Solutions to Exercise Sheet 6

Exercise 1. Suppose $\mathbf{w}$ is a codeword for this code, then from the definition of the parity matrix H, we have

$$H\mathbf{w} = 0. \tag{1}$$

Because $\mathbf{w}$ is a column vector with binary entries, we can rewrite the matrix product (1) as

$$\sum_i H_i \cdot w_i = 0, \tag{2}$$

where H_i denotes the i-th column of H and w_i is the i-th entry (i.e. row) of $\mathbf{w}$.

Since the code corrects up to $e-1$ errors and detects up to e errors, the minimum weight of the code is $d = 2e$, which means that there exists a codeword $\mathbf{w}_d$ such that the number of nonzero entries in $\mathbf{w}_d$ is $2e$. From (2), we can deduce that for $\mathbf{w}_d$, there are $2e$ columns in H (corresponding to the nonzero entries of $\mathbf{w}_d$) which are linearly dependent. Because $2e$ is the minimum weight (therefore also the minimum number of linearly dependent columns), all sets of $2e-1$ columns must be linearly independent.

Exercise 2.

(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 minimum distance between the codewords is 3, therefore single errors can be corrected.

(b) To correct one error and detect two, the minimum Hamming distance has to be $d = 4$. This corresponds to having 3 linearly independent columns in H (see Exercise 1). In particular, the last column of H can neither be equal to another column of H nor to a linear combination of any two columns. There are 5 columns and the number of possible 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 entries $h_{i,6}$ 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.

Exercise 3.

(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 $2^k = 8$ 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 minimum distance between the words is $d = 2$. This code detects single errors without correcting them.

(b) In this case, $n = 4$ 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 two 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 trade off between the transmission rate and the possibility to correct errors.

Exercise 4. The *if* part is trivial. We will only prove the *only if* part. First note that because the codewords form a closed linear subspace, the sum (modulo 2) and the difference (modulo 2) of two codewords are also codewords:

$$\forall i, j, i \neq j, \exists l, m, \mathbf{w}_i + \mathbf{w}_j = \mathbf{w}_l, \mathbf{w}_i - \mathbf{w}_j = \mathbf{w}_m. \tag{3}$$

The phrase "the minimum Hamming distance is d" can be expressed mathematically as

$$d = \min_{i,j} d(\mathbf{w}_i, \mathbf{w}_j), \tag{4}$$

where $d(\mathbf{a}, \mathbf{b})$ gives the Hamming distance between $\mathbf{a}$ and $\mathbf{b}$.

By using the linearity condition (3), we can rewrite (4) as

$$\begin{aligned} d &= \min_{i,j} d(\mathbf{w}_i, \mathbf{w}_j) && (5) \\ &= \min_{i,j} d(\mathbf{w}_i - \mathbf{w}_j, 0) && (6) \\ &= \min_{k} d(\mathbf{w}_k, 0). && (7) \end{aligned}$$

The last line shows that d is defined as the minimum Hamming distance between any codeword and 0, so there is at least one codeword which has this distance and all other codewords having this distance as a lower bound.