Binary Linear Codes

In coding theory, a linear code is an error-correcting code for which any linear combination of codewords is also a codeword. Linear codes allow for more efficient encoding and decoding algorithms than other codes such as syndrome decoding.

A subset C of $\mathbb{K}^n$ is called a *linear code*, if C is a subspace of $\mathbb{K}^n$ (i.e., C is closed under addition). A linear code of dimension k contains precisely 2^k codewords.

1) The fact that the zero vector is a member of any subspace of a vector space, the zero vector is always a codeword.

2) The fact that any subspace of a vector space is closed under addition, the sum of two codewords is another codeword.

3) The number of codewords in a linear code C is 2^k.

Proposition 1. *In a linear code C, the minimum distance is equal to the minimal weight among all non-zero codewords.*

Proof. Let x and y be codewords in C, then $x - y \in C$. We then have $d(x, y) = d(x - y, 0)$ which is the weight of $x - y$. □

♠ **Generator Matrix.** In coding theory, a generator matrix is a matrix whose rows form a basis for a linear code. The codewords are all of the linear combinations of the rows of this matrix, that is, the linear code is the row space of its generator matrix.

A $k \times n$ matrix G is a *generator matrix* for some linear code C, if the rows of G are linearly independent; that is if the rank of G equals k. A linear code generated by a $k \times n$ generator matrix G is called a $[n, k]$ code. An $[n, k]$ code with distance d is said to be an $[n, k, d]$ code. If G_1 is row equivalent to G, then G_1 also generates the same linear code C. If G_2 is column equivalent to G, then the linear code C_2 generated by G_2 is not equal to C, but equivalent to C.

Consider the 3×5 generator matrix

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

of rank 3. By interchanging the first row and the third row, we obtain another generator matrix

$$G_1 = \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 \\ 0 & 1 & 0 & \vdots & 1 & 0 \\ 0 & 0 & 1 & \vdots & 1 & 0 \end{pmatrix} = [\, I_3 \quad B\,]$$

for the same linear code. Note that G and G_1 are in *reduced row echelon form* ($\mathcal{RREF}$). This linear code has an information rate of $3/5$ (i.e., G and G_1 accept all the messages in $\mathbb{K}^3$ and change them into words of length 5). The generator matrix $G_1 = [\, I_3 \quad B\,]$ is said to be in *standard form,* and the code C generated by G is called a *systematic code.* Not all linear codes have a generator matrix in standard form. For example, the linear code $C = \{000, 100, 001, 101\}$ has six generator matrices

$$G_1 = \begin{pmatrix} 100 \\ 001 \end{pmatrix} \ , \ G_2 = \begin{pmatrix} 001 \\ 100 \end{pmatrix} \ , \ G_3 = \begin{pmatrix} 100 \\ 101 \end{pmatrix} \ , \ G_4 = \begin{pmatrix} 001 \\ 101 \end{pmatrix} \ , \ G_5 = \begin{pmatrix} 101 \\ 100 \end{pmatrix} \ , \text{ and } G_6 = \begin{pmatrix} 101 \\ 001 \end{pmatrix}.$$

None of these matrices are in standard form. Note that the matrix $G' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ in standard form generates the code $C' = \{000, 100, 010, 110\}$ which is equivalent to C. If G is in $\mathcal{RREF}$, then any column of G which is equal to the vector e_i is called a *leading column.* If $\mathbf{m} \in \mathbb{K}^k$ is the message and $v = \mathbf{m}G \in \mathbb{K}^n$ is the codeword of a systematic code, then the first k digits of v which represent the message $\mathbf{m}$ are called *information digits*, while the last $n-k$ digits are called *redundancy* or *parity-check* digits. If C is not a systematic code, then to recover the message from a codeword we select the digits corresponding to the leading columns $e_1,\ e_2, \cdots, e_k$. For example, if $G = \begin{pmatrix} 001 \\ 100 \end{pmatrix} = [\, e_2 \quad \theta \quad e_1\,]$ and $v = 001$, then we recover the message $\mathbf{m} = \mathbf{10}$ from the last digit and the first digit of v respectively.

Let S be a subset of $\mathbb{K}^n$. The set of all vectors orthogonal to S is denoted by $S^\perp$ and called the *orthogonal complement* of S. It can readily be shown that $S^\perp$ is a linear code. If $C =< S >$, then $C^\perp =< S^\perp >$ which is also a linear code is called the *dual code* of C.

