9. Coding Theory - An Introduction

Coding theory is the study of methods in transferring information accurately from one location to another. Nowadays, we can not live without communications. How to transmit information over noisy channels from distant sources to expected sinks is an important issue to study.

There is a long list of applications for coding theory, such as communications of space spacecrafts, computers, networks, and the manufacture of compact discs. The readers may refer to [1] for a more detail presentation.

In a word, codes are studied by various scientific disciplines for the purpose of designing efficient and reliable data transmission methods. Mainly, there are four types of coding :

(1). Source coding (data compression),

(2). Channel coding (error control),

(3). Cryptographic coding (security control), and

(4). Line coding (data transport).

Since Discrete Mathematics are involved in channel coding, this chapter will put all the effort in the explanation of this part. For the rest of codings, only basic ideas will be introduced.

9.1 Basics of error-correcting codes (channel coding)

Throughout of this section, we let F_q be the finite field with q elements. Clearly, q is then a prime power. We also denote the set of n-tuples (vectors) with coordinates in F_q by F_q^n. Hence, $< F_q^n, \bigoplus, \bigodot >$ is a vector space where $\bigoplus$ is the vector addition and $\bigodot$ is the scalar multiplication of a scalar in F_q and a vector in F_q^n.

Definition 9.1.1 (q-ary code of length n)

A q-ary code C of length n is a set of vectors in F_q^n. If $q = 2$, then C is a binary code.

For convenience in presentation, we use (sometimes) $a_1a_2\cdots a_n$ to denote the vector $(a_1, a_2, \cdots, a_n)$. For example, $C_1 = \{000, 111\}$ is a binary code of length 3 and $C_2 = \{012, 120, 111, 222\}$ is a ternary ($q = 3$) code of length 3.

Definition 9.1.2 (Hamming distance)

Let $\vec{x} = (x_1, x_2, \cdots, x_n)$ and $\vec{y} = (y_1, y_2, \cdots, y_n)$ be two vectors in F_q^n. Then, the Hamming distance of $\vec{x}$ and $\vec{y}$, denoted by $d(\vec{x}, \vec{y})$ is the number of coordinates in which $\vec{x}$ and $\vec{y}$ differ.
That is, $d(\vec{x}, \vec{y}) = |\{i \mid x_i \neq y_i, i = 1, 2, \cdots, n\}|$.

For example, $d(000, 111) = 3$ and $d(012, 111) = 2$.

It is not difficult to check that the Hamming distance is indeed a metric.

Lemma 9.1.3

The Hamming distance satisfies

(a) $d(\vec{x},\vec{y}) = 0$ if and only if $\vec{x} = \vec{y}$;

(b) $d(\vec{x},\vec{y}) = d(\vec{y},\vec{x})$ for all $\vec{x},\vec{y} \in F_q^n$; and

(c) $d(\vec{x},\vec{y}) \leq d(\vec{x},\vec{z}) + d(\vec{z},\vec{y})$ for all $\vec{x},\vec{y},\vec{z} \in F_q^n$.

Definition 9.1.4 (Code distance)

Let C be a q-ary code of length n. Then, the code distance of C is defined as $\min\{d(\vec{c}_1,\vec{c}_2) \mid \vec{c}_1,\vec{c}_2 \in C \text{ and } \vec{c}_1 \neq \vec{c}_2\}$.

If we consider all codewords in C as the points in an n-dim. space, then $d(C)$ measures the shortest (Hamming) distance among these points. The importance of this notion can be seen from the following assumption.

First, we consider the binary codes. Therefore, the channel we use for communications is a binary channel. One of the simplest channels is the binary symmetric channel (B.S.C.). It has no memory and it receives and transmits two symbols, 0 and 1.

We assume that the B.S.C. has the property : a symbol is transmitted correctly with probability q, and with probability $p = 1 - q$ it will not be. Moreover, in the transmission of a codeword, we assume each symbol is transmitted independently. Without loss of generality, we let $p < \frac{1}{2}$ and p is called the error rate of transmitting a symbol.

Lemma 9.1.5

Let C be a binary code of length n. Let $d(\vec{x},\vec{y}) = d$ where $\vec{x},\vec{y} \in C$.

