Chapter 5

Matrix Description for Linear Block Codes

In the previous chapters we were able to describe the maximum-likelihood decoding (MLD) method and the information theoretical bounds as well as the performance of block codes and the computation of the error probabilities of coded transmission by using only an enumeration of the block code.

For a reasonable and practical definition of codes of a large block length and for the easy processing by the encoder and decoder a compact description with matrices is required. This and other important concepts such as the dual code and the method of syndrome decoding will be discussed in this chapter.

However, to obtain really powerful and useful coding techniques a further structure than just linearity is required. The resulting cyclic codes and their representation by polynomials will be discussed in the next chapter.

5.1 The Generator Matrix

5.1.1 Basic Properties

Given a linear $(n, k)_q$ code $\mathcal{C}$, according to Definition 3.3 the code $\mathcal{C}$ forms a vector space with q^k words or vectors. The set $\mathcal{C}$ can also be interpreted as a subspace of the vector space $\mathbb{F}_q^n$ of all q^n possible words. The basics of vector spaces are given in Section A.5. In particular, every *linear combination* of codewords is another codeword, i.e.,

$$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_l \in \mathcal{C}, \quad \alpha_1, \ldots, \alpha_l \in \mathbb{F}_q \quad \Longrightarrow \quad \sum_{i=1}^{l} \alpha_i \boldsymbol{a}_i \in \mathcal{C}.$$

The maximum number of linearly independent words corresponds to the *dimension* of the vector space or of the code and is denoted $\dim(\mathcal{C})$. Obviously $\dim(\mathcal{C}) \leq n$ and $\dim(\mathcal{C}) = k$ since a vector space of dimension k over $\mathbb{F}_q$ contains

q^k words. Every selection of $\dim(\mathcal{C})$ linearly independent words forms a *basis* for the code.

$\mathbb{F}_q^n$ denotes the words or vectors of length n with elements in $\mathbb{F}_q$. Correspondingly, $\mathbb{F}_q^{k,n}$ denotes the set of all (k,n)-dimensional matrices with elements in $\mathbb{F}_q$. All vectors are to be understood as row vectors so that formally we have $\mathbb{F}_q^n = \mathbb{F}_q^{1,n}$.

Definition 5.1. *A matrix $\boldsymbol{G} \in \mathbb{F}_q^{k,n}$ is called a* generator matrix *for the linear $(n,k)_q$ code $\mathcal{C}$, if*

$$\mathcal{C} = \Big\{ \boldsymbol{u}\boldsymbol{G} \;\Big|\; \boldsymbol{u} \in \mathbb{F}_q^k \Big\}. \tag{5.1.1}$$

The generator matrix creates the code and simultaneously provides an encoding rule with which the codewords are generated as follows:

$$\begin{aligned}(a_0,\ldots,a_{n-1}) &= (u_0,\ldots,u_{k-1}) \cdot \begin{pmatrix} g_{0,0} & \cdots & g_{0,n-1} \\ \vdots & & \vdots \\ g_{k-1,0} & \cdots & g_{k-1,n-1} \end{pmatrix} \\ &= (u_0 g_{0,0} + \cdots + u_{k-1} g_{k-1,0}, \ldots, u_0 g_{0,n-1} + \cdots + u_{k-1} g_{k-1,n-1}) \\ &= u_0 (g_{0,0},\ldots,g_{0,n-1}) + \cdots + u_{k-1}(g_{k-1,0},\ldots,g_{k-1,n-1}).\end{aligned}$$

The rows of the generator matrix are to be linearly independent and therefore form a basis for $\mathcal{C}$ with $\dim(\mathcal{C}) = k$.

The rows of the generator matrix are obviously codewords for the unit vectors as information words. Every code generated by a generator matrix is linear, because $\boldsymbol{a}_i = \boldsymbol{u}_i \boldsymbol{G}$ implies that

$$\sum_{i=1}^{l} \alpha_i \boldsymbol{a}_i = \sum_{i=1}^{l} \alpha_i \Big(\boldsymbol{u}_i \boldsymbol{G} \Big) = \underbrace{\Big(\sum_{i=1}^{l} \alpha_i \boldsymbol{u}_i \Big)}_{\boldsymbol{u}} \cdot \boldsymbol{G} = \boldsymbol{u} \cdot \boldsymbol{G}.$$

The row rank (column rank) of a matrix corresponds to the number of linearly independent rows (columns) or to the dimension of the vector space created by them. The row and column rank are equal and are therefore simply called *rank*. Since $n > k$ the columns of the generator matrix are, of course, linearly dependent.