♠ **Parity-Check Matrix.** A matrix H is called a *parity-check matrix* for a linear code C of length n generated by the matrix G, if the columns of H form a basis for the dual code $C^\perp$. If v is a word in C, then $vH = \theta$.

The parity check matrix for a given binary linear code can be derived from its generator matrix (and vice versa). If the generator matrix for an [n,k]-code is in standard form

$G = (\, I_k | P \,)$, then the parity check matrix is given by

$$H = \begin{bmatrix} P \\ I_{n-k} \end{bmatrix}, \qquad \text{because} \quad G\,H = P + P = Z_{k,n-k}\,.$$

For example, if a binary code has the generator matrix

$$G = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}, \text{ then its parity check matrix is } H = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

A code C is called *self-dual* if $C = C^{\perp}$. In this case n must be even and C must be an $(n, n/2)$ code. If G is a generator matrix of a self-dual code, then $H = G^t$. Both the generator matrices

$$G = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} \text{ and } G_1 = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

generate self-dual codes but only G_1 is in $\mathcal{RREF}$. If $G = [I \;\, B]$ is a generator of a self-dual code, then $B^2 = I$.

Theorem 1. *Let H be a parity-check matrix for a linear code C generated by the $k \times n$ matrix G. Then*

(i) the rows of G are linearly independent;

(ii) the columns of H are linearly independent;

(iii) $GH = Z_k$, where Z_k is the $k \times k$ zero matrix;

(iv) by permuting columns of H, we obtain another parity-check matrix corresponding to G,

(v) $\dim(C) = \mathrm{rank}(G)$, $\dim(C^{\perp}) = \mathrm{rank}(H)$, and $\dim(C) + \dim(C^{\perp}) = n$;

(vi) H^t is a generator matrix for $C^{\perp}$ with G^t its parity-check matrix;

(vii) if C is self-dual with $G = [I_k \;\, B]$ its generator, then $G_1 = [B \;\, I_k]$ also generates C;

(viii) C has distance d if and only if any set of $d-1$ rows of H is linearly independent, and at least one set of d rows of H is linearly dependent.

♣ Algorithms for Finding Generator and Parity-Check Matrices.

Example 1. *Let $S = \{01100, 01010, 11100, 00110\}$ be a subset of $I\!\!K^5$ generating the linear code C. By using the words in S, we define the matrix*

$$M = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \end{bmatrix} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

Note that

$$M_1 + M_2 = 0\,1\,1\,0\,0 + 0\,1\,0\,1\,0 = 0\,0\,1\,1\,0 = M_4.$$

Thus the linear binary code C generated by S has dimension 3; so the matrix

$$G = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \end{bmatrix} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}$$

is a generator matrix.

Now we use some row operations on G, to obtain a generator matrix in standard form. Let

$$E_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad \text{and } E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

be elementary matrices, then

$$\begin{aligned} G_1 = E_3\,E_2\,E_1\,G &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & \vdots & 0 & 0 \\ 0 & 1 & 0 & \vdots & 1 & 0 \\ 0 & 0 & 1 & \vdots & 1 & 0 \end{pmatrix}. \end{aligned}$$

is a generator matrix in standard form.

To obtain a parity-check matrix of the linear code C, we form the matrix

$$B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}$$

from the last two columns of G_1; the matrix

$$H_1 = \begin{bmatrix} B \\ I_2 \end{bmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ \cdots & \cdots \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

will be a parity-check matrix associated to the generator matrix G_1 .

Example 2. *Let $S = \{1010010101, 0001010001, 0000100100, 0000001001, 0000000011\}$ be a linearly independent set generating C. The generator matrix*

$$G = \begin{array}{c} \begin{array}{cccccccccc} e_1 & & & e_2 & e_3 & & e_4 & & e_5 & \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

is in $\mathcal{RREF}$ but not in standard form.

We permute the columns of G into order $1, 4, 5, 7, 9, 2, 3, 6, 8, 10$ to form the matrix

$$G_1 = G * P = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \left(\begin{array}{ccccc:ccccc} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

Then we form the matrix H_1 and finally rearrange the rows of H_1 into their natural order to form the parity-check matrix H.

$$H_1 = \begin{bmatrix} B \\ I_5 \end{bmatrix} = \begin{array}{c} \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} 1 \\ 4 \\ 5 \\ 7 \\ 9 \\ 2 \\ 3 \\ 6 \\ 8 \\ 10 \end{array} ; \quad H = P * H_1 = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array}$$

The columns of H form a basis for $C^{\perp}$.