Then, the word error rate p_{err} for transmitting $\vec{x}$ and receiving $\vec{y}$ is $p_{err}(\vec{x}, \vec{y}) = q^{n-d} \cdot p^d = (1-p)^{n-d} \cdot p^d$.

Proof. It follows that d errors in symbols have occurred. □

Since p is less than $\frac{1}{2}$, $p_{err}(\vec{x}, \vec{y}) < p_{err}(\vec{x}, \vec{z})$ provided $d(\vec{x}, \vec{y}) < d(\vec{x}, \vec{z})$. This fact tells that receiving $\vec{z}$ is more probable than receiving $\vec{y}$. The idea used in this set up is generally referred as the maximum likelihood decoding (MLD).

For convenience, if $\vec{x}$ is transmitted, $\vec{y}$ is received and $d(\vec{x}, \vec{y}) = s$, then we say that s errors occurred in the transmission of $\vec{x}$. We shall say s errors occurred in the transmission of a code C if for each codeword in C, each transmission has at most s errors occurred.

Theorem 9.1.6

(a) A code C can detect up to s errors in any codeword if $d(C) \geq s+1$.
(b) A code C can correct up to t errors if $d(C) \geq 2t+1$.

Proof. The first part comes from the fact that s errors will not turn one codeword to another. Since the case $s > 0$ happens, the received word is not a codeword in C. The proof of second part follows from the fact that the received word is closer to one codeword than any others. □

Exercise 1.

Verify Lemma 9.1.3.

Exercise 2.

Use Lemma 9.1.3 to show the result obtained in Theorem 9.1.6(b).

Exercise 3.

Find a set $\tilde{e}$ of 16 words (binary) of length 7 such that $d(\tilde{C}) = 3$.

For binary codes, we can use $\vec{x} + \vec{y}$ to denote the error pattern whenever $\vec{x}$ is transmitted and $\vec{y}$ is received. In fact, the number of 1's in $\vec{x} + \vec{y}$ is equal to $d(\vec{x}, \vec{y})$. So, for convenience, we can use $wt(\vec{x})$ to denote the number of 1's in $\vec{x}$.

Definition 9.1.7 (Linear codes)

Let C be a linear subspace of the vector space $< F_q^n, \bigoplus, \bigodot >$. Then, C is a linear code. If C is of length n, distance d and dimension k, then C is referred to an $[n, k, d]_q$-code.

It is not difficult to see that an $[n, k, d]_q$-code has q^k codewords. Subsequently, if C is a binary code, then it is denoted by an $[n, k, d]$-code (omitting q).

Definition 9.1.8

If C is an $[n, k, d]$-code, then $\frac{k}{n}$ is defined as the code rate, i.e., each codeword contains k information bits (symbols).

On the other hand, if C is not a linear code and $|C| = M$, $d(C) = d$, then we refer C as an $(n, M, d)_q$-code. We also use (n, M, d)-code for

binary codes. Hence, an $[n, k, d]$-code is an $(n, 2^k, d)$-code. But, not every code is linear.

Fundamental problem of coding theory (q-ary code)

(a) Given n and d, determine the maximum M such that we have an $(n, M, d)_q$-code. M is denoted by $A_q(n, d)$.

(b) Given n and d, determine the maximum k such that we have an $[n, k, d]_q$-code. q^k is denoted by $A_q[n, d]$. The above problem is in general not solved except some special cases. The most well-known result is the following.

Theorem 9.1.9 (Sphere-Packing bound)

Let C be an $(n, M, d)_q$-code. Then,

$$M \leq \frac{q^n}{\sum_{i=0}^{t}(q-1)^i\binom{n}{i}}$$

where $t = \lfloor \frac{d}{2} \rfloor$.

