Example
Consider the (2, 3) parity check code. Then the encoding function E is

$$00 \to 000, 01 \to 011, 10 \to 101, 11 \to 110$$

Thus C, the set of all code words, is

$$\{000, 011, 101, 110\}$$

There are only three possible error vectors of weight 1

$$001, 010, 100$$

and any one of these added to any code word does not yield a code word. Thus every single error is detected by this code.

Suppose that 011 is transmitted and the channel adds to it the error vector 100. Then the received word is 111. But the same word will be received if 101 were transmitted and the channel had added the error vector 010. The received word is in fact equidistant from three code words and, so, this error is not corrected.

Theorem 1.3
A binary code with minimum distance $2k + 1$ is capable of correcting any pattern of k or fewer errors.

Proof
Let C be a code of length n with minimum distance $2k + 1$ and e be an error pattern of weight at most k. Let b be a code word that is transmitted so that the received word is $r = b + e$. If b^* is any code word $b^* \neq b$, then

$$\begin{aligned} d(r, b^*) &= d(b + e, b^*) \\ &= \text{wt}(b + e + b^*) \\ &= \text{wt}(b + b^* + e) \end{aligned}$$

Now

$$\text{wt}(b + b^*) = d(b, b^*) \geq 2k + 1$$

and $\text{wt}(e) \leq k$. Therefore

$$\text{wt}(b + b^* + e) \geq k + 1$$

i.e.

$$d(r, b^*) \geq k + 1$$

Also

$$d(r, b) = \text{wt}(r + b) = \text{wt}(b + b + e) = \text{wt}(e) \leq k$$

Hence b is the nearest code word to r, the received word and by the maximum likelihood decoding principle $D(r) = b$. Thus, the error vector e with $\text{wt}(e) \leq k$ is corrected.

Definition 1.11 – (m, $3m$) triple repetition code
The code of length $3m$ in which the encoding function E: $\mathbb{B}^m \to \mathbb{B}^{3m}$ is defined by

$$E(a) = E(a_1 a_2 \cdots a_m)$$
$$= a_1 a_2 \cdots a_m a_1 a_2 \cdots a_m a_1 a_2 \cdots a_m$$

where $a = a_1 a_2 \cdots a_m \in \mathbb{B}^m$, is called a **triple repetition code**.

If $a = a_1 a_2 \cdots a_m$, $b = b_1 b_2 \cdots b_m$ are distinct words of length m, then $d(a, b) \geq 1$ and, therefore,

$$d(E(a), E(b)) \geq 3$$

Thus, the triple repetition code is capable of detecting any two errors and correcting any single error.

1.2 MATRIX ENCODING TECHNIQUES

One systematic algebraic technique for encoding binary words is by **matrix multiplication**.

Recall that if $\mathbf{A} = (a_{ij})$ is an $m \times n$ matrix and $\mathbf{B} = (b_{rs})$ is an $n \times k$ matrix, then the product $\mathbf{AB}$ of $\mathbf{A}$ and $\mathbf{B}$ is an $m \times k$ matrix (c_{ij}) where

$$c_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj} \quad 1 \leq i \leq m, 1 \leq j \leq k$$

Also recall that if $\mathbf{A}$, $\mathbf{B}$ are two groups then a map $f: \mathbf{A} \to \mathbf{B}$ satisfying the property

$$f(xy) = f(x) f(y) \forall x, y \in \mathbf{A}$$

is called a **homomorphism**. A homomorphism $f: \mathbf{A} \to \mathbf{B}$ is called:

(i) a **monomorphism** if the map f is one–one; and
(ii) an **isomorphism** if the map f is both one–one and onto.

An $m \times n$ matrix, with $m < n$ over $\mathbb{B}$ is called an **encoding matrix** (or **generator matrix**) if the first m columns of it form the identity matrix $\mathbf{I}_m$. Given a generator matrix $\mathbf{G}$, we define an encoding function $E: \mathbb{B}^m \to \mathbb{B}^n$ by

$$E(x) = x\mathbf{G} \quad x \in \mathbb{B}^m$$

Since the first m columns of $\mathbf{G}$ form an identity matrix, the initial part of $x\mathbf{G}$ is x itself. Thus, the matrix encoding method gives distinct code words corresponding to different message words and the map E is one–one. Again, both

$\mathbb{B}^m$ and $\mathbb{B}^n$ are additive Abelian groups and for $x, y \in \mathbb{B}^m$

$$E(x+y) = (x+y)\mathbf{G} = x\mathbf{G} + y\mathbf{G} = E(x) + E(y)$$

Thus E is a homomorphism and we have the following proposition.