♡ **Matlab.** *To permute the columns of G into order $1, 4, 5, 7, 9, 2, 3, 6, 8, 10$ to form the matrix G_1, first we define the permutation matrix P as follows:*

$\gg$ $P = eye(10)$; $T = P$; < *Return key* >

$\gg$ $P(: ,\ 2) = T(: ,\ 4)$; $P(: ,\ 3) = T(: ,\ 5)$; $P(: ,\ 4) = T(: ,\ 7)$; $P(: ,\ 5) = T(: ,\ 9)$; $P(: ,\ 6) = T(: ,\ 2)$; $P(: ,\ 7) = T(: ,\ 3)$; $P(: ,\ 8) = T(: ,\ 6)$; $P(: ,\ 9) = T(: ,\ 8)$; < *Return key* >

$\gg$ $G1 = G * P$ < *Return key* >

Finally H is obtained as follows:

$\gg$ $B = G1(: ,\ 6 : 10)$; $H1 = [B\ ;\ eye(5)\ ]$; $H = P * H1$ < *Return key* >

♠ **Maximum Likelihood Decoding (MLD) for Linear Codes.** *We will describe a procedure for either $\mathcal{CMLD}$, Complete Maximum Likelihood Decoding (no need for a retransmission) or $\mathcal{IMLD}$, Incomplete Maximum Likelihood Decoding (ask for a retransmission)* for a linear code.

If $C \in I\!\!K^n$ is a linear code of dimension k, and if $u \in I\!\!K^n$, we define the *coset of C determined by u denoted $\widehat{u}$ as follows:*

$$\widehat{u} = C + u = \{v + u : v \in C\}.$$

There are as many as 2^{n-k} distinct cosets of C in $I\!\!K^n$ of order 2^k, *where every word in $I\!\!K^n$ is contained in one of the cosets.*

Note. *A coset leader is a member of the coset with minimum weight. If a coset contains more than one coset leader and if the received word is in that coset, then a retransmission is required.*

Theorem 2. *Let C be a linear code. Then*

(i) $\widehat{\theta} = C$;

(ii) if $v \in \widehat{u} = C + u$, then $\widehat{v} = \widehat{u}$;

(iii) $u + v \in C$ if and only if u and v are in the same coset.

The parity-check matrix and cosets of the code play fundamental roles in the decoding process.

Let C be a linear code. Assume the codeword v in C is transmitted and the word w is received, resulting in the *error pattern* $u = v + w$. Then $w + u = v$ is in C, so **the error pattern u and the received word w are in the same coset of C.** Since error patterns of small weight are the most likely to occur, we choose a word u of least weight in the coset $\widehat{u}$ (which must contain w) and conclude that $v = w + u$ was the word sent.

Let $C \in I\!\!K^n$ be a linear code of dimension k and let H be a parity-check matrix. For any word $w \in I\!\!K^n$, the *syndrome of* w is the word $s(w) = w\,H$ in $I\!\!K^{n-k}$.

Theorem 3. *Let H be a parity-check matrix for a linear code C. Then*

(i) $wH = \theta$ if and only if w is a codeword in C.

(ii) $w_1H = w_2H$ if and only if w_1 and w_2 lie in the same coset of C.

(iii) If u is the error pattern in a received word w, then uH is the sum of the rows of H that correspond to the positions in which errors occurred in transmission.

A table which matches each syndrome with its coset leader, is called a *standard decoding array,* or $\mathcal{SDA}$. To construct an $\mathcal{SDA}$, first list all the cosets for the code, and choose from each coset word of least weight as coset leader u. Then find a parity-check matrix for the code and, for each coset leader u, calculate its syndrome uH.

