ $(0 \leq \deg h < n)$ is an element of I. Then by the polynomial division algorithm, there exist polynomials s(x) and t(x) of $F[x]$ such that

$$h(x) = g(x)\ s(x) + t(x),\ 0 \leq \deg t < \deg g$$

i.e. $\{h(x)\} = \{g(x)\}\{s(x)\} + \{t(x)\}$,

whence $\{t(x)\} \in I$, which implies $\deg t = 0$, by definition of g(x). But then $t(x) \equiv 0$, otherwise $\{t(x)\}$ is a unit, which would imply I is equal to the whole ring. Hence the theorem is proved. Q.E.D.

Thus the set of all cyclic codes of length n may be described by the set of all polynomials of $F[x]$ which divide $x^n - 1$. If I and g(x) are as in Theorem 2.2, then g(x) is known as the generating polynomial of I. The next result determines the number of information digits of a cyclic code.

Theorem 2.3: Let F be a field and f(x) a polynomial of degree n in F[x]. Let I be a nontrivial ideal of F[x]/(f(x)), with generating polynomial g(x), such that f(x) = g(x)h(x), where h(x) ϵ F[x]. Suppose deg h = m. Then I has dimension m over F.

Proof: Consider

$B = \{\{g(x)\}, \{x\ g(x)\} \ \dots \ \{x^{m-1}g(x)\}\}.$

This set is linearly independent for suppose there are scalars $\lambda_0 \dots \lambda_{m-1}$ ϵ F such that

$$\sum_{i=0}^{m-1} \lambda_i\{x^i g(x)\} \qquad = \quad 0$$

$$\therefore \{ \sum_{i=0}^{m-1} \lambda_i x^i g(x)\} \qquad = \quad 0$$

Now $\sum \lambda_i x^i g(x)$ is a polynomial of degree less than n. Thus the above linear combination is not zero, unless all the λ_i's are zero, since $g(x) \not\equiv 0$. Thus B is linearly independent. Clearly, I $\supseteq$ Span B. Now let {s(x)} ϵ I, where deg s < n. Then, by definition of g(x), $g(x)|s(x)$, i.e. s(x) = g(x)k(x) for some k(x) ϵ F[x]. Now deg k $\leq$ m (otherwise deg s > n) So let

$$k(x) = \sum_{i=0}^{m} a_i x^i$$

Then $s(x) = g(x)k(x) = g(x) \sum_{i=0}^{m} a_i x^i$

$$= \sum_{i=0}^{m} a_i x^i g(x)$$

Whence $\{s(x)\} = \sum_{i=0}^{m} a_i\{x^i g(x)\}$, i.e. {s(x)} is a linear combination of elements of B. This proves the theorem. Q.E.D.

Example: In the binary case

$$1 - x^7 = (1 - x)(1 + x + x^3)(1 + x^2 + x^3) \text{ over GF } (q^2).$$

Consider the code generated by $g(x) = 1 + x + x^3$. This will be a (4,7) code by the previous theorem.

$$\begin{array}{llccccccc} \text{Now} & \{g(x)\} = & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ & \{xg(x)\} = & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ & \{x^2g(x)\} = & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ & \{x^3g(x)\} = & 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{array}$$

Since these vectors form a base for the code, a generator matrix is

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

We can also describe a code in terms of the roots of the generator polynomial $g(x)$, possibly in an extension field of $GF(q)$.

Let V be a cyclic (m,n) code with generator polynomial $g(x)$. Let $\alpha_1 \ldots \alpha_k$ be the roots of $g(x)$. Then $\{f(x)\} \in V$ if and only if $f(\alpha_i) = 0$ $(1 \leq i \leq k)$.

Conversely, let $\alpha_1 \ldots \alpha_r$ be elements of some finite extension field of $GF(q)$. Then define a code V by the statement that $\{f(x)\}$ is a code vector if and only if $f(\alpha_i) = 0$ $(1 \leq i \leq r)$.

We calculate the length of this code as follows. Let α_i have order ℓ_i $(1 \leq i \leq r)$. Then each α_i must be a root of the generating polynomial $g(x)$, and since $g(x)|x^n - 1$, also of $x^n - 1$. Hence each ℓ_i must be a divisor of n, and we can set

$$n = \text{LCM}\ (\ell_1 \ldots \ell_n)$$

Now let $m_i(x)$ be the irreducible monic polynomial over $GF(q)$ which has α_i as a root. Then

$$m_i(x)|x^n - 1 \quad (1 \leq i \leq r)$$

Let $g(x) = \text{LCM}(m_1(x) \ldots m_r(x))$. Clearly, $g(x)|x^n - 1$ and V is the cyclic code generated by $g(x)$.

Matrix representation of cyclic codes

Let V be an (m,n) cyclic code, described by $\{f(x)\} \in V$ if and only if $\alpha_1 \ldots \alpha_r$ are roots of $f(x)$. Then we can construct a parity-check matrix H (with entries from $GF(q^k)$) for V in the following manner.

Let $f(x) = \sum_{i=0}^{n-1} a_i x^i \quad (a_i \in GF(q))$

Then $f(\alpha_i) = 0 \Rightarrow \sum_{i=0}^{h-1} a_i\alpha^i = 0$. We thus see that V is the nullspace of the matrix

$$H = \begin{bmatrix} 1 & \alpha_1\alpha_1^2 & \ldots & \alpha_1^{n-1} \\ 1 & \alpha_2\alpha_2^2 & \ldots & \alpha^{n-1} \\ 1 & & & \\ 1 & & & \\ 1 & \alpha_r\alpha_r^2 & \ldots & \alpha_r^{n-1} \end{bmatrix}$$

This representation will be useful for later work on cyclic codes.

Performance of cyclic codes

Theorem 2.4: If V is a binary cyclic code of length n generated by $g(x)$ and $g(x)$ does not divide $x^k - 1$ for $k < n$, then V has minimum distance 3.

Proof: Suppose this is not true. The minimum distance of V is the minimum Hamming weight of a non-zero member of V. Suppose there exists a member of V with weight 2. Then this element of V has form $\{e(x)\}$, where $w(e(x)) = 2$, i.e. $e(x) = x^j + x^k$ for positive integer $j \neq k$. Now suppose $j < k$. Then

$$e(x) = x^j(1 + x^{k-j}).$$

Now since $\{e(x)\} \varepsilon V$, $g(x)|e(x)$. But $g(x)|x^n - 1$ implies the constant term of $g(x)$ is non-zero, i.e. it must be 1. Whence $g(x)|1 + x^{k-j}$. But $0 < k - j < n$, which contradicts $g(x)$ does not divide $x^r - 1$ from $r < n$ (noting $x^t - 1 = 1 + x^r$ in the binary case). Thus there is no code vector of weight 2.

Suppose there is a code vector of weight 1. Then this has form $\{e(x)\}$, where $e(x) = x^j$ for some $j \geq 0$. Now $g(x)|x^j$. But $g_o = 1$ implies $j = 0$ i.e. $e(x) = 1$, which implies $g(x) = 1$, which is a contradiction. Q.E.D.

CHAPTER 3

THE BOSE-CHAUDHURI-HOCQUENGHEM CODES

The codes described in this section are among the most powerful error-correcting codes known. They were discovered separately by Bose and Chaudhuri,and by Hocquenghem, about 1960. Most other codes developed since them are appreciably behind them in performance. They are a class of cyclic codes and are best described as such.

We will first describe the generalised q-ary Bose-Chaudhuri-Hocquenghem (BCH) codes. Let α be a non-zero element of an extension field of GF(q). Consider the set

$M = \{\alpha^{m_o}, \alpha^{m_o+1}, \alpha^{m_o+2} \; \ldots \; \alpha^{m_o+d-2}\}$ where m_o and d are positive integers. Let V be the cyclic code for which $\{f(x)\}$ is a code vector if and only if all elements of M are roots of f(x). A BCH code is any code which is defined in this way. The length n of the code is equal to the LCM of the orders of the roots. Let e be the order of α. We must have

$$(\alpha^{m_o})^n = 1 \qquad \text{and} \qquad (\alpha^{m_o+1})^n = 1$$

$$\therefore \quad \alpha^{m_o n} = \alpha^{m_o n+n} = 1$$

$$\therefore \quad 1 = \alpha^n, \text{ hence } e|n.$$

Conversely, if $\alpha^e = 1, (\alpha^j)^e = 1$ and we have that n is the LCM of the orders of the $\alpha^j (0 \leq j \leq d-2)$. Consequently, $n|e$, and hence n=e. The number of information symbols is found by the methods of Chapter 2. The following result from matrix algebra is necessary to calculate the minimum distance of BCH codes.

Lemma 3.1: Let F be a field, $x_1, x_2 \ldots x_s \in F$.