Example 5.1. Consider the $(7,4)_2$ Hamming code with the enumerating description of $\mathcal{C}$ according to Example 1.2. A suitable generator matrix is given by

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

with the systematic encoder

$$(u_0, u_1, u_2, u_3) \mapsto (u_0, u_1, u_2, u_3, u_1 + u_2 + u_3, u_0 + u_2 + u_3, u_0 + u_1 + u_3).$$

Here is a simple example with numbers:

$$(1, 1, 0, 1) \mapsto (1, 1, 0, 1, 0 + 1 + 1, 1 + 0 + 1, 1 + 1 + 1) = (1, 1, 0, 1, 0, 0, 1).$$

The maximum number of linearly independent rows is 4, as can be recognized from the first four components (the other three components are of no importance for the dimension, but only for the minimum distance). The maximum number of linearly independent columns is 4 as well since the first four columns are obviously independent and the last three columns are a linear combination of the first four. The rows of the generator matrix form the basis for the code, i.e., 1000011, 0100101, 0010110 and 0001111 are four linearly independent codewords. ■

5.1.2 Elementary Row Operations

Theorem 5.1 (Elementary Row Operations). *Let $\boldsymbol{G}$ be a generator matrix for the $(n, k)_q$ code $\mathcal{C}$. The following so-called* elementary row operations *are allowed to be used on $\boldsymbol{G}$ without actually changing the code.*

- *The interchange of two rows.*
- *Multiplication of a row by a non-zero scalar.*
- *Addition of a non-zero scalar multiple of another row.*

Thus the four generator matrices

$$\begin{pmatrix} \vdots \\ \boldsymbol{G}_i \\ \vdots \\ \boldsymbol{G}_j \\ \vdots \end{pmatrix} \quad \begin{pmatrix} \vdots \\ \boldsymbol{G}_j \\ \vdots \\ \boldsymbol{G}_i \\ \vdots \end{pmatrix} \quad \begin{pmatrix} \vdots \\ \alpha\boldsymbol{G}_i \\ \vdots \\ \boldsymbol{G}_j \\ \vdots \end{pmatrix} \quad \begin{pmatrix} \vdots \\ \boldsymbol{G}_i \\ \vdots \\ \boldsymbol{G}_j + \alpha\boldsymbol{G}_i \\ \vdots \end{pmatrix}. \tag{5.1.2}$$

create the same code. These elementary row operations (and possibly additional column permutations) can transform $\boldsymbol{G}$ into the row-echelon form *(also called standard form or systematic form), denoted*

$$\boldsymbol{G} = \Big(\boldsymbol{I}_k \,\Big|\, \boldsymbol{P}\Big), \tag{5.1.3}$$

where $\boldsymbol{I}_k \in \mathbb{F}_q^{k,k}$ is the (k, k)-dimensional identity matrix and $\boldsymbol{P} \in \mathbb{F}_q^{k,n-k}$.

The row-echelon form corresponds to a systematic encoder which can thus always be obtained. If column permutations have to be used to achieve the form (4.1.3), the code is modified and therefore it is no longer identical to the original code, but called equivalent (see Definition 1.6). However, the distance properties and the weight distributions of equivalent codes are identical.

The proof of this theorem for Galois fields is similar to the proof for real numbers in basic linear algebra. If $\boldsymbol{G}$ did not have the maximum rank k, the elementary row operations would create an all-zero row with the consequence that $|\mathcal{C}| \leq q^{k-1}$, however, this would contradict the elementary definition of a block code (where q^k information words are to be mapped to q^k different codewords).

Example 5.2. Again consider the $(7,4)_2$ Hamming code. According to Examples 4.1 or 1.2, 1111111, 1011010, 0110011 and 1110000 are codewords. If these codewords are linearly independent (which can not easily be seen), they form a basis. The generator matrix

$$\boldsymbol{G}_1 = \begin{pmatrix} 1\,1\,1\,1\,1\,1\,1 \\ 1\,0\,1\,1\,0\,1\,0 \\ 0\,1\,1\,0\,0\,1\,1 \\ 1\,1\,1\,0\,0\,0\,0 \end{pmatrix}$$

defined by this basis must create an identical code. We will now try to transform $\boldsymbol{G}_1$ into $\boldsymbol{G}$ of Example 4.1 by using elementary row operations. Add row 1 to row 2 and row 4:

$$\boldsymbol{G}_2 = \begin{pmatrix} 1\,1\,1\,1\,1\,1\,1 \\ 0\,1\,0\,0\,1\,0\,1 \\ 0\,1\,1\,0\,0\,1\,1 \\ 0\,0\,0\,1\,1\,1\,1 \end{pmatrix}.$$