Proof. Let $\vec{x}$ be any codeword of C. Then, $\sum_{i=0}^{t}(q-1)^i\binom{n}{i}$ measures the number of words $\vec{y}$ with $d(\vec{x}, \vec{y}) \leq t$. The set of $\vec{y}$'s satisfying $d(\vec{x}, \vec{y}) \leq t$ is denoted by $B_q(\vec{x}; t)$, i.e., the n-dim. ball centered at $\vec{x}$ with radius t. Since for any two codewords $\vec{x_1}$, $\vec{x_2}$ in C, $B_q(\vec{x_1}; t) \bigcap B_q(\vec{x_2}; t) = \varnothing$ provided $t = \lfloor \frac{d}{2} \rfloor$.
This implies that $M \cdot (\sum_{i=0}^{t}(q-1)^i\binom{n}{i}) \leq q^n$, the proof follow. □

Example 9.1.10

Let $n = 7$, $q = 2$, and $d = 3$. Then,

$$M \leq \frac{2^7}{\sum_{i=0}^{1} \binom{7}{i}} = \frac{2^7}{2^3} = 2^4.$$

Indeed, we can find a binary code of length 7 and distance 3 which has exactly 16 codewords. Hence, we have the equality.

Definition 9.1.11 (Perfect codes)

An $(n, M, d)_q$-code satisfying $M = \frac{q^n}{\sum_{i=0}^{t} (q-1)^i \binom{n}{i}}$ is called a perfect code.

We remark here that only finitely many perfect codes exist.[2] We shall present some of them later.

Definition 9.1.12 (Generator Matrices)

Let C be a linear code, an $[n, k, d]_q$-code. Let $\{\vec{g}_1, \vec{g}_2, \cdots, \vec{g}_k\}$ be a basis of C. Then, the following matrix is a generator matrix of C.

$$G = \begin{bmatrix} \\ \vec{g}_1 \\ \\ \vec{g}_2 \\ \\ \vdots \\ \\ \vec{g}_k \\ \\ \end{bmatrix}_{k \times n}$$

Therefore, C is the row space of G (over $GF(q)$).

Now, we consider binary codes. Clearly, G is then a $(0, 1)$-matrix. Moreover, C is a linear code with 2^k codewords.

Definition 9.1.13 (Dual codes)

Let C be an $[n, k, d]$-code. Then, $C^{\perp} = \{\vec{u} \mid \vec{u} \cdot \vec{v} = 0, \forall \vec{v} \in C\}$ is called a dual code of C.

Theorem 9.1.14 (Binary codes)

If C is an $[n, k, d]$-code, then $C^{\perp}$ is a linear code with dimension $n - k$.

Proof. The fact that $C^{\perp}$ is a linear code is easy to verify. We prove the dimension part.
Since C is an $[n, k, d]$-code, let G be a generator matrix of C. Consider G as a linear transformation from $GF(2)^n$ into $GF(2)^k$ defined by $G\vec{v}, \vec{v} \in GF(2)^n$. Then, it is clear that $Ker(G) = C^{\perp}$ by the definition of $C^{\perp}$. Now, from the fundamental theorem of Linear Algebra, the nullity of G is equal to $n - k$ since the row rank of G is k. □

To construct a good linear code, we can also apply this important observation.

Definition 9.1.15 (Parity-check matrix)

A matrix H is called a parity-check matrix of a linear $[n, k, d]$-code C, if for each $\vec{v} \in C$, $\vec{v}H = \vec{0}_{1\times k}$.

Proposition 9.1.16

If H is a parity-check matrix of C, an $[n, k, d]$-code, then H^t is a

generator matrix of $C^{\perp}$, furthermore, if G is a generator matrix of C, then GH is a $k \times k$ zero matrix.

Example 9.1.17

Let H be the matrix obtained by listing all non-zero 3-dim. binary vectors as its 7 rows. Then, $\{\vec{v} \in GF(2)^7 \mid \vec{v}H = \vec{0}\}$ is a $[7, 4, 3]$-code.

The fact that $d = 3$ in the above code is interesting to know, we leave it as part of exercise. For more information on this part, see [2].

Exercise 4.

Find $A(8, 3)$, $A(9, 3)$ and $A(10, 3)$.

Exercise 5.

Prove that the linear code obtained in Example 9.1.16 is in fact a perfect code.

Exercise 6.