Then

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_s \\ x_1^2 & x_2^2 & \dots & x_s^2 \\ x_1^{s-1} & x_2^{s-1} & \dots & x_s^{s-1} \end{vmatrix} = \prod_{i>j} (x_i - x_j)$$

A determinant of this form is known as a Van der Monde determinant.

Theorem 3.2: The BCH code satisfying {f(x)} is a code vector if and only if every element of

$M = \{\alpha^{m_o}, \alpha^{m_o+1} \dots \alpha^{m_o+d-2}\}$ has minimum distance at least d.

Proof: Using the methods of Chapter 2, the BCH code described in the nullspace of the matrix

$$H = \begin{bmatrix} 1 & \alpha^{m_o} & (\alpha^{m_o})^2 & \dots & (\alpha^{m_o})^{n-1} \\ 1 & \alpha^{m_o+1} & & \dots & (\alpha^{m_o+1})^{n-1} \\ 1 & & & & \\ 1 & & & & \\ 1 & \alpha^{m_o+d-2} & (\alpha^{m_o+d-2})^2 & \dots & (\alpha^{m_o+d-2})^{n-1} \end{bmatrix}$$

We will show that any d-1 column vectors of this matrix are linearly independent.

Construct a (d-1) × (d-1) matrix of any d-1 column vectors of H. Consider its determinant

$$\begin{vmatrix} (\alpha^{m_o})^{j_1} & (\alpha^{m_o})^{j_2} & (\alpha^{m_o})^{j_{d-1}} \\ (\alpha^{m_o+1})^{j_1} & (\alpha^{m_o+1})^{j_2} & (\alpha^{m_o+1})^{j_{d-1}} \\ \vdots & \vdots & \\ (\alpha^{m_o+d-2})^{j_1} & (\alpha^{m_o+d-2})^{j_2} & (\alpha^{m_o+d-2})^{j_{d-1}} \end{vmatrix}$$

$$= \alpha^{m_o(j_1+j_2\ +\ \dots\ +\ j_{d-1})} \begin{vmatrix} 1 & 1 & \dots & 1 \\ \alpha^{j_1} & \alpha^{j_2} & \dots & \alpha^{j_{d-1}} \\ (\alpha^{j_1})^2 & (\alpha^{j_2})^2 & \dots & (\alpha^{j_{d-1}})^2 \\ \vdots & & & \\ (\alpha^{j_1})^{d-1} & & & (\alpha^{j_{d-1}})^{d-1} \end{vmatrix}$$

If we let $x_i = \alpha^{j_i}$ $(1 \leq i \leq d-1)$, the determinant has the form of that in Lemma 3.1, and hence is equal to

$$\prod_{i>k} (\alpha^{j_i} - \alpha^{j_k}).$$

$\alpha^{j_i} \neq \alpha^{j_k}$ for distinct columns of H and hence the determinant is non-zero. By Corollary 1.10, the Code has minimum distance d, as required. Q.E.D.

We now consider the binary BCH codes. Let α be a primitive element of $GF(2^m)$ for some integer m.

Consider the binary BCH code defined by $\{f(x)\}$ is a code vector if and only if every element of

$M = \{\alpha\ \alpha^2\ \dots\ \alpha^{2t}\}$ for some positive integer t.

This is a BCH code with $m_o = 0$ and $d = 2t + 1$. The length of the code is equal to the order of α, which is $2^m - 1$. Let $m_i(x)$ denote the unique irreducible polynomial with coefficients from GF(2) having α^j as a root $(1 \leq j \leq 2t)$. Then the generating polynomial g(x) of the code is given by

$$g(x) = \mathrm{LCM}(m_1(x)\ \dots\ m_{2t}(x)) \qquad (1)$$

But from the theory of finite fields, if β is a root of f(x) in a field of characteristic 2, so are β^2, β^4, etc. Hence $m_1(\alpha) = m_1(\alpha^2) = 0$, $m_3(\alpha^3) = m_3(\alpha^6) = \beta$ etc. Thus we have that $m_{2j}(x) = m_j(x)$ $0 \leq j \leq t$. Thus we may eliminate $m_j(x)$ where j is even on the right hand side of (1). We now have

$$g(x) = \mathrm{LCM}(m_1(x),\ m_3(x),\ m_5(x)\ \dots\ m_{2t-1}(x)) \qquad (2)$$

Since $GF(2^m)$ is of degree m over GF(2), deg $m_j(x) \leq m$. There are t polynomials $m_i(x)$ on the right hand side of (2), whence deg $g \leq mt$. This is the number of parity-check symbols in the code.

The code described has minimum distance 2t + 1, and hence, by Theorem 1.2, can correct t errors.

We have thus proved:

Theorem 3.3: For each positive integer m, and each positive integer t, there is a binary BCH code of length $2^m - 1$ which corrects t errors and has a maximum of mt parity-check symbols.

The Reed-Solomon Codes: We will consider a class of codes here that were known before the BCH codes, and show that they are in fact a special case of the BCH codes.

We consider a q-ary BCH code where α is a primitive element of GF(q). Consider the BCH code defined by {f(x)} is a code vector if and only if every element of

$$\{\alpha, \alpha^2, \alpha^3 \ldots \alpha^{d-1}\}$$

is a root of F(x). This defines a BCH code with minimum distance d. The generating polynomial is

$$g(x) = (x - \alpha)(x - \alpha^2) \ldots (x - \alpha^{d-1}).$$

Since deg g = d - 1, this code has d - 1 parity-check symbols and length n = q - 1, the order of α.

Now the Reed-Solomon codes are constructed as follows. Let α be a primitive element of GF(q). Then let $a_0 \ldots a_{m-1}$ be the m - 1 information symbols to be transmitted. Define

$$F(x) = a_0 + a_1x \ldots + a_{m-1}x^{m-1}.$$

Encode to

$$(F(1), F(\alpha), F(\alpha^2) \ldots F(\alpha^{q-2})) \ (d > m)$$

Now consider the polynomial $f(x) = \sum_{j=0}^{q-2} F(\alpha^j)x^j$.

We will show that $\alpha, \alpha^2 \ldots {}^{d-1}$ are roots of $f(x)$

$$f(x) = \sum_{j=0}^{d-2} F(\alpha^j)x^j$$

$$= \sum_{j=0}^{q-2} \left(\sum_{k=0}^{m-1} \alpha_k \alpha^{jk} \right) x^j$$

$$= \sum_{j=0}^{q-2} \sum_{k=0}^{m-1} a_k (\alpha^k x)^j$$

$$= \sum_{k=0}^{m-1} a_k \sum_{j=0}^{q-2} (\alpha^k x)^j \qquad Ⓐ$$

We note that every non-zero element of GF(q) is a zero of

$$1 - x^{q-1} = (1 - x)\ (1 + x + \ldots + x^{q-2})$$

Hence every non-zero element of GF(q) except 1 is a zero of $1 + x + \ldots + x^{q-2}$.

Put $\phi(x) = 1 + x + \ldots .\ x^{q-2}$

$$\phi(1) = q - 2$$
$$\equiv - 1 \pmod{q}$$

Ⓐ becomes $f(x) = \sum_{k=0}^{m-1} a_k \phi(\alpha^k x)^j$

$$f(1) = \sum_{k=0}^{m-1} a_k \phi(\alpha^k) = a_0 \phi(1) = - a_0$$

$$f(\alpha^{-1}) = \sum_{k=0}^{m-1} a_k \phi(\alpha^{k-1}) = - a_1$$

$$f(\alpha^{-j}) \qquad\qquad = - a_j$$

$$f(\alpha^{-(m-1)}) \qquad\qquad = - a_{m-1}$$

This yields a decoding procedure for the code.

For $m - 1 < j \leq 1 - 2$, $f(\alpha^{-j}) = 0$

But $\alpha^{-j} = \alpha^{q-1-j}$ $(m-1 < j \leq q - 2)$

$\therefore$ $f(\alpha^{q-1-j}) = 0$ $(m-1 < j \leq q - 2)$

Hence all the elements of the set

$\{\alpha,\alpha^2 \ldots \alpha^{q-m-1}\}$ are roots of $f(x)$. Thus the code described is a BCH code with length $q - 1$, minimum distance $q - m$ and $q - m - 1$ check digits.

CHAPTER 4

THE DECODING PROBLEM

One of the most important problems in coding theory is the construction of decoding algorithms which may be easily implemented by communications systems. It is undesirable to use large amounts of computer or machine store space and time to encode and decode messages. Thus the decoding process should be as fast as possible, without sacrificing the error-correcting ability of the code. At the same time, the decoding device must be as simple as possible, or it may itself become a source of errors.

Some useful concepts

Let V be a q-ary (m,n) block code. We showed in Chapter 1 that V could be considered to be an m-dimensional subspace of the n-dimensional vector space over GF(q). V can also be expressed as the row space of an $m \times n$ encoding matrix E over GF(q) or equivalently as the nullspace of an $(n - m) \times n$ parity-check matrix H over GF(q).

