Solution of exercise sheet n° 6 :

6-1.

(a) The size of the matrix is given by $n = 6$ and $m = 4$. n corresponds to the size of the codewords. The rank of H is equal to $m = 4$ and corresponds to the number of parity bits. We define a codeword vector $\mathbf{x}$ of components x_i, $i = 1, 2, ..., n$. The condition $H\mathbf{x} = 0$ can be written in terms of the system of equations :

$$\begin{cases} x_1 + x_5 + x_6 = 0 \\ x_1 + x_2 + x_6 = 0 \\ x_2 + x_3 + x_6 = 0 \\ x_1 + x_4 + x_6 = 0 \end{cases}$$

We find the following solutions :

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The minimal distance between the codewords is 3 thus one can only correct single errors.

(b) To correct 1 error and to detect two errors, the minimal Hamming distance has to be $d = 4$. This corresponds to having the 3 linearly independent columns of H (see exercise 6-3). In particular, the last column of H can neither be equal to another column of H nor equal to a linear combination of any two columns. There are 5 columns given and the number of linear combinations of two of them is $\frac{5\cdot(5-1)}{2} = 10$. Remember that the column with all zeros is also forbidden, so this gives us $5+10+1 = 16$ different forbidden columns. Because the $h_{i,6}$ are bits, only $2^4 = 16$ different combinations are possible, and they are all excluded by the argument above. Therefore the Hamming distance d can not be 4.

6-2.

(a) The first 3 columns of G_1 are linearly independent and correspond to $k = 3$ bits of information, while the last two columns correspond to $m = 2$ parity bits. There are thus $2^k = 8$ codewords. The Hamming matrix has 2 rows (number of parity bits) and 5 columns (lengths of the codewords) and contains at least two linearly independent columns. H can be found by solving the equation $H\mathbf{w} = 0$ for a codeword $\mathbf{w}$. We can try a solution of the form :

$$\begin{pmatrix} h_{1,1} & h_{1,2} & h_{1,3} & 1 & 0 \\ h_{2,1} & h_{2,2} & h_{2,3} & 0 & 1 \end{pmatrix}$$

We find :

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

Since all codewords have at least two bits the minimal distance between the words is $d = 2$. This code allows the detection of single errors without correcting them.

(b)

In this case G_2, $n = 4$ and $k = 1$. The number of codewords is $2^k = 2$. $m = n - k = 3$, which corresponds to 3 parity bits. Therefore 3 columns of H can be written as the identity matrix. We find :

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

The code has only two codewords :

$$\mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

The Hamming distance is $d = 4$. This is a repetition code which corrects single errors and detects double errors.

Remark : The transmission rate is given by $R = k/n$. We observe that $R_{G_1} = 3/5$, but it does not permit to correct any error (can only detect single errors). In contrary $R_{G_2} = 1/4$ but permits to correct single errors and detect double errors. There is a tradoff between the transmission rate and the possibility to correct errors.

6-3.

A Hamming code corrects up to $e-1$ errors and detects (but not necessarily correct) up to e errors iff the minimal Hamming distance is $d = 2e$. It remains to show that $d = 2e$ is equivalent to the requirement that all sets of $2e-1$ columns of the parity matrix H are linearly independent.

If $\mathbf{w}_i$ is a codeword one has $H\mathbf{w}_i = \mathbf{0}$. Let $\mathbf{w}_j = \mathbf{w}_k + \mathbf{z}$, thus the number of 1s in $\mathbf{z}$ (the weight $W(\mathbf{z})$) is equal to the distance d_{jk} between $\mathbf{w}_j$ and $\mathbf{w}_k$. One has $H\mathbf{w}_j = H\mathbf{w}_k + H\mathbf{z} = 0$ and $H\mathbf{w}_k = 0$ since it is a codeword. One obtains $H\mathbf{z} = 0$. This only holds if there are d_{jk} columns of H that are linearly dependent.

But the minimal Hamming distance is $d = min_{ij}\{d_{ij}\}$, that means : d is the smallest number of linearly **dependent** columns in H. This again means that one requires all sets of $d - 1$ columns of H to be linearly independent. Thus, $d = 2e$ is equivalent to require that all sets of $2e - 1$ columns of H are linearly independent.

6-4.

Let $\mathbf{w}_i$ be a codeword and $W(\mathbf{w}_i)$ its weight. The weight can be written as $W(\mathbf{w}_i) = d(\mathbf{w}_i, \mathbf{0})$. We are going to use the distance property : $d(\mathbf{w}_i, \mathbf{w}_j) = d(\mathbf{w}_i - \mathbf{w}_k, \mathbf{w}_j - \mathbf{w}_k)$. By replacing k by j one obtains $d(\mathbf{w}_i, \mathbf{w}_j) = d(\mathbf{w}_i - \mathbf{w}_j, \mathbf{0})$, $\mathbf{w}_i - \mathbf{w}_j$ which is also a codeword because all codewords form a group.

Proof.

The definition of the minimal Hamming distance $d = min_{i,j}\{d(\mathbf{w}_i, \mathbf{w}_j)\}$ implies in particular $\mathbf{w}_j = \mathbf{0}$ ($\mathbf{0}$ is always a codeword) $\forall i : d(\mathbf{w}_i, \mathbf{0}) = W(\mathbf{w}_i) \geq d$. But it also implies that there is at least one couple (l, m) such that $d(\mathbf{w}_l, \mathbf{w}_m) = d$ (since there are at least two codewords which attain the minimum), which offers the possibility to write $d(\mathbf{w}_l - \mathbf{w}_m, \mathbf{0}) = d$. There is thus a

k $\mathbf{w}_k = \mathbf{w}_l - \mathbf{w}_m$ satisfies $W(\mathbf{w}_k) = d$.

Inversely, if on assumes that $W(\mathbf{w}) \geq d$ then $\forall i, j\ \exists k : \mathbf{w}_k = \mathbf{w}_i - \mathbf{w}_j$ such that $d(\mathbf{w}_i, \mathbf{w}_j) = d(\mathbf{w}_i - \mathbf{w}_j, \mathbf{0}) = W(\mathbf{w}_k) \geq d$. As there is a z such that $W(\mathbf{w}_z) = d$, one also has a pair (a, b) such that $d(\mathbf{w}_a, \mathbf{w}_b) = d(\mathbf{w}_a - \mathbf{w}_b, \mathbf{0}) = W(\mathbf{w}_z) = d$. On obtains thus $d = min_{i,j}\{d(\mathbf{w}_i, \mathbf{w}_j)\}$.