Add row 2 to row 1 and row 3:

$$\boldsymbol{G}_3 = \begin{pmatrix} 1\,0\,1\,1\,0\,1\,0 \\ 0\,1\,0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1\,1\,0 \\ 0\,0\,0\,1\,1\,1\,1 \end{pmatrix}.$$

Add row 3 to row 1:

$$\boldsymbol{G}_4 = \begin{pmatrix} 1\,0\,0\,1\,1\,0\,0 \\ 0\,1\,0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1\,1\,0 \\ 0\,0\,0\,1\,1\,1\,1 \end{pmatrix}.$$

Add row 4 to row 1:

$$\boldsymbol{G}_5 = \begin{pmatrix} 1\,0\,0\,0\,0\,1\,1 \\ 0\,1\,0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1\,1\,0 \\ 0\,0\,0\,1\,1\,1\,1 \end{pmatrix}.$$

This is the original generator matrix of Example 4.1. ■

The minimum distance is obviously smaller than or equal to the minimum Hamming weight of all rows of the generator matrix since the rows are codewords. However, the following example will show that $d_{\min}$ can also be smaller

than the minimum row weight.

$$\boldsymbol{G} = \begin{pmatrix} 1\,1\,1\,0 \\ 0\,1\,1\,1 \end{pmatrix} \quad \text{creates} \quad \mathcal{C} = \{0000, 1110, 0111, 1001\}.$$

5.1.3 Linear Random Codes

The systematic form of the generator matrix can be used to define what is called a linear random code. The average weight distribution computed in the next theorem has already been quoted in Section 3.5 and compared to the weight distribution of general random codes. The readers only interested in the standard theory can skip this theorem without any loss.

Theorem 5.2 (Linear Random Codes). *For the matrix $\boldsymbol{P} \in \mathbb{F}_2^{k,n-k}$ all $k(n-k)$ coefficients are chosen statistically independent with a uniform distribution. Now, the generator matrix $\boldsymbol{G} = \left(\boldsymbol{I}_k \middle| \boldsymbol{P}\right)$ creates a linear $(n,k)_2$ random code, its average weight distribution is approximately binomial:*

$$E(A_r) = \left\{ \begin{array}{ll} 1 & \textit{for} \quad r = 0 \\ 2^{k-n}\left[\binom{n}{r} - \binom{n-k}{r}\right] & \textit{for} \quad 1 \le r \le n-k \\ 2^{k-n}\binom{n}{r} & \textit{for} \quad n-k < r \le n \end{array} \right\}.$$

Proof (in part taken from an idea in [205]). By using the binomial formula (A.2.2) the property (3.5.6) of the form $\sum_{r=0}^{n} E(A_r) = 2^k$ can be verified. According to the premises, each one of the $2^{k(n-k)}$ possible matrices $\boldsymbol{P}$ is chosen with the same probability $2^{-k(n-k)}$. The codewords are denoted $\boldsymbol{a} = \boldsymbol{u}\boldsymbol{G} = (\boldsymbol{u}, \boldsymbol{u}\boldsymbol{P})$.

Of the 2^k possible information words only the all-zero word leads to $w_H(\boldsymbol{a}) = 0$, therefore $A_0 = 1$ is always valid, thus $E(A_0) = 1$.

The number of all possible words of weight r is $\binom{n}{r}$ and the number of all possible words $(\boldsymbol{u}, \boldsymbol{p})$ with $\boldsymbol{u} = \boldsymbol{0}$ of weight r is $\binom{n-k}{r}$. Therefore the number of all words $(\boldsymbol{u}, \boldsymbol{p})$ with $\boldsymbol{u} \neq \boldsymbol{0}$ of weight r is exactly $\binom{n}{r} - \binom{n-k}{r} = B_r$. Thus B_r is the number of all words in $\mathbb{F}_2^n$ of weight r where the first k components are not identical to the all-zero word.

Now, let $r > 0$, the probability of $(\boldsymbol{u}, \boldsymbol{p})$ being a codeword with $\boldsymbol{u} \neq \boldsymbol{0}$ and a random choice of $\boldsymbol{P}$ is

$$\begin{aligned} P(\,(\boldsymbol{u}, \boldsymbol{p}) \in \mathcal{C} \mid \boldsymbol{u} \neq \boldsymbol{0}) &= \frac{\text{number of matrices } \boldsymbol{P} \