We note that the row space V' of H is an (n - m, n) q-ary code. Furthermore we have:

Proposition 4.1: Let V be an (m,n) q-ary block code, with encoding matrix E and parity check matrix H. Then the nullspace of E is equal to the row-space of H.

Proof: Let V' denote the nullspace of E and V" the row-space of H. If $\overline{h}_k = (h_{k1} \dots h_{kn})$ is any column vector of H and $(e_{j1} \dots e_{jn})$ is a column vector of E, we have, since $(e_{j1} \dots e_{jn}) \in V$ that

$$\sum_{i=1}^{n} e_{ji}h_{ik} = 0$$

Hence $E(h_k)^t = 0$ and since the column vectors of H span V", we have

$$V" \subseteq V'.$$

To establish equality, we show dim V" = dim V'. By the method of constructing H used in chapter 1, we note that the column vectors of H are linearly independent, and hence dim V" = n - m. Since the now space of E has dimension m, its nullspace V' has dimension n - m.

Hence dim V' = dim V" and the result follows. Q.E.D.

The nullspace V' of the encoding matrix E is known as the dual code of V. Thus every (m,n) q-ary code V is associated with an (n - m, n) q-ary dual code V'. It is clear that the dual code of V' is V. The concept of dual codes will be useful in constructing decoding algorithms.

Majority-Logic Decoding

We now construct a method of decoding that is extremely easy to implement and which is due to Massey (1963).

Let V be an (m,n) q-ary block code. Then a parity check A_i is a sum of the form $a_{i1}x_1 + \ldots + a_{in}x_n$, which is zero for every code vector $(x_1 \ldots x_n)$ in V, and where $a_{1i} \ldots a_{ni}$ are fixed elements of GF(q). A_i is said to check the code digit x_k if $a_{ik} \neq 0$. Clearly if A_i does not check x_k, the value of the left hand side of the equation $A_i = 0$ will be unaffected by the value of x_k.

A set of parity checks $\{A_1 \ldots A_r\}$ on a code V is said to be orthogonal on the kth digit if

(i) every one of the A_i checks the kth digit $(1 \leq i \leq r)$

(ii) every other digit of the code is checked by at most one of the A_i.

Clearly, if a set of parity checks is orthogonal on the kth digit, we can construct a set of parity checks orthogonal on the kth digit whose kth coefficients are all 1. Thus, without loss of generality, we will assume that the kth coefficients are 1's.

Theorem 4.2: (Reed-Massey). If a linear code has at least d - 1 parity checks orthogonal on each digit, then the code has minimum distance at least d.

Proof: We will show that every non-zero code vector has weight at least d. Suppose $\bar{c}$ is a non-zero code vector, and that the kth digit is non-zero. Let $A_1 \ldots A_{d-1}$ be parity checks on the kth digit. Then each of them must check at least one other non-zero digit, and since $A_1 \ldots A_{d-1}$ are orthogonal on the kth digit these digits are distinct. Hence there are at least d non-zero digits in $\bar{c}$. Q.E.D.

Note that if V is a cyclic code and $A = \sum_{k=1}^{n} a_k x_k$ is a parity check on V, which checks the jth digit, then

$(x_1 \ldots x_n)$ a code vector implies $(x_n x_1 \ldots x_{n-1})$

is also a code vector, hence

$$a_1 x_n + a_2 x_1 \ldots + a_n x_{n-1} = 0$$

$$\therefore \quad a_2 x_1 \ldots + a_n x_{n-1} + a_1 x_1 = 0.$$

Thus we have a parity check which checks the j-1th digit, obtained by a cyclic shift of the coefficients of A to the left. Hence the following result follows immediately from Theorem 4.2.

Corollary 4.3: If it is possible to construct a set of d - 1 parity checks orthogonal on any digit of a cyclic code, then the code has minimum distance d.

The existence of J parity checks orthogonal on each digit of a code V gives a method of decoding. To decode the kth digit of a received vector F, apply the d - 1 parity checks to that digit. Let r_k denote the kth received digit and let $r_k = c_k + e_k$, where c_k is the transmitted code digit and e_k is the error digit. We calculate c_k by giving it that value which occurs most often when the J parity checks orthogonal on the kth digit are applied to the received vector, or zero if no majority occurs. This process is known as <u>single step majority-logic decoding</u>. This is an effective method of decoding, as the following theorem illustrates.

<u>Theorem 4.4</u>: Suppose V is a q-ary code with J parity checks orthogonal on each digit. Then single step majority decoding will correct every combination of J/2 or fewer errors.

<u>Proof</u>: Let $\bar{r}$ be the received code vector and put $\bar{r} = \bar{c} + \bar{e}$, where $\bar{c}$ is the transmitted code vector and $\bar{e}$ is the error vector. Since any parity check applied to $\bar{c}$ yields zero, and by linearity of parity checks, we obtain the same result if we apply a parity check to $\bar{e}$ as when we apply it to $\bar{r}$.

Suppose $A_1 \ldots A_J$ are orthogonal on the jth digit. Suppose The kth error digit is zero. Then since $A_1 \ldots A_J$ are orthogonal on the kth digit, the J/2 or power non-zero error digits are checked at most once each, and hence $A_i\bar{r} \neq 0$ for at most J/2 error digits. In this case correct decoding will occur.

Now suppose the kth error digit is non-zero. Then, if all the other digits checked by any one of the A_i are zero, then $A_i\bar{r} = e_k$, the kth error digit. But since at most J/2 - 1 of the remaining error digits are non-zero, and each is checked at most once, it is clear that at least J/2 + 1 of the parity checks are unaffected, which is a clear majority.

Hence correct decoding occurs in this case also. This proves the theorem. Q.E.D.

Majority-logic decoding is extremely easy to implement in a communications system, but may not decode with the maximum efficiency for a given code. The theoretical error connecting capability of a code given by Theorem 1.2 may not be reached. We now prove a result that sets an upper bound on the number of errors that may be connected by single step majority decoding.

Theorem 4.5: Let $\bar{d}$ denote the minimum distance of the dual code of an (m,n) code V. Then the number of errors that can be corrected by single step majority decoding, t_1 is bounded by

$$t_1 \leq \frac{n-1}{2(\bar{d}-1)}$$

Proof: Note firstly that if $a_1x_1 + \ldots + a_nx_n$ is a parity check on V, then $(a_1 \ldots a_n)$ is an element of the dual code of V. Hence it has at least $\bar{d}$ of the a_i's non-zero. Suppose a set of parity checks is orthogonal on the kth digit. Each parity check checks at least $\bar{d} - 1$ other digits, and no other digit is checked more than once. Hence there are at most $(n - 1) / (\bar{d} - 1)$ parity checks. By the previous theorem, this means that at most half this number of errors may be corrected. Hence

$$t_1 \leq \frac{n-1}{2(\bar{d}-1)}, \text{ as required.}$$ Q.E.D.

Suppose a code has minimum distance d. Then by Theorem 1.2, the number of errors t it can correct is given by

$$t = \frac{d-1}{2}$$

By theorem 4.4, for $\frac{d-1}{2}$ errors to be corrected by single step majority logic decoding, it is sufficient that there be d - 1 parity checks orthogonal on each digit. Codes which satisfy this condition are called completely orthogonalizable in one step.

Unfortunately, many important codes cannot be efficiently decoded in this manner. For example, if a Reed-Solomon code has minimum distance d, its dual code can be shown to be a Reed-Solomon code with minimum distance n - d + 2, which is a large number. Massey showed that all codes with 3 or fewer information digits are completely orthogonalizable in one step, but such codes are of very little practical interest. He also showed this to be true of the binary (7,15) BCH codes. It is not, however true of the binary BCH codes as a class. To date no significant condition for codes to be completely orthogonalizable has been discovered.

Systematic Codes: A systematic (m,n) code is one in which the first m symbols are used as information symbols, and the rest as parity-check symbols. The generating matrix of a systematic code may be expressed in the form

$[I_m:P]$, where I_m is the m × m identity matrix, and P is an m × (n-m) matrix. Systematic codes are particularly useful for majority loqic decodinc, since only the information symbols need be decoded.

We define two block codes V_1 and V_2 to be equivalent if there is a distance-preservinq bijection between them. Clearly, equivalent codes have the same error-correctinq capabilities, number of information symbols, etc. That systematic codes are a useful class to study is illustrated by the following result.

Theorem 4.6: Every (m,n) linear code V is equivalent to a systematic code.