Example. Consider the code $C = \{0000,\ 1011,\ 0101,\ 1110\}$ generated by the generator matrix $G = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ with a parity-check matrix $H = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$. Here are members of $I\!\!K^4$ which are not in C:

$$I\!\!K^4 - C = \{1000,\ 0100,\ 0010,\ 0001,\ 1100,\ 0110,\ 0011,\ 1001,\ 1101,\ 1101,\ 0111,\ 1111\}$$

From the fact that $16/4 = 4$, we conclude that there are 4 cosets associated to the code C. We need the word 0000 and 3 members of $I\!\!K^4 - C$ with least weights as our coset leaders. We choose 1000, 0100, and 0010. Here are our cosets:

$$\begin{aligned} \widehat{0000} &= \{0000, 1011, 0101, 1110\} \\ \widehat{1000} &= \{1000, 0011, 1101, 0110\} \\ \widehat{0100} &= \{0100, 1111, 0001, 1010\} \\ \widehat{0010} &= \{0010, 1001, 0111, 1100\} \end{aligned}$$

Notice that $0001 \in \widehat{0010}$. Thus

$$\widehat{0100} = \{0100, 1111, 0001, 1010\} = \widehat{0001}.$$

Here is the $\mathcal{SDA}$ for the code:

Coset leader u	Syndrome uH
0000	00
1000	11
0100 or 0001	01*
0010	10

The syndrome with a $*$ indicates a retransmission in the case of $\mathcal{IMLD}$. Notice that the set of error patterns that can be corrected using $\mathcal{IMLD}$ is equal to the set of unique coset leaders.

If $w = 1101$ is received, then the syndrome of w is $s(w) = wH = 11$. Notice that the word of least weight in the coset $\widehat{w}$ is $u = 1000$ and the syndrome of u is $s(u) = uH = 11 = wH$. Furthermore, $\mathcal{CMLD}$ concludes $v = w + u = 1101 + 1000 = 0101$ was sent, so there was an error in the first digit. Notice also that $s(w) = 11$ picks up the first row of H corresponding to the location of the most likely error; also the coset leader in the $\mathcal{SDA}$ is 1000. The calculations

$$\begin{aligned} d(0000, 1101) &= 3 \quad d(0101, 1101) = 1 \\ d(1011, 1101) &= 2 \quad d(1110, 1101) = 2 \end{aligned}$$

which give the distances between w and each codeword in C, show that indeed $v = 0101$ is the closest word in C to w.

For the received $w = 1111$, however, the same calculations

$$\begin{array}{ll} d(0000, 1111) = 4 & d(0101, 1111) = 2 \\ d(1011, 1111) = 1 & d(1110, 1111) = 1 \end{array}$$

reveal a tie for the closest word in C to w. This is not surprising, since there was a choice for a coset leader for the syndrome $1111H = 01$. In the case of $\mathcal{CLMD}$, we arbitrarily choose a coset leader, which in effect arbitrary selects one codeword in C closest to w. Using $\mathcal{IMLD}$, we ask for retransmission.

Here is a binary table for the characters in the English language:

Character	Number	Message	Character	Number	Message
$\flat$ (*space*)	0	0 0 0 0 0	$\mathcal{P}$	16	0 0 0 0 1
$\mathcal{A}$	1	1 0 0 0 0	$\mathcal{Q}$	17	1 0 0 0 1
$\mathcal{B}$	2	0 1 0 0 0	$\mathcal{R}$	18	0 1 0 0 1
$\mathcal{C}$	3	1 1 0 0 0	$\mathcal{S}$	19	1 1 0 0 1
$\mathcal{D}$	4	0 0 1 0 0	$\mathcal{T}$	20	0 0 1 0 1
$\mathcal{E}$	5	1 0 1 0 0	$\mathcal{U}$	21	1 0 1 0 1
$\mathcal{F}$	6	0 1 1 0 0	$\mathcal{V}$	22	0 1 1 0 1
$\mathcal{G}$	7	1 1 1 0 0	$\mathcal{W}$	23	1 1 1 0 1
$\mathcal{H}$	8	0 0 0 1 0	$\mathcal{X}$	24	0 0 0 1 1
$\mathcal{I}$	9	1 0 0 1 0	$\mathcal{Y}$	25	1 0 0 1 1
$\mathcal{J}$	10	0 1 0 1 0	$\mathcal{Z}$	26	0 1 0 1 1
$\mathcal{K}$	11	1 1 0 1 0	.	27	1 1 0 1 1
$\mathcal{L}$	12	0 0 1 1 0	,	28	0 0 1 1 1
$\mathcal{M}$	13	1 0 1 1 0	;	29	1 0 1 1 1
$\mathcal{N}$	14	0 1 1 1 0	?	30	0 1 1 1 1
$\mathcal{O}$	15	1 1 1 1 0	!	31	1 1 1 1 1