Let $wt(\vec{x})$ (weight of $\vec{x}$) denote the number of non-zero coordinates in the codeword $\vec{x}$. Assume that C is an $[n, k, d]$-code. Show that the minimum weight of $\vec{x} \in C \setminus \{\vec{0}\}$ is in fact equal to $d(C)$.

Exercise 7.

Verify Proposition 9.1.16.

Exercise 8.

Estimate $A(10, 4)$ by using combinatorial designs.

9.2 Special linear codes

In this section, we consider only binary codes and q-ary codes are treated accordingly.

A good way to represent a binary vector in $GF(2)^n =_{def} F^n$ is consider the support of the vector $\vec{x} = (x_0, x_1, \cdots, x_{n-1})$ or simply $x_0x_1x_2\cdots x_{n-1}$.

Definition 9.2.1 (Support of a (0,1)-vector or codeword)

Let $\vec{x} = x_0x_1\cdots x_{n-1}$ be a binary vector (codeword) is defined as $supp(\vec{x}) = A =_{def} \{i \in \mathbb{Z}_n \mid x_i = 1\}$.

For example, 0100011 has support $\{1, 5, 6\}$. In the sense of support, a binary code can be viewed as a collection of sets. Therefore, if C is a code of length n, then $supp(C) = \{supp(\vec{x}) \mid \vec{x} \in C\}$ is a collection of subsets of $\mathbb{Z}_n$. This idea brings together a code and a design. That is, we obtain various codes simply from the knowledge of combinatorial designs. Unfortunately, they are not linear codes.

So, we need a better way to describe a codeword for linear codes.

Definition 9.2.2 (Cyclic codes)

A code C is cyclic if for each codeword $x_0x_1\cdots x_{n-1}$ in C, $x_{n-1}x_0x_1\cdots x_{n-2}$ is also a codeword in C.

For example, $C = \{1011000, 0101100, 0010110, 0001011, 1000101, 1100010, 0110001\}$ is a cyclic code.

By observation, the above set of codewords can be written as $C = \{1+x^2+x^3, x+x^3+x^4, x^2+x^4+x^5, x^3+x^5+x^6, x^4+x^5+1, x^5+x^6+x, x^6+1+x^2\}$ satisfying $x^7 = 1$. We can further simplify C as $\{x^i \cdot (1 + x^2 + x^3) \mid i \in \mathbb{Z}_7 \text{ and } x^7 = 1\}$. (Polynomial representation)

Theorem 9.2.3

Let $f(x)$ be defined as a polynomial of degree $n-k$ in $\mathbb{Z}_2[x]$ such that $f(x) \mid x^n + 1$. Then $C = \{g(x)f(x) \mid g(x) \in \mathbb{Z}_2[x] \text{ and } deg(g(x)) \leq k-1\}$ is a cyclic linear code of length n with dimension k.

Proof. The linearity and dimension of the code is easy to see. On the other part, let $h(x) \in C$, then $h(x) = g_1(x)f(x)$ for some $g_1(x)$ of degree at most $k-1$. Now, consider $x \cdot h(x) = x \cdot g_1(x)f(x)$. If $x \cdot g_1(x)$ is of degree at most $k-1$, then we are done. Otherwise, $x \cdot g_1(x)$ is of degree k and thus $h(x)$ is of degree n. By the fact that $x^n + 1 = f(x) \cdot g_2(x)$, $xg_1(x) = g_2(x) + g_3(x)$ such that $g_3(x)$ is of degree at most $k-1$.
Hence, $h(x) = xg_1(x)f(x) = g_2(x)f(x) + g_3(x)f(x) = g_3(x)f(x) \in C$.
This concludes the proof. □

By adapting the notion of Algebra, we have the following result.

Corollary 9.2.4

If C is a linear cyclic code of length n, then C is an ideal of $\langle \mathbb{Z}_2[x]/ < x^n + 1 >, +, \cdot \rangle$. Moreover, C is generated by a polynomial which is a factor of $x^n + 1$.