Proof: Let V be generated by the m × n matrix $M = (m_{ij})$. We note that the usual operations on rows of M (addition of rows, multiplication of a row by a non-zero scalar, and transposition of rows) do not affect the row space of M. Furthermore, transposition of any two columns of M results in a matrix whose row space is a code equivalent to V, as may be easily verified. From the theory of matrices, it is possible to reduce M to a matrix of the form $[I_m:A]$, whose row space V is a systematic code. In view of the precedinq remarks, V' is equivalent to V. Q.E.D.

he next result shows a useful relationship between the generatinq matrix and parity-check matrix of a binary systematic code.

Theorem 4.7: Let $M = [I_m:P]$ be the generating matrix of a binary systematic code V. Then

$H = [P^T:I_{n-m}]$ is a parity-check matrix for V.

Proof: We have

$$M = \left[\begin{array}{c|ccc} & p_{11} & \ldots & p_{1\ n-m} \\ I_m & & & \\ & p_{m1} & \ldots & p_{m\ n-m} \end{array}\right]$$

$$H = \left[\begin{array}{ccc|c} p_{11} & \ldots & p_{m1} & \\ & & & I_{n-m} \\ p_{1\ n-m} & \ldots & p_{m\ n-m} & \end{array}\right]$$

Let $\overline{m}^j$ be the jth row vector of M. Then

$$\overline{m}^j = (\underbrace{0 \ldots \underset{\text{jth position}}{1} \ldots 0}_{\text{m digits}}\ p_{j1} \ldots p_{n\ n-m})$$

The kth row vector $\overline{h}^k$ of H is

$$\overline{h}^k = (p_{1k} \quad \ldots \quad p_{mk} \quad \underbrace{0 \quad \ldots \quad \underset{\text{(k+m)th posn}}{1} \quad \ldots \quad 0}_{\text{n-m digits}})$$

The kth co-ordinate of $H(\overline{m}^j)^T$ is

$$\overline{m}^j . \overline{h}^k = p_{jk} + p_{jk} = 0$$

Thus $H(\overline{m}^j)^T = \overline{0}$ and hence

$$H(\overline{c})^T = \overline{0} \text{ for every } \overline{c} \in V.$$

Hence H is a parity-check matrix for V. Q.E.D.

L-step majority decoding

As we have noted, Theorem 4.5 imposes a severe restriction on the usefulness of single-step majority-logic decoding. We now discuss a generalisation of this procedure which is very much more powerful than single-step decoding. We define a set of parity checks to be orthogonal on a set of digits.

$A = \{i_1 \ldots i_k\}$ of an (m,n) code if

(i) For each $i_j \in A$, the i_j-th term has the same coefficient in all of the parity checks.

(ii) No other term has a non-zero coefficient in more than one of the parity checks.

Example: The set

$$x_1 + 2x_2 \quad \ldots + x_5$$
$$x_1 + 2x_2 + x_3 \quad \ldots$$
$$x_1 + 2x_2 \quad \ldots + x_4$$

is orthogonal on the first and second positions.

Suppose an (m,n) code has minimum distance d, and that there are d-1 parity checks orthogonal on some combination A of digits of the code. If $\frac{d-1}{2}$ or fewer errors occur in the transmission, then it may be argued by the method of Theorem 4.4 that the combination of error digits described by A will be correctly estimated by majority logic.

Now, if it is possible to find several sets of positions $A_1, A_2 \ldots A_p$, and sets of d-1 parity checks orthogonal on each of the A_i. If $\frac{d-1}{2}$ or fewer errors occur, then each of the combinations of error digits described by the A_i will be an additional parity check. If we can select parity checks orthogonal on other sets of digits and estimate these and repeat the process L times until we obtain parity checks orthogonal on each single digit, then we say the code is L-step orthogonalizable.

Example: Consider the (4;7) Hamming code with parity-check matrix

$$H = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} = [P:I]$$

Note that this matrix differs from that constructed for a Hamming Code in Chapter 2, by a permutation of columns. Consequently the code described is equivalent to that constructed by the method previously discussed. Furthermore, the code discussed now is a systematic one. Since the Hamming codes are perfect single-error correcting, they have minimum distance d = 3. Let the error vector in a received word be

$$\bar{e} = (e_1, e_2 \ldots e_7)$$

If 1 or no errors occur, then since the second and third rows of H form d - 1 = 2 parity checks orthogonal on the second and third positions $e_2 + e_3$ can be correctly estimated from them by majority logic. Similarly, $e_3 + e_4$ can be deduced from the first and second rows of H.

Let $(e_2 + e_3)^*$ and $(e_3 + e_4)^*$ be the estimates of these sums obtained in this way. (Which will be correct if 1 or no errors occur). Then if we add $(e_2 + e_3)^*$ to the result of the parity check defined by the second column of H and $(e_3 + e_4)^*$ to that defined by the third column, we obtain the parity checks.

$$\begin{matrix} x_1 + & \ldots & x_4 & \\ x_1 & \ldots & & + x_7 \end{matrix}$$

which are orthogonal on the first digit. The other 3 information digits may be estimated in the same way. Thus this code is 2-step orthogonalizable.

This example is generalised in the following Theorem.

Theorem 4.8: The $(2^m - m - 1, 2^m - 1)$ Hamming code may be completely orthogonalised in m - 1 steps, for all positive integers $m \geq 2$.

For m = 2, it is trival to show that the Hamming (1,3) code can be 1-step orthogonalised.

Proof: We have just shown that, for m = 3, the (4,7) Hamming code is completely orthogonalizable in 2 steps. We now proceed by induction on m.

Let H be the parity-check matrix of the $(2^m - m - 1, 2^m - 1)$ code, expressed in the form $[P:I]$, as before. Suppose the first position is non-zero in an odd number of the rows of P, i.e. the first column has odd weight. Let $\bar{r}_i$ be the i-th row vector of P, and $\bar{\gamma}_i$ the sum of the remaining rows. Suppose both $\bar{r}_i$ and $\bar{\gamma}_i$ have a 1 in the k-th position. Clearly, a necessary and sufficient condition for this to hold is that the k-th column of P has a 1 in the i-th row and is of even weight. Thus $\bar{r}_i$ and $\bar{\gamma}_i$ form a set of d - 1 = 2 parity checks orthogonal on these positions. We can form m such pairs of parity checks (1 from each column of P). These can be used to eliminate all 1's from columns of even weight in p. The only non-zero columns of P after this process are those which had odd weight. There are

$$2^{m-1} - m = 2^{m-1} - (m-1) - 1 \text{ of these.}$$

If we now remove all the zero columns and the last row of P, we obtain the matrix P' corresponding to the $(2^{m-1} - (m-1) - 1, 2^{m-1} - 1)$ Hamming code. The first position is still checked, since it was in a column of odd weight.

Alternatively, suppose the first column of P is of even weight. We again construct m pairs of parity-checks orthogonal on sums of digits as above. These checks contain exactly those information noise bits in columns of even weight. These sums correspond to a matrix made up of all the columns of P of even weight. Omit any row of this matrix (unless only two check the first position, in which case omit a row which does not check the first position). The remaining m - 1 columns form a matrix P' corresponding to the $(2^{m-1} - (m-1) - 1, 2^{m-1} - 1)$ Hamming code.

We repeat this process m - 3 times, to arrive at the (4,7) Hamming code, and the first digit is still checked. Since this code is 2-step orthogonalizable, e_1 can be calculated with m - 1 levels of majority logic. A similar argument applies to the rest of the information digits, and the theorem is proved.

Conclusion: Majority logic decoding has been found to be applicable to several important classes of codes. All BCH codes of length 15 or less have been found to be L-step orthogonalizable. Rudolph (1964), in a thesis to the University of Oklahoma, described a class of codes based on projective and Euclidean geometries that can be efficiently decoded by L-step majority logic. Since the complexity of decoders increases exponentially with the number of levels of majority logic required, the cost of equipment becomes prohibitively high when L is large, and other methods of decoding must be devised.

CHAPTER 5

TOPOLOGICAL CODES

Saltzer (1968) has described a class of binary linear codes based on the notions of algebraic topology. Although these codes are relatively inefficient in terms of the ratio of the number information symbols to the code length, they have the advantage that they can be decoded by majority logic methods. In particular, we will consider a subclass which can be decoded by single-step majority logic.

Preliminary topological concepts

We will begin with a finite set of objects

$\{x_0^0, x_1^0 \; \ldots \; x_{r_0-1}^0\}$ of objects known as vertices.

A p-simplex is a set of p + 1 vertices, which we will denote

$$x_i^p = (x_{i_0}^0, x_{i_1}^0 \; \ldots \; x_{i_p}^0).$$

Geometrically, a 1-simplex can be thought of a a line segment, a 2-simplex as a triangle, etc. A non-empty subset of a simplex containing q + 1 vertices ($q < p$) of the simplex is called a q-face of the simplex.

A simplicial complex K is a family of simplexes such that, if K contains a given simplex x_i^p, then it also contains all its faces.