Example. Let $\mathcal{C}$ be a binary linear code defined by the following generator matrix $\mathcal{G}$ with the parity-check matrix $\mathcal{H}$:

$$\mathcal{G} = \begin{bmatrix} \mathcal{G}_1 \\ \mathcal{G}_2 \\ \mathcal{G}_3 \\ \mathcal{G}_4 \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \quad \text{and} \quad \mathcal{H} = \begin{bmatrix} \mathcal{H}_1 \\ \mathcal{H}_2 \\ \mathcal{H}_3 \\ \mathcal{H}_4 \\ \mathcal{H}_5 \\ \mathcal{H}_6 \\ \mathcal{H}_7 \\ \mathcal{H}_8 \\ \mathcal{H}_9 \end{bmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Note. The code $\mathcal{C} \subset \mathbb{K}^9$ is generated by the matrix $\mathcal{G}$ of rank 5 with $2^5 = 32$ words; and there are exactly

$$\frac{|\mathbb{K}^9|}{|\mathcal{C}|} = \frac{2^9}{2^5} = 2^4 = 16 \text{ cosets.}$$

Also, it is not difficult to see that $\mathcal{C}$ is a $(9,5,3)$ code.

The generator matrix $\mathcal{G}$ encodes any binary message associated with an English character in the above Alphabet table, into a codeword of length 9.

Notice that the product of $e_k \in I\!K^9$ by the parity-check matrix $\mathcal{H}$ is $\mathcal{H}_k$, the k-th row of $\mathcal{H}$. Any transmission error involving the last 4 digits of the codeword will not alter the meaning of the original message. Thus any syndrome that is a linear combination of the last four columns of $\mathcal{H}$ will be ignored. The fact that $\mathcal{C}$ is a $(9,5,3)$ code, implies that any received word with a syndrome equal to one of the first five rows of $\mathcal{H}$ will be corrected without involving any coset. Therefore there will be fewer number of cosets involved in our error-correcting schemes.

Suppose now, we need to encode and transmit the message $\mathcal{M}\ \mathcal{A}\ \mathcal{T}\ \mathcal{H}$ with the above generator matrix $\mathcal{G}$.

Step 1. First we change our message into a binary message matrix

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

Step 2. Next, we encode the matrix $\mathcal{M}$ with $\mathcal{G}$ to obtain the codeword matrix $\mathcal{V}$:

$$\mathcal{V} = \mathcal{M} * \mathcal{G} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}$$
$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

Step 3. Suppose after the transmission, the message is received as:

$$\mathcal{W} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Step 4. Since $\mathcal{G}$ is in standard form, the received binary message matrix $\mathcal{N}$ will be obtained from the first five columns of $\mathcal{W}$:

$$\mathcal{N} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Step 5. The Alphabet table produces the message $\mathcal{O}\ \mathcal{A}\ \mathcal{T}\ \flat$.

Step 6. The syndrome matrix:

$$\mathcal{S} = \mathcal{W} * \mathcal{H} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

indicates that the first and last characters were not transmitted correctly.

Step 7. Suppose both cosets leaders u_1 and u_4 associated with the syndromes $s_1 = 1\ 0\ 1\ 0$ and $s_4 = 0\ 1\ 1\ 0$, respectively are unique, then $w_1 + u_1$ and $w_4 + u_4$ will produce the correct characters and the message $\mathcal{O}\ \mathcal{A}\ \mathcal{T}\ \flat$ will be changed into the message $\mathcal{M}\ \mathcal{A}\ \mathcal{T}\ \mathcal{H}$.

Step 8. Suppose conditions in Step 7 are not met. In the case of $\mathcal{CMLD}$, an arbitrary coset leader is selected in order to rectify the error. In the case of $\mathcal{IMLD}$, a retransmission of any character with a non-zero syndrome, associated with the coset with multiple leaders will be needed.

Note. The fact that our code $\mathcal{C}$ has a weight of 3 and $s_1 = 1\ 0\ 1\ 0 = \mathcal{H}_2$ and $s_4 = 0\ 1\ 1\ 0 = \mathcal{H}_4$, we conclude that

$$\begin{aligned}\mathcal{V} = \mathcal{W} + \begin{bmatrix} e_2 \\ \theta \\ \theta \\ e_4 \end{bmatrix} &= \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix},\end{aligned}$$

which produces the message $\mathcal{M}\ \mathcal{A}\ \mathcal{T}\ \mathcal{H}$.