Proof

If $e = b^*$ is a code word, then for any code word b, $b + e$ is again a code word and the error pattern e goes undetected.

Conversely, suppose that an error pattern e goes undetected. Then there exists a code word b such that $b + e = b^*$ (say) is again a code word. This shows that $e = b + b^*$ – again a code word.

Example

Consider the 3×6 generating matrix

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

All the code words of the code generated by this matrix are:

000 →	000000	001 →	001111
100	100110	110	110101
010	010011	101	101001
011	011100	111	111010

This is a group code with 4 code words of weight 3, 3 code words of weight 4 and only the zero code word of weight 0. Thus, the minimum distance of the code is 3. Hence, this is a code which is capable of correcting any single error and detecting any error of weight 2.

Next, we give a decoding procedure for group codes with the help of which the probability of an error passing undetected is minimized. This decoding procedure uses decomposition of a finite group into cosets which we describe briefly.

Recall that a non-empty subset N of a group M is called a **subgroup** of M if:

(i) the composition in M induces a composition in N, i.e. wherever $a, b \in N$, then $ab \in N$; and
(ii) N is a group w.r.t. the induced composition.

For example, if $\mathbf{G}$ is an $m \times n$ matrix over $\mathbb{B}$ then $\{a\mathbf{G} \mid a \in \mathbb{B}^m\}$ is a subgroup of $\mathbb{B}^n$. The order of the subgroup is at most 2^m whereas if $\mathbf{G}$ is a generator matrix, then the order of the subgroup is precisely 2^m (Proposition 1.1).

Let $n > 1$ and C be a subgroup of $\mathbb{B}^n$. For $a \in \mathbb{B}^n$, $a + C = \{a + c : c \in C\}$ is a subset (and not in general a subgroup) of $\mathbb{B}^n$ called a **coset** of C in $\mathbb{B}^n$. If $b \in a + C$, then $b = a + c$ for some $c \in C$. Therefore, for any $c' \in C$,

$$b + c' = a + (c + c') \in a + C$$

Thus $b + C \subseteq a + C$. Again $b = a + c$ implies $a = b + c$ and, as above, it follows that $a + C \subseteq b + C$. Hence $a + C = b + C$. On the other hand, if $a + C = b + C$, then $b = b + 0 \in b + C = a + C$. We thus have

$$b \in \mathbb{B}^n \text{ is in } a + C \text{ iff } a + C = b + C \tag{1.1}$$

Now, consider two cosets $a+C$ and $b+C$ of C in $\mathbb{B}^n$. If $(a+C)\cap(b+C)\neq\emptyset$, there exists an $x\in a+C$ and $x\in b+C$. It follows from (1.1) above that $a+C=x+C$ and $b+C=x+C$ and therefore, $a+C=b+C$. So

$$\text{two cosets of } C \text{ in } \mathbb{B}^n \text{ are either disjoint or identical} \tag{1.2}$$

Observe that the number of elements in any coset $a+C$ of C in $\mathbb{B}^n$ is equal to the order of the subgroup C (which equals the number of elements in C). Every element of $\mathbb{B}^n$ is in some coset (in fact in a unique coset in view of (1.2) above) of C in $\mathbb{B}^n$, e.g. if $a\in\mathbb{B}^n$, then $a\in a+C$. Also, the group $\mathbb{B}^n$ being finite, the number of distinct cosets of C in $\mathbb{B}^n$ is finite. If $a^1+C,\ldots,a^k+C$ are all the distinct cosets of C in $\mathbb{B}^n$, then we have

$$\mathbb{B}^n=\bigcup_{i=1}^{k}(a^i+C) \qquad \text{and} \qquad (a^i+C)\cap(a^j+C)=\emptyset \text{ for } i\neq j \tag{1.3}$$