We will use r_p to denote the number of p-simplexes in a complex. A p-chain y^p is a formal sum of the form

$$y^p = \sum_{i=0}^{r_p-1} \alpha_i \; x_i^p$$

where the α_i's are elements of some field. For our purposes, the field GF(2) will be used.

Addition and scalar multiplication of p-chains are defined in the obvious way:-

$$y_1^p = \sum_{i=0}^{r_p-1} \alpha_{i1}^p x_i^p$$

$$y_2^p = \sum_{i=o}^{r_p-1} \alpha_{i2} x_i^p$$

$$y_1^p + y_2^p = \sum_{i=0}^{r_p-1} (\alpha_{i1} + \alpha_{i2}) x_i^p$$

$$\gamma y_1^p = \sum_{i=0}^{r_p-1} \gamma\alpha_{i1} x_i^p$$

Clearly, these definitions make the set of p-cycles of a complex into a vector space V^p of dimension r_p over GF(2), with the set of all p-simplexes of the complex as a base.

Boundary and coboundary operators

Let the p-simplex $x_j^p = (x_{j_0}^0, x_{j_1}^0 \dots x_{j_p}^0)$ and define the boundary operator

$$\rho : V^p \rightarrow V^{p-1}$$

by

$$\rho(x_{j_0}^0 \quad \dots \quad x_{j_p}^0) = \sum_{q=0}^{p} (x_{j_0}^0 \quad \dots \quad x_{j_{q-1}}, x_{j_{q+1}}, \quad x_{j_p}^0)$$

Thus a p-simplex is mapped by the boundary operator onto the sum of its (p-1) faces. Conventionally, V^{-1} is taken as the space {0}.

Let Z^p denote the kernel of ρ. The elements of Z^p are known as p-cycles.

If the boundary operator is applied twice to a simplex x_j^p, we obtain a sum containing all the (p-2) faces of the simplex, where each (p-2) face appears exactly twice. Hence

$$\rho^2 \, y^p = 0.$$

This implies that ρV^{p+1} is a subspace of Z^p. The elements of ρV^{p+1} are called bounding cycles of V^p.

The boundary operator may also be described by the p-th incidence matrix $(a_{ij}{}^p)$, where

$$a^p_{ij} = \begin{cases} 1 \text{ if } x_j{}^{p-1} \text{ is a face of } x^p_i \\ 0 \text{ if not} \end{cases}$$

for $i = 0, \ldots, (r_p - 1)$ $j = 0 \ldots (r_{p-1} - 1)$.

Then

$$\rho x^p_i = \sum_{j=0}^{r_{p-1}-1} a^p_{ij} \; x^{p-1}_j$$

The co-boundary operator $\delta : V^{p-1} \to V^p$ is a linear operator defined by

$$\delta x^{p-1}_j = \sum_{i=0}^{r_p-1} a_{ij} \; x^p_i$$

Then we have that the co-boundary of a p-1 simplex consists of the sum of those p-simplexes whose boundaries contain the given (p-1)-simplex. We denote the kernel of δ in V^p by $\overline{Z}^p$ and the elements of $\overline{Z}^p$ are known as p-co-cycles.

Quasi-direct sums

We define the inner product of two vectors in V^p,

$$y^p_1 = \sum_{i=0}^{r_{p-1}} \alpha_{i1} \; x^p_i \quad \text{and} \quad y^p_2 = \sum_{i=0}^{r_{p-1}} \alpha_{i2} \; x^p_i$$

by $(xy^p_i, y^p_2) = \sum_{i=0}^{r_p-1} \alpha_{i1} \; \alpha_{i2}$.

This definition of inner product will obey all the usual laws for inner products. In particular, we can represent a vector $y^p \in V^p$ by its "Fourier Series"

$$y^p = \sum_{i=0}^{r_{p-1}} (y^p, x_i^p)\ x_i^p$$

We note that

$$(\rho x_i^p,\ x_k^{p-1}) = (\sum_{j=0}^{r_{p-1}-1} a_{ij}^p\ x_j^{p-1},\ x_k^{p-1}) = a_{ik}^p$$

$$(x_i^p,\ \delta x_k^{p-1}) = (x_i^p,\ \sum_{j=0}^{r_p-1} a_{jk}^p\ x_j^p) = a_{ik}^p$$

from which it follows that for any p-chain y_1^p and any p-1 chain y^{p-1} that

$$(\rho y_1^p,\ y_2^{p-1}) = (y_1^p,\ \delta y_2^{p-1}),$$

whence $(y_1^{p+1},\ \delta^2 y_2^{p-1}) = (\rho y_1^{p+1},\ \delta y_2^{p-1})$

$$= (\rho^2 y_1^{p+1},\ y_2^{p-1})$$

$$= (0,\ y_2^{p-1}) = 0$$

Since this is true for every (p+1) chain y_1^{p+1}, it follows that $\delta^2 y_2^{p-1} = 0$.

Thus δV^{p+1} is a subspace of $\overline{Z}^p$.

We define orthogonality of vectors in the usual way, i.e.

y_1^p is orthogonal to y_2^p if and only if

$(y_1^p,\ y_2^p) = 0.$

Note that a non-zero vector may be orthogonal to itself, since we are dealing with residue arithmetic.

Let W be an n-dimensional vector space over a field F. Let U and V be subspaces of W. We say that W is the quasi-direct product of U and V, written

$W \sim U \oplus V$ if:-

(i) Every vector of U is orthogonal to every vector of V.

(ii) dim W = dim U + dim V.

Note that $U \cap V$ is not necessarily the zero vector.

We will now show the following hold in V^p:-

(a) $V^p \sim Z^p \oplus \delta V^{p-1}$

(b) $V^p \sim \overline{Z}^p \oplus \rho V^{p+1}$

Let $y_1^p \in Z^p$, $y_2^{p-1} \in V^{p-1}$

$$\begin{aligned}(y_1^p, \delta y_2^{p-1}) &= (\rho y_1^p, y_2^{p-1}) \\ &= (0, y_2^{p-1}) = 0.\end{aligned}$$

Conversely, suppose that

$$(y_3^p, \delta y_2^{p-1}) = 0 \qquad \forall\ y_2^{p-1} \in V^{p-1}$$

$$\therefore\ (\rho y_3^p, y_2^{p-1}) = 0 \qquad \text{"}$$

$$\therefore\ \rho y_3^p = 0$$ and hence y_3^p is a p-cycle.

Hence $V^p \sim Z^p \oplus \delta V^{p-1}$. The proof that $V^p \sim \overline{Z}^p \oplus \rho V^{p+1}$ is similar.

Simplicial codes. From the foregoing considerations we can associate four binary codes with a given simplical complex K and a given value of p.

The p-cycle code Z^p, the p-cocycle code $\overline{Z}^p$, the p-boundary code, ρV^{p+1}, and the p-coboundary code δV^{p-1}. From the previous section, it is seen that Z^p and

δV^{p-1} are dual codes, as are $\overline{Z}^p$ and ρV^{p-1}. We will firstly consider the case that the complex K is a simplex. In this case, it is known that

$$Z^p = \rho V^{p+1}$$
$$\overline{Z}^p = \delta V^{p-1}.$$

Thus there are only two codes to consider for each p, the p-cycle codes and the p-cocycle codes. In this case, the codes are known as simplicial codes. The following results show that these codes are completely orthogonalizable in one step. Let K be an m-simplex in the following two theorems.

Theorem 5.1: The p-cycle simplicial code Z^p is completely orthogonalizable in one step and has minimum distance p + 2.

Proof: In view of Theorem 4.2, it is sufficient to exhibit a code word of weight p+2, and for each digit to construct a set of p+1 parity checks orthogonal on that digit.

For the first part, let x_j^{p+1} be any (p+1) simplex. Since a(p+1)-simplex has p+2 (p+1)-faces, the weight of $\rho\ x_j^{p+1}$ is p+2, and since $\rho x_j^{p+1} \in Z^p$, we have an element of Z^p of weight p+2, as required.