Proposition 1.1
For any $m \times n$ generator matrix $\mathbf{G}$, the encoding function $E: \mathbb{B}^m \to \mathbb{B}^n$ given by $E(x) = x\mathbf{G}$, $x \in \mathbb{B}^m$, is a monomorphism.

A code given by a generating matrix is called a **matrix code**.

Definition 1.12
When the code words in a block code form an additive group, the code is called a **group code**.

Corollary
A matrix code is a group code.

To consider other examples of group codes we first need another definition.

Definition 1.13
An $(m, m+1)$ parity check code is the code given by the encoding function $E: \mathbb{B}^m \to \mathbb{B}^{m+1}$ defined by

$$E(a_1 a_2 \cdots a_m) = a_1 a_2 \cdots a_m a_{m+1}$$

where

$$a_{m+1} = \begin{cases} 1 & \text{if } \mathrm{wt}(a) = \mathrm{wt}(a_1 a_2 \cdots a_m) \text{ is odd} \\ 0 & \text{if } \mathrm{wt}(a) \text{ is even} \end{cases}$$

Lemma 1.3
$(m, m+1)$ parity check code is a group code.

Proof
Let $a = a_1 a_2 \cdots a_m$, $a' = a'_1 a'_2 \cdots a'_m$ be two message words and $b = b_1 \cdots b_m b_{m+1}$, $b' = b'_1 \cdots b'_m b'_{m+1}$ be the corresponding code words in the parity check scheme. Now, $b + b' = c_1 \cdots c_m c_{m+1}$, where $c_i = b_i + b'_i$, $1 \leq i \leq m+1$.

$$c_1 + c_2 + \cdots + c_m = (b_1 + \cdots + b_m) + (b'_1 + \cdots + b'_m)$$

which is odd iff one of $(b_1 + \cdots + b_m)$ and $(b'_1 + \cdots + b'_m)$ is odd and the other is even. But, in this case, either $b_{m+1} = 1$ and $b'_{m+1} = 0$ or $b_{m+1} = 0$ and $b'_{m+1} = 1$. Hence, in this case $c_{m+1} = 1$.

Again $c_1 + \cdots + c_m$ is even iff either both of $b_1 + \cdots + b_m$ and $b'_1 + \cdots + b'_m$ are odd or both are even. If this is so, then $b_{m+1} + b'_{m+1} = 0$, i.e. $c_{m+1} = 0$.

This proves that c is a code word under the parity check scheme. This gives a composition in the set of all code words. The 0 word is the identity and every word is its own inverse. Hence, the set of all code words forms a group.

Lemma 1.4
The triple repetition $(m, 3m)$ code is a group code.

Proof
If

$$b = a_1 \cdots a_m a_1 \cdots a_m a_1 \cdots a_m$$

and

$$b' = a'_1 \cdots a'_m a'_1 \cdots a'_m a'_1 \cdots a'_m$$

are two code words, then

$$b + b' = c_1 \cdots c_m c_1 \cdots c_m c_1 \cdots c_m$$

where $c_i = b_i + b'_1$, which is again a block of length m repeated three times. Hence $b + b'$ is a code word. 0 is the identity and every code word is its own inverse. Hence, the code is a group code.

Exercise 1.1

1. Prove that an $(m, m + 1)$ parity check code is a matrix code. What is the generator matrix of this code?
2. Give an example of a group code which is not a matrix code.

Proposition 1.2
For a group code, the minimum distance equals the minimum of the weights of the non-zero code words.

Proof
Let d be the minimum distance of the group code. Then there exist code words b, b' such that $d = d(b, b') = \mathrm{wt}(b + b')$. But the code being a group code, $b + b'$ is a code word. Let t be the minimum of the weights of non-zero code words. Then, by the above,

$$d \geq t$$

t being the minimum among weights of non-zero code words, there exists a non-zero code word b'' such that

$$t = \mathrm{wt}(b'') = d(b'', 0) \geq d$$

Hence $d = t$.

Remark
In group codes, the error patterns that pass undetected are precisely those which correspond to non-zero code words.