Hence, the ideal (cyclic code) mentioned in Theorem 9.2.3 can be written as $< f(x) >$. As examples,
since $x^7 + 1 = (x + 1)(x^3 + x + 1)(x^3 + x^2 + 1)$, linear cyclic codes of length 7 can be obtained by $< x + 1 >$, $< x^3 + x + 1 >$, $< x^3 + x^2 + 1 >$ or their combinations. Especially, the one, $< x^3 + x + 1 >$ has 16 codewords is a perfect code.

In the sense of isomorphism, this code can also be obtained by using the matrix as its parity-check matrix for a $[7, 4, 3]$-code.

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

Using a similar argument, we have a class of "Hamming" codes.

Proposition 9.2.5 (Hamming codes)

Let $n = 2^k - 1$ and $k \geq 3$, and H be the matrix obtained from all non-zero vectors of length k as its row vectors. Then, the linear code C with H as its parity-check matrix is an $[n, n - k, 3]$-code.

Proof. The dimension of C is easy to see, and the distance $d(C) = 3$ is a direct consequence of the fact that the minimum weight of non-zero codewords is 3. □

We can extend the Hamming code to a more general class of BCH-code in honor of the authors Bose-Chaudhuri-Hocquenghem.[3]

Proposition 9.2.6 (Double-error-correcting BCH-code)

Let F be a finite field of order $n = 2^k$ and $F^* = \{1, \alpha, \alpha^2, \cdots, \alpha^{n-2}\}$. Then, the linear code obtained by using the matrix in the following is an $[n-1, n-2k-1, d]$-code C where $d \geq 5$.

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

parity-check matrix of a BCH-code

Proof. The dimension of C is clearly $n-1-2k$. Hence, it's left to prove that $d(C) \geq 5$. We claim any four rows are independent vectors. Let i, j, h and l be the indices of four rows which are satisfying $a(\alpha^i\alpha^{2i}) + b(\alpha^j\alpha^{2j}) + c(\alpha^h\alpha^{2h}) + d(\alpha^l\alpha^{2l}) = \vec{0}$.
Since we only consider binary codes, the following equations hold.
(1) $\alpha^i + \alpha^j + \alpha^h + \alpha^l = \vec{0}$ and (2) $(\alpha^i + \alpha^j + \alpha^h + \alpha^l)^2 = \vec{0}$.
Since all i, j, h, l are distinct, we have (3) $1 + \alpha^{j-i} + \alpha^{h-i} + \alpha^{l-i} = \vec{0}$ and (4) $(1 + \alpha^{j-i} + \alpha^{h-i} + \alpha^{l-i})^2 = \vec{0}$. For convenience, let $j-i$, $h-i$, $l-i$ be j', h', and l' respectively. Now, consider the 4×4 Vandermonde matrix. Therefore, its determinant is not equal to zero as long as $\alpha^{j'}$, $\alpha^{h'}$ and $\alpha^{l'}$ are distinct. This implies that the first two columns are not dependent. The proof follows. □

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ \alpha^{j'} & \alpha^{2j'} & \alpha^{3j'} & \alpha^{4j'} \\ \alpha^{h'} & \alpha^{2h'} & \alpha^{3h'} & \alpha^{4h'} \\ \alpha^{l'} & \alpha^{2l'} & \alpha^{3l'} & \alpha^{4l'} \end{bmatrix}$$

4×4 Vandermonde matrix

Since $d(C) \geq 5$, C can correct at least two errors. By a similar idea, we can construct linear codes with larger code distance. But, there is a trade off between the dimension and the distance.

Exercise 9. (Singleton Bound)

Prove that if an $[n, k, d]_q$-code exists, then $n - k \geq d - 1$.

Exercise 10.

Give some examples to show that $n - k$ can be equal to $d - 1$.

(Maximum Distance Separable Codes, MDS-codes)

Exercise 11. (Varshanov-Gilbert Bound)

Prove that if $\sum_{i=1}^{d-2} \binom{n-1}{i} < 2^m - 1$ for some integers n, d and m, then there exists an $[n, n - m, d]$-code. (Binary codes)

Exercise 12. (Dual code)

Let C be an $[n, k, d]$-code and $C^{\perp}$ be its dual code. Find the dimension of $C^{\perp}$ and $d(C^{\perp})$.