Now consider the code word $(\alpha_0, \alpha_1\ \ldots\ \alpha_{n-1})$. This corresponds to the chain

$$y^p = \alpha_0\ x_0^p + \alpha_1\ x_1^p + \ \ldots\ + \alpha_{n-1}\ x_{n-1}^p.$$

Let the vertices of x_k^p be denoted by $x_{j_0}^0\ \ldots\ x_{j_p}^0$. Denote the vertices not in x_k^p by $x_{j_{p+1}}^0\ \ldots\ x_{j_m}^0$. Now let the (p-1)-faces of x_k^p be denoted by $x_0^{p-1}, x_1^{p-1}\ \ldots\ x_p^{p-1}$ be denoted by $x_0^{p-1}, x_1^{p-1}\ \ldots\ x_p^{p-1}$, where

$$x_q^{p-1} = (x_{j_0}^0\ \ldots\ x_{j_{q-1}}^0\quad x_{j_{q+1}}^0\ \ldots\ x_{j_p}^0).$$

The coboundary of x_q^{p-1} is obtained by adjoining one vertex of $(x_0^0\ \ldots\ x_m^0)$ which is not in x_q^{p-1}, i.e.

$$\delta x_q^{p-1} = (x_{j_o}\ \ldots\ x_{j_p})$$

$$+ \sum_{w=0}^{m-p-1} (x_{j_0}^0\ \ldots\ x_{j_{q-1}}^0 \ldots x_{q+1}^0 \ldots x_{j_p},\ x_{j_{p+1+w}}^0).$$

By the orthogonality relations previously established, this will be a parity check for Z^p, and there are p+1 such checks, obtained by lettering q=0...p. Furthermore, any two simplexes in the last term which correspond to distinct values of q and κ are distinct, and hence the sum described is orthogonal on the k-th digit. Q.E.D.

Theorem 5.2: The p-cocycle simplicial code has minimum distance (m-p+1) and is completely orthogonalizable in one step.

Proof: We exhibit a code vector of weight m-p+1 and a set of m-p parity checks orthogonal on each digit.

Let x_k^{p-1} be any (p-1) face of the m-simplex. Then δx_k^{p-1} is and element of the p-cocycle code. But there are m-(p-1) = m-p+1 vertices of the complex not in x_k^{p-1}. Hence x_k^{p-1} is a p-1 face of m-p+1 p-simplexes and has weight m-p+1.

Let $(\alpha_0 \ldots \alpha_{n-1})$ be a code vector of the p-cocycle code. This word corresponds to the p-chain

$$y^p = \sum_{i=0}^{n-1} \alpha_i \; x_i^p$$

Again, consider the k-th position. Let the vertices of x_k^p again be given by $x_k^p = (x_{j_0}^0 \ldots x_{j_p}^0)$ and the remaining vertices by $x_{j_{p+1}} \ldots x_{j_m}$.

Put $x_q^{p+1} = (x_{j_0} \ldots x_{j_p}, x_{j_{p+q+1}})$ q = 0 ... m-p-1.

$$\text{Now } \rho x_q^{p+1} = (x_{j_0} \ldots x_{j_p})$$

$$+ \sum_{w=0}^{p} (x_{j_0}^e \ldots x_{j_{w-1}}^0, x_{j_{w+1}}^0 \ldots x_{j_p}^0, x_{j_{p+q+1}}^0)$$

Again, this is a parity check by the previous section, and by letter q range from 0 to m-p-1, we obtain m-p parity checks orthogonal on the k-th position. Q.E.D.

The Euler-Poincaré Formula and Code Efficiency

We now consider an arbitrary simplicial complex K and the codes associated with it.

Let H^p denote the quasi-orthogonal complement of ρ^{p+1} in Z^p and $\overline{H}^p$ the quasi-orthogonal complement of δV^{p-1} in $\overline{Z}^p$, i.e.

$$Z^p \sim \rho V^{p+1} \oplus H^p \qquad \textcircled{1}$$

$$\overline{Z}^p \sim \delta V^{p-1} \oplus \overline{H}^p \qquad \textcircled{2}$$

Then we have

$$V^p \sim \delta V^{p-1} \oplus \rho V^{p+1} + H^p \qquad \textcircled{3}$$

$$V^p \sim \delta V^{p-1} \oplus \overline{H}^p + \rho V^{p+1} \qquad \textcircled{4}$$

where the quasi-direct sum of three spaces is defined similarly to that of two spaces. From the above formulae, clearly

$\dim H^p = \dim \overline{H}^p$. We denote this number B_p, the <u>p-th Betti number</u> of the complex. Let

$$\sigma_p = \dim Z^p \text{ and } \overline{\sigma}_p = \dim \overline{Z}^p.$$

From $\textcircled{1}$ $\sigma_p = \dim \rho V^{p+1} + B_p$ $\qquad \textcircled{5}$

But $\dim \rho V^{p+1} = \dim V^{p+1} - \dim Z^{p+1}$ from the theory of operators

$$\therefore \quad \sigma p - Bp = r_{p+1} - \sigma_{p+1}$$

$$\therefore \quad r_{p+1} = \sigma_p + \sigma_{p+1} - B_p \qquad p = 0, 1, 2 \ldots \qquad \textcircled{6}$$

Define $\mu_p = \sum_{j=0}^{p} (-1)^j r_j$

$$\therefore \quad \mu_p = r_0 + \sum_{j=1}^{p} (-1)^j r_j$$

$$= r_0 + \sum_{j=1}^{p} (-1)^j (\sigma_{j-1} + \sigma_j - B_{j-1}) \quad \text{by } \textcircled{5}$$

$$= r_0 - \sigma_0 + (-1)^p \sigma_p - \sum_{j=1}^{p} (-1)^j B_{j-1}$$

$$\therefore \sum_{j=0}^{p} (-1)^j r_j = r_0 - \sigma_0 + (-1)^p \sigma_p + \sum_{j=0}^{p-1} (-1)^j B_j$$

$$\therefore \sigma_p = (-1)^p \, [\sigma_o - r_0 + \sum_{j=0}^{p-1} (-1)^j (r_j - B_j)] + r_p$$

Now the boundary of a 0-chain is 0 and hence $\sigma_0 = r_0$, whence

$$\sigma_p = r_p + (-1)^p \sum_{j=0}^{p-1} (-1)^j (r_j - B_j) \qquad \textcircled{7}$$

This is the Euler-Poincaré formula.

By $\textcircled{2}$ $\bar{\sigma}_p = \dim \delta V^{p-1} + B_p$

and by $\textcircled{3}$, $r_p = B_p + \dim \rho V^{p+1} + \dim \delta V^{p-1}$

$$\therefore r_p - \bar{\sigma}_p = \dim \rho V^{p+1}$$

$$= \sigma_p - B_p \text{ by } \textcircled{5}$$

$$\therefore \bar{\sigma}_p = r_p + B_p - \sigma_p \qquad \textcircled{8}$$

This gives us the dimension of the p-cycle and p-cocycle codes in terms of the number of simplexes of dimension $\leq$ p and the Betti numbers of the complex.

Performance of simplicial codes.

We will now apply the above theory to the case where the complex K is an m-simplex and compute the number of information symbols for the p-cycle and p-cocycle codes in these cases.

Firstly, observe that in this case,

$r_p = C^{m+1}_{p+1}$, and hence both the p-cycle and the p-cocycle codes have length C^{m+1}_{p+1}.

It is known from the theory of simplicial complexes that for an m-simplex

$$B_0 = 1; \quad B_p = 0 \text{ for } 1 < p \leq m.$$

Applying the Euler-Poincaré formula,

$$\begin{aligned}
\sigma_p &= C^{m+1}_{p+1} + (-1)^p \sum_{j=0}^{p-1} (-1)^j (C^{m+1}_{j+1} - B_j) \\
&= C^{m+1}_{p+1} + (-1)^p \left[-1 + \sum_{j=0}^{p-1} (-1)^j \, C^{m+1}_{j+1}\right] \\
&= (-1)^p \left[-1 + \sum_{j=0}^{p} (-1)^j C^{m+1}_{j+1}\right] \\
&= (-1)^p \sum_{j=0}^{p+1} (-1)^{j-1} \, C^{m+1}_j \qquad \textcircled{9}
\end{aligned}$$

We claim that $\sigma_p = C^m_{p+1}$ $\qquad \textcircled{10}$

For p = 0, equation $\textcircled{9}$ becomes

$$\begin{aligned}
\sigma_p &= (-1)^0 \, (-C^{m+1}_0 + C^{m+1}_1) \\
&= -1 + (m+1) = m = C^m_1 \ .
\end{aligned}$$

Thus the assertion is true for p=0. We proceed by induction on p.