Consider an (m,n) group code and let C be the set of all code words of this code. Then order of C is 2^m. $\mathbb{B}^n$ is the set of all words of length n which as seen earlier is a group and C is a subgroup of it. We can then write $\mathbb{B}^n$ as a disjoint union of cosets of C in $\mathbb{B}^n$. In each coset of C in $\mathbb{B}^n$, we choose a word b^i of least weight and call it a **coset leader**. Observe that

$$\mathrm{wt}(b^i)\leq \mathrm{wt}(b^i+c^i)\forall c^i\in C$$

Any element c of $\mathbb{B}^n$ can be uniquely written as $c=b^i+c^j$ for some $c^j\in C$. We define a decoding function D by putting

$$D(c)=c^j$$

For any code word, $c^k\neq c^j$, we have

$$\begin{aligned}
d(c,c^k)&=d(b^i+c^j,c^k)=\mathrm{wt}(b^i+c^j+c^k)\\
&\geq \mathrm{wt}(b^i)\\
&=d(b^i+c^j,c^j)=d(c,c^j)
\end{aligned}$$

Thus, there is no code word lying within the circle with centre at c and radius equal to $d(c,c^j)$.

This decoding procedure or process is known as **decoding by coset leaders.**

Theorem 1.4
In group codes, decoding by coset leaders corrects precisely those error patterns which are coset leaders.

Proof
Suppose that an error pattern e is corrected by this method of decoding. Let c^i be a code word transmitted so that the received word is $b=c^i+e$. Then $b=b^k+c^r$ for some code word c^r and coset leader b^k. By the decoding process,

$D(b) = c^r$ and since the error is corrected, we must also have $D(b) = c^i$. Hence, $c^r = c^i$. Thus, $b^k + c^i = c^i + e$ or $e = b^k$ – a coset leader.

Conversely, suppose that $e = b^k$ is a coset leader. Then, for any code word, c^i, the received word is $c^i + e = b^k + c^i$ and $D(b^k + c^i) = c^i$. Hence the error pattern is corrected.

Example

Consider first the (3, 4)-parity check code $\mathscr{C}$:

$$\begin{array}{ll} 000 \longrightarrow 0000 & 011 \longrightarrow 0110 \\ 001 \longrightarrow 0011 & 101 \longrightarrow 1010 \\ 010 \longrightarrow 0101 & 110 \longrightarrow 1100 \\ 100 \longrightarrow 1001 & 111 \longrightarrow 1111 \end{array}$$

Coset decomposition of $\mathscr{C}$ in $\mathbb{B}^4$ is:

Coset leader	The coset							
0000	0000	0011	0101	1001	0110	1010	1100	1111
0001	0001	0010	0100	1000	0111	1011	1101	1110

Observe that we could have taken 0010 or 0100 as coset leader for the coset $0001 + \mathscr{C}$. Having chosen 0001 as coset leader, if 1011 is the word received, then decoding by coset leaders decodes this word into $1011 + 0001 = 1010$ which is the word at the head of the column in which the received word 1011 lies.

Next, consider the (2, 6) triple repetition code $\mathscr{C}$:

$$00 \rightarrow 000000,\ 01 \rightarrow 010101,\ 10 \rightarrow 101010,\ 11 \rightarrow 111111$$

Here $\mathscr{C}$ is a subgroup of $\mathbb{B}^6$ and, using Lagrange's theorem, we may represent the elements of $\mathbb{B}^6$ in tabular form as follows:

Coset leader	The coset			
000000	000000	010101	101010	111111
000001	000001	010100	101011	111110
000010	000010	010111	101000	111101
000100	000100	010001	101110	111011
001000	001000	011101	100010	110111
010000	010000	000101	111010	101111
100000	100000	110101	001010	011111
000011	000011	010110	101001	111100
001001	001001	011100	100011	110110
100001	100001	110100	001011	011110
000110	000110	010011	101100	111001
010010	010010	000111	111000	101101
001100	001100	011001	100110	110011
100100	100100	110001	001110	011011
011000	011000	001101	110010	100111
110000	110000	100101	011010	001111

In this case also, we find that if the received word $r = b + e$, then decoding by coset leaders decodes this word into b, the code word which lies at the head of the column in which r lies.

The above observed principle holds in general. Thus, if all the words of $\mathbb{B}^n$ are written in a tabular form each row being a coset $b^i + C$ of C in $\mathbb{B}^n$ with b^i a coset leader and the first row representing the words of C, to decode a received word r we locate it in the table. The word r is decoded into the code word which appears at the head of the column in which r occurs.