9.3 Quadratic residue codes

In this section, we introduce the idea of constructing linear codes from quadratic residues, a well-known notion in number theory.

Definition 9.3.1 (Quadratic residue)

Let F be a finite field of order p (a prime). Then, in $< F^*, \cdot >$, the set $Q = \{y \mid y = x^2, x \in F^*\}$ is called the quadratic residues of F (or simply of p).

Example 9.3.2

Consider $\mathbb{Z}_7$. $\{1, 2, 4\}$ is the set of quadratic residues of 7. In $\mathbb{Z}_{11}$, $\{1, 4, 9, 5, 3\}$ is the set of quadratic residues.

It is not difficult to check that in F^*, $|Q| = \frac{p-1}{2}$. Note here that if p is a prime, then $< \mathbb{Z}_p, +, \cdot >$ is a finite field. Moreover, we can prove that 2 is a quadratic residue of p if and only if $p \equiv \pm 1 (\mathrm{mod} 8)$, see [ ] for reference. For such primes, the power of 2 is also a quadratic residue of $\mathbb{Z}_p^*$. Now, we are ready to define the codes.

Definition 9.3.3 (Quadratic residue codes)

Let p and l be two primes, p is odd, and l is a quadratic residue of p. Then, the cyclic code generated by $f(x) = \prod_{j \in Q}(x - \alpha^j)$ where Q is the set of quadratic residues of p and α is a primitive p^{th} root of unity in some extension field of $GF(l)$ is a quadratic residue code of length p.

In fact, the quadratic residue code defined above is an ideal of the ring $\mathbb{Z}_l[x]/ < x^p - 1 >$. The dimension of the code is therefore equal to $\frac{p+1}{2}$ since l is a quadratic residue of p. This fact is important for the existence of binary quadratic residue codes, i.e., $p \equiv \pm 1 (\text{mod } 8)$.

Example 9.3.4

In the case when $p = 7$, the quadratic residues are $1, 2, 4$. Let α be a primitive 7^{th} root of 1, i.e., $\alpha^7 = 1$. Then, the cyclic code generated by $(x - \alpha)(x - \alpha^2)(x^2 - \alpha^4)$ over $GF(2) \cong \mathbb{Z}_2$ is of length 7 and dimension 4. This is also known as the Hamming $[7, 4, 3]$-code.

There are famous quadratic residue codes, such as binary Golay code, $[23, 12, 7]$-code and ternary Golay code of length 11 and dimension 6. We note here that by adding a parity check bit to each codeword of a code C with $d(C) = 2t + 1$, we obtain an extended code C^* whose distance is $2t + 2$. This is by the following fact.

(•) If all codewords of C are of even weights (or odd weights resp.), then $d(C)$ is even.

Exercise 13.

Show that a $[23, 12, 7]$-code (Golay code) is a perfect code.

Exercise 14.

Show that the $[11, 6, 5]_3$-code (ternary Golay code) is a perfect code.

Exercise 15.

Prove that the Hamming $[7, 4, 3]$-code and the above two Goaly codes are the only three perfect codes.

9.4 Reed-Solomon Codes

Reed-Solomon codes are a group of error-correcting codes that were introduced by Irving S. Reed and Gustave Solomon in 1960. Their applications are too many to mention. The most prominent of which include the technologies of CD, DVD, QR codes and satellite communication, see [4] for a more substantial introduction.

Reed-Solomon codes are classes of q-ary codes where q is a prime power. Here, we use polynomial representation to explain their ideas.

Definition 9.4.1 (Reed-Solomon Codes)

Let q be a prime power 2^r and F be a finite field of order q. A Reed-Solomon code, $RS(2^r, \delta)$ is a cyclic code of length $n = 2^r - 1$ generated by $g(x) = (x + \beta^{m+1})(x + \beta^{m+2})\cdots(x + \beta^{m+\delta-1})$ where β is a primitive element of F.

Notice that an $RS(2^r, \delta)$ is a $(2^r - 1)$-ary code. For example, if $r = 3$, then a codeword might be $\vec{v} = \beta^3\beta^4 10000$. The reason that $q = 2^r$ is selected comes from the practical application, it is easier to convert an $RS(2^r, \delta)$ code to a binary code of length $r \cdot (q - 1)$.