$$\sigma_{p+1} = (-1)^{p+1} \left(\sum_{j=0}^{p+2} (-1)^{j-1} C_j^{m+1} \right)$$

$$= (-1) \; (-1)^p \left(\sum_{j=0}^{p+1} (-1)^{j-1} C_j^{m+1} \right)$$

$$+ \; (-1)^{p+1} \; (-1)^{p+1} \; C_{p+2}^{m+1}$$

$$= -\sigma_p + C_{p+2}^{m+1}$$

$$= - \; C_{p+1}^{m} + C_{p+2}^{m+1}$$

$$= - \frac{m!}{(m-p-1)!(p+1)!} + \frac{(m+1)!}{(m-p-1)!(p+2)!}$$

$$= \frac{-m!(p+2) \; + \; (m+1)!}{(m-p-1)!(p+2)!}$$

$$= \frac{m!(-p-2 \; +m+1)}{(m-p-1)!(p+2)!}$$

$$= \frac{m!(m-p-1)}{(m-p-1)!(p+2)!}$$

$$= \frac{m!}{(m-p-2)!(p+2)!} \quad = C_{p+2}^{m}, \text{ as required.}$$

By equation 8, $\overline{\sigma}_p = r_p + B_p - \sigma_p$

$$= C_{p+1}^{m+1} - C_{p+1}^{m}$$

$$= \frac{(m+1)!}{(m-p)!(p+1)!} - \frac{m!}{(m-p-1)!(p+1)!}$$

$$= \frac{(m+1)! \; - \; m!(m-p)}{(m-p)!(p+1)!}$$

$$= \frac{m!(m+1-m+p)}{(m-p)!(p+1)!}$$

$$= \frac{m!}{(m-p)!p!} = C_p^m$$

Summary of results

We have shown that the p-cycle and p-cocycle codes associated with an m-simplex are codes completely orthogonalizable in one step with the following parameters:

	p-cycle code	p-cocycle code
Word length (n)	C_{p+1}^{m+1}	C_{p+1}^{m+1}
Number of information digits (k)	C_{p+1}^{m}	C_p^m
Minimum weight (d)	p+2	m+p-1
Information rate $(\frac{k}{n})$	$\frac{m-p}{m+1}$	$\frac{p+1}{m+1}$

Construction of simplicial codes

Since for the complex K an m-simplex it is true that $Z^p = \rho V^{p+1}$ and $\overline{Z}^p = \delta V^{p-1}$ $(1 \leq p \leq m - 1)$, we have associate with the m-simplex, the following exact sequences:

$$V^0 \overset{\rho}{\leftarrow} V^1 \overset{\rho}{\leftarrow} V^2 \quad \ldots \quad \leftarrow V^m$$

$$V^0 \overset{\delta}{\rightarrow} V^1 \rightarrow V^2 \rightarrow \ldots \quad \rightarrow V^m$$

We now consider as an example, the 3- simplex. We describe the 1-cycle and 1 cocycle codes.

By the formulae previously established these codes have the following parameters:

p-cycle code: $n = C_2^4 \qquad = 6$

$k = C_2^3 \qquad = 3$

$d = 1 + 2 \qquad = 3$

p-cocycle code: $n = C_2^4 \qquad = 6$

$k = C_1^3 \qquad = 3$

$d = 3 - 1 + 1 = 3$

Denote the 4 vertices of the complex by

$x_0^0, \quad x_1^0, \quad x_2^0, \quad x_3^0$

Denote the 6 1-faces by

$x_0^1 = (x_0^0, x_1^0) \qquad x_1^1 = (x_1^0, x_2^0)$

$x_2^1 = (x_2^0, x_3^0) \qquad x_3^1 = (x_0^0, x_2^0)$

$x_4^1 = (x_0^0, x_3^0) \qquad x_5^1 = (x_1^0, x_3^0)$

Denote the 4 2-faces by

$$x_0^2 = (x_0^0, x_1^0, x_2^0) \qquad x_1^2 = (x_1^0, x_2^0, x_3^0)$$

$$x_2^2 = (x_0^0, x_2^0, x_3^0) \qquad x_3^2 = (x_0^0, x_1^0, x_3^0)$$

The map $\rho\colon V^2 \to V^1$ maps the 2-faces as follows:

$$\begin{array}{llllllll} x_0^2 & \to & x_0^1 & + x_1^1 & & + x_3^1 & & \\ x_1^2 & \to & & x_1^1 & + x_2^1 & & & + x_5^1 \\ x_2^2 & \to & & & x_2^1 & + x_3^1 & + x_4^1 & \\ x_3^2 & \to & x_0^1 & & & & + x_4^1 & + x_5^1 \end{array}$$

Whence the 1-cycle code is the row space of the 2-incidence matrix

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Since the 1-cycle code has dimension 3, only 3 rows of this matrix are linearly independent. Consequently the generating matrix for this code may be given by deleting one of the rows, i.e.

$$M = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

To describe the 1-cocycle code, consider the boundary operator $\rho V^1 \to V^0$:

$$\begin{array}{llllll} x_0^1 & \to & x_0^0 & + x_1^0 & & \\ x_1^1 & \to & & x_1^0 & + x_2^0 & \\ x_2^1 & \to & & & x_2^0 & + x_3^0 \\ x_3^1 & \to & x_0^0 & & + x_2^0 & \\ x_4^1 & \to & x_0^0 & & & + x_3^0 \\ x_5^0 & \to & & x_1^0 & & + x_3^0 \end{array}$$

Thus the 1-incidence matrix is given by

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

From this matrix we note that the 1-cycle code is the subspace of V^1 spanned by the row space of

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

Now again only 3 of these vectors are linearly independent, so omitting the last row we obtain the generating matrix M^1 for the 1-cocycle code.

$$M' = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

We note that this is a parity-check matrix for the 1-cycle code and that the generating matrix for the 1-cycle code is a parity-check matrix for the 1-cocycle code.

Efficiency of simplical codes.

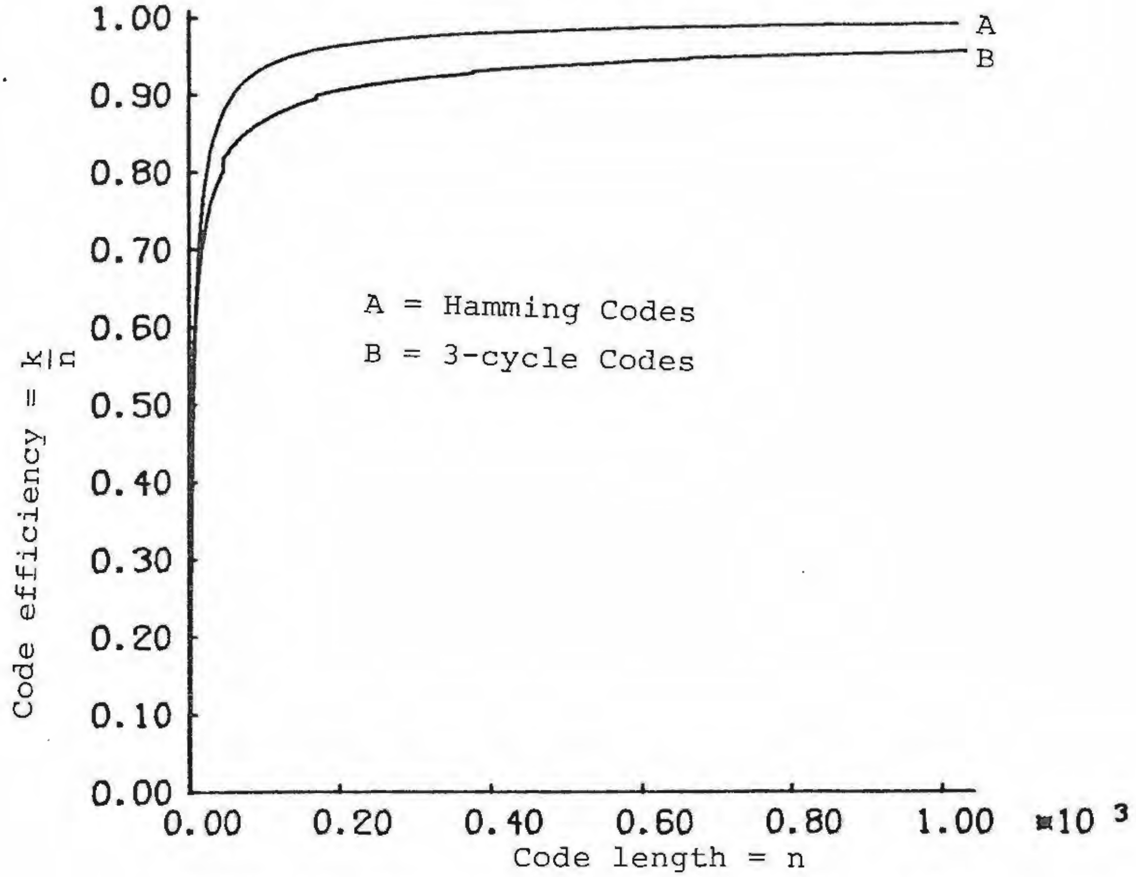
The p-cycle code of an m-simplex has an information rate (ratio of number of information digits to code length) of $\frac{m-p}{m+1}$. This approaches 1 as m approaches infinity, and hence codes of any required efficiency can be constructed by choosing m arbitrarily large. The minimum distance is given by $d = p + 2$, and hence the number of errors which may be corrected is given by the formula