1.3 GENERATOR AND PARITY CHECK MATRICES

Let us consider, once again, the matrix code given by the generator matrix

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

If $a_1a_2a_3a_4a_5a_6$ is the code word in this code corresponding to the message word $a_1a_2a_3$, then

$$(a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \quad a_6) = (a_1 \quad a_2 \quad a_3)\mathbf{G}$$

and thus

$$a_4 = a_1 + a_3$$
$$a_5 = a_1 + a_2 + a_3$$
$$a_6 = a_2 + a_3$$

These equations may be rewritten as

$$a_1 + a_3 + a_4 = 0$$
$$a_1 + a_2 + a_3 + a_5 = 0$$
$$a_2 + a_3 + a_6 = 0$$

These are called **parity check equations**. In matrix form, these equations may be written as

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{pmatrix} = 0$$

The matrix

$$\mathbf{H}=\begin{pmatrix}1 & 0 & 1 & 1 & 0 & 0\\ 1 & 1 & 1 & 0 & 1 & 0\\ 0 & 1 & 1 & 0 & 0 & 1\end{pmatrix}$$

is called the **parity check matrix** of the code. Let

$$\mathbf{A}=\begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 1\\ 1 & 1 & 1\end{pmatrix}$$

Then $\mathbf{G}=(\mathbf{I}_3 \quad \mathbf{A})$, where $\mathbf{I}_3$ is the identity matrix of order 3. Also

$$\mathbf{A}'=\mathbf{A}^t=\begin{pmatrix}1 & 0 & 1\\ 1 & 1 & 1\\ 0 & 1 & 1\end{pmatrix}$$

and $\mathbf{H}=(\mathbf{A}' \quad \mathbf{I}_3)$.

We shall later prove this relation between the generator matrix and the corresponding parity check matrix in the general case. The matrix $\mathbf{H}$ has the property that for any code word a, $\mathbf{Ha}=0$. (Note that $\mathbf{a}$ is the vector formed by taking the elements of the code word a.)

We observe that the $(m, m+1)$ parity check code we considered earlier has $a_1a_2\text{---}a_{m+1}$ as a code word provided

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

Then $a_1+a_2+\cdots+a_m+a_{m+1}=0$.

Observe that this code is a matrix code given by the generator matrix

$$\mathbf{G}=\begin{pmatrix}1 & 0 & \cdots & 0 & 1\\ 0 & 1 & \cdots & 0 & 1\\ \vdots & \vdots & \cdots & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & 1\end{pmatrix}$$

The parity check matrix of this code is the $1\times(m+1)$ matrix $\mathbf{H}=(1 \quad 1 \quad \cdots \quad 1)$.

Next, we consider the (3, 6) matrix code given by the generator matrix

$$\mathbf{G}=\begin{pmatrix}1 & 0 & 0 & 1 & 1 & 1\\ 0 & 1 & 0 & 0 & 1 & 1\\ 0 & 0 & 1 & 1 & 0 & 1\end{pmatrix}$$

For any code word $a_1a_2\text{---}a_6$ in this code we have

$$(a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \quad a_6)=(a_1 \quad a_2 \quad a_3)\mathbf{G}$$

and so

$$a_4 = a_1 + a_3$$
$$a_5 = a_1 + a_2$$
$$a_6 = a_1 + a_2 + a_3$$

or

$$a_1 + a_3 + a_4 = 0$$
$$a_1 + a_2 + a_5 = 0$$
$$a_1 + a_2 + a_3 + a_6 = 0$$

In matrix notation, these parity check equations may be written as

$$\mathbf{H}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{pmatrix} = 0$$

where

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Again, observe that if

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

then

$$\mathbf{A}^t = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

and

$$\mathbf{G} = (\mathbf{I}_3 \quad \mathbf{A}) \qquad \text{while} \qquad \mathbf{H} = (\mathbf{A}^t \quad \mathbf{I}_3)$$

We now define a parity check matrix in general.

Definition 1.14 – parity check matrix
If $m < n$, then any $(n - m) \times n$ matrix $\mathbf{H}$, whose last $n - m$ columns form the identity matrix $\mathbf{I}_{n-m}$ is called a **parity check matrix**.