It is worth of noting that an $RS(2^r, \delta)$ code has distance δ, length $2^r - 1$, and dimension $2^r - \delta$. Therefore, it is an MDS-code satisfying $d(C) = n - k + 1$. In what follows, we introduce the idea of applying RS codes in manufacturing compact disc.

Definition 9.4.2 (Burst length)

The burst length of a binary vector $\vec{v} = (v_0, v_1, \cdots, v_{n-1})$ is defined as the covering range of the first non-zero coordinate and the last non-zero coordinate. If they are v_i and v_j, then the burst length of $\vec{v}$, $bl(\vec{v}) = j - i + 1$.

For example, if $\vec{v} = 1000101000$, then $bl(\vec{v}) = 7 = 6 - 0 + 1$.

Definition 9.4.3 (Cyclic Burst Length)

Let $\pi^k(\vec{v})$ denote the k^{th} cyclic shift of $\vec{v}$. Then, the cyclic burst length $cbl(\vec{v}) = \min\{bl(\pi^k(\vec{v})) \mid k = 1, 2, \cdots, n\}$ where $\vec{v}$ is a binary vector with n coordinates (n-dim. vector).

We can use polynomials to describe the above mentioned length. For example, if $\vec{v} = 1000101000$, then the polynomial used in representing $\vec{v}$ is $1 + x^4 + x^6$. Hence, $bl(\vec{v}) = 6 + 1$. Now, for the cyclic idea, we use the same codeword of length 10. Thus, to find $cbl(\vec{v})$, we check $x^k(1 + x^4 + x^6)$ mod $(1 + x^{10})$ and find the polynomial with smallest gap between lowest degree and highest degree. So, $cbl(\vec{v}) = 7$. But, for $\vec{v} = 1000100$, $bl(\vec{v}) = 4$ and $cbl(\vec{v}) = 3$.

Due to the fact that most errors occurred in a compact disc are burst errors, we had better use the so-called (cyclic) burst error-correcting codes. As a matter of fact, even we can not correct l errors in a code (designed), but it is possible to correct burst errors of length l.

The reason for the above statement can be seen from the following example. Consider a code of length 15 which has at most 3 errors. Then, there are $\binom{15}{0} + \binom{15}{1} + \binom{15}{2} + \binom{15}{3} = 576$ error patterns. In order to correct all the errors, we need a code C with $d(C) \geq 7$. That is to say, if we use RS codes, then $m - k + 1 = 15 - k + 1 = 7$ and thus $k = 9$. On the other hand, if cyclic burst errors are considered, then we have 61 error patterns. Clearly, it needs a code with smaller distance to correct all the errors, i.e., with the same code, we are able to correct a burst error with larger burst length than the number of errors. For more details, the readers can refer to [4].

Exercise 16.

Let C be an RS$(2^4, 7)$-code and $g(x) = \prod_{i=0}^{5}(x + \beta^i)$ where β is a primitive element of $GF(2^4) = \mathbb{Z}_2[x]/ < 1 + x + x^4 >$.

Let $\vec{w} = w(x) = 1 + \beta^4 x + \beta x^3 + \beta^9 x^5 + x^6$ be received in transmission. Find the error pattern and the codeword transmitted.

Exercise 17.

Using the same code, find the error patterns of the following two received words :

(a) $0\beta^3\beta\beta^5\beta^3\beta^2\beta^6\beta^{10}\beta 000000$, (b) $\beta 0\beta^7 0\beta^{12}\beta^3\beta^3 10000000$.

References

1. Elwyn R.Berlekamp, Algebraic Coding Theory, [1968,1984], Laguna Hills, Aegean Park Press.
2. F.J.MacWilliams and N.J.A.Sloane, The Theory of Error-Correcting Codes, 1977, New York, NY : North-Holland Publishing Company.
3. R.C.Bose and D.K.Ray-Chaudhuri, On a class of error-correcting binary group codes, Inform. Control, 3(1960), 68-79.
4. CMU Tutorials, Introduction to Reed-Solomon codes : principles, architecture and implementation.