$$p + 2 = 2t + 1 \qquad \text{(Theorem 1.2)}$$

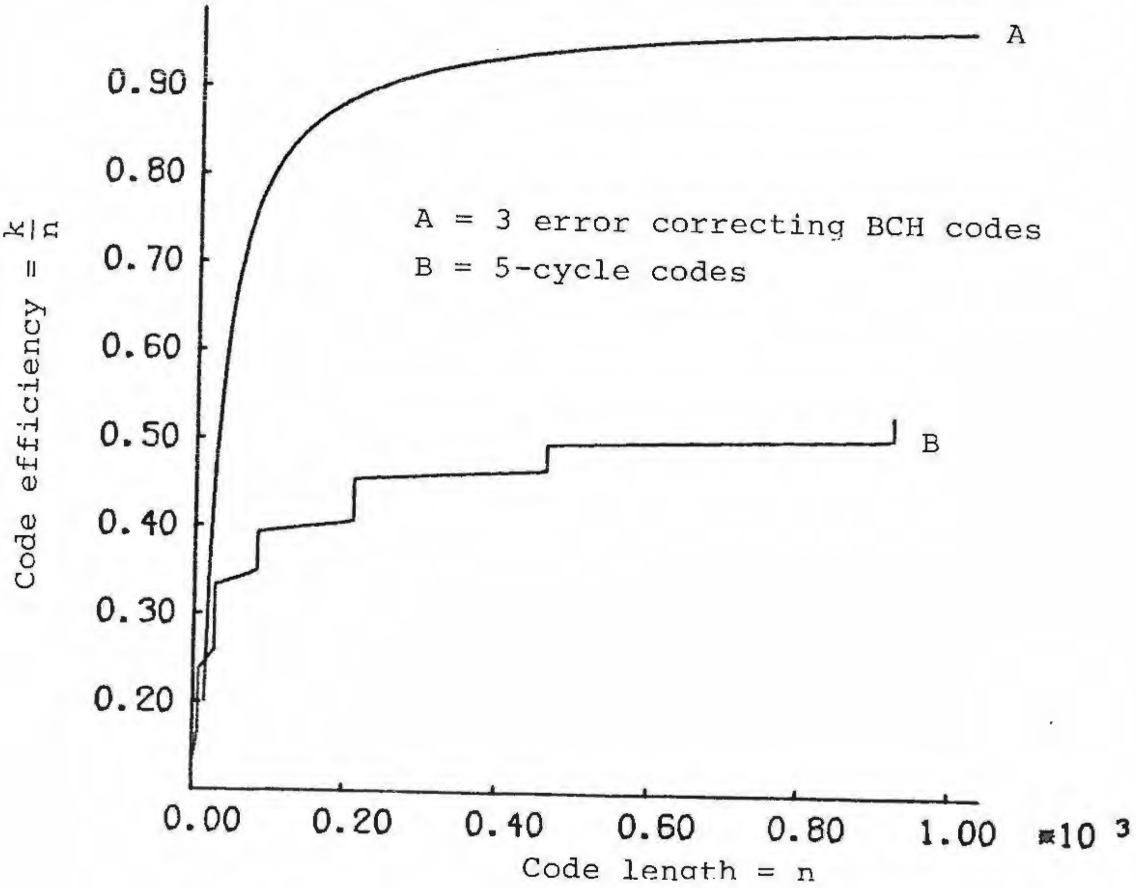
$$\therefore \quad t = \frac{p+1}{2} \text{ (or the largest integer} < \frac{p+1}{2})$$

Thus t is dependent on p only. If we set p=1, we obtain a code with minimum distance 3, which is then capable of correcting single errors. If we set m=2, p=1, we obtain a (1;3) code which consists of the code vectors (0 0 0) and (1 1 1). This trivial example is the Hamming (1;3) code. However, as Graph 1 shows, the 1-cycle codes are, in general, less efficient than the Hamming codes of equivalent length. They do however, have an advantage over the Hamming codes in that they can all be decoded with single-step majority logic, whereas the Hamming $(2^m - m - 1,\ 2^m - 1)$ code may require up to m-1 levels of majority logic.

The efficiency of the p-cocycle codes, which are the dual codes of the p-cycle codes, approaches zero as the code length becomes large. Consequently, these codes are of little practical use, except as parity-checks on the p-cycle codes.



Graph 1: Hamming Codes and 3-cycle codes



Graph 2: 3-Error Correcting BCH Codes and 5-cycle Codes

The t-error correcting simplicial codes are less efficient than the corresponding t-error correcting BCH codes, but again the decoding of BCH codes is a very much more complicated process than that of decoding the simplicial codes. In graph 2 we plot the graphs of efficiency of the 5-simplicial codes, which are 3-error correcting and the lower bound of efficiency of the 3-error correcting BCH codes against code length. (The lower bound for efficiency of the BCH codes is given in Theorem 3.3).

It will be noted that the actual code described by the procedure used in the example depends on the numbering of the vertices, but it is easy to show that the different codes obtained by different numberings of the faces will be equivalent. In some cases, it is possible to choose the numbering in such a way that the codes involved are cyclic. I have not been able to determine whether this is in fact true for all p-cycle codes for all values of m, nor am I aware of any literature in this regard.

CHAPTER 6

CONCLUSION

The study of error-correcting codes is now approximately 25 years old. The first known publication on the subject was in 1949 by M. Golay, who later did much research into the subject of perfect codes. It has been recently established that all the perfect codes are known.

R.W. Hamming presented his perfect single-error correcting codes in 1950, in an article in the Bell System Technical Journal. These codes turned out to be a special case of the powerful Bose-Chaudhuri codes which were discovered around 1960. Various work has been done on the theory of minimal redundancy of codes for a given error-correcting performance, by Plotkin, Gilbert, Varshamov and others, between 1950 and 1960. The binary BCH codes were found to be so close to the theoretical bounds that, to date, no better codes have been discovered.

Although the BCH codes are extremely efficient in terms of ratio of information to check digits, they are not easily, decoded with a minimal amount of apparatus. Petersen in 1961 described an algorithm for decoding BCH codes, but this was cumbersome compared with the majority-logic methods of Massey and others. Thus the search began for codes which are easily decoded with comparatively simple apparatus. The finite geometry codes which were described by Rudolph in a 1964 thesis were examples of codes which are easily decoded by a small number of steps of majority logic. The simplicial codes of Saltzer are even better in this respect, since they can be decoded by a single step of majority logic, but are rather inefficient.

The applications of coding theory have changed over the years, as well. The first computers were huge circuits of relays, which were unreliable and prone to errors. Error-correcting codes were required to minimise the possibility of incorrect results. As vacuum tubes and later transistorised circuits made computers more reliable, the need for sophisticated and powerful codes in the computer world diminished. Other used presented themselves however, for example the control systems of unmanned spacecraft. Because of the difficulty of sending and receiving messages in this case, very powerful codes were required. Other uses were found in transmission lines and telephone exchanges.

The codes considered in this dissertation have, for the most part, been block codes for use on the binary symmetric channel. There are, however, several other applications, such as codes for use on an erasure channel, where bits are corrupted so as to be unrecognizable, rather than changed. There are also codes for burst-error correction, where chennel noise is not randomly distributed, but occurs in "bursts" a few bits long. Certain cyclic codes are of application in these cases.

The theory of error correcting codes has risen from virtual non-existence in 1950 to a major and sophisticated part of communication theory. Judging from the articles in journals, it promises to be the subject of a great deal of research for some years to come.

References

Berlekamp, E.R.: "Algebraic coding theory", McGraw-Hill, 1968.

Birkhoff Garrett and Bartee, Thomas C.: "Modern Applied Algebra", McGraw-Hill, 1970.

Blake, Ian F. and Mullin, Ronald C.: "The Mathematical Theory of Coding", Academic Press, 1975.

Bose, R.C. and Ray-Chaudhuri, D.K.: "On a class of error-correcting binary group codes", Information and Control 3 (68), 1960.

Bose, R.C. and Chaudhuri, D.K.: "Further results on Error-Correcting Binary Group Codes", Information and Control 3 (279-290), 1960.

Chien, Robert T.: "A new proof of the BCH bound", IEEE Trans. on Information Theory, IT-18, (541), 1972.

Hamming, R.W.: "Error detecting and error correcting codes", Bell System Tech.J.,XXVI-2(147-160), 1950.

Massey, James L.: "Threshold Decoding", M.I.T.Press, 1963.

Petersen, W.W.: "Error-correcting Codes", M.I.T.Press, 1963.

Rudolph, L.D.: "A class of majority-logic decodable codes", IEEE Trans. on Information Theory, IT-13, (305-307), 1967.

Saltzer, Charles: "Topological Codes", Proceedings of Symposium on Error-Correcting Codes, Mathematics Research Centre, U.S. Army, University of Wisconsin, 1968.

Weldon, E.J., Jr.: "Some results on majority-logic decoding", Proceedings of Symposium on Error-Correcting Codes, Mathematics Research Centre, U.S. Army, University of Wisconsin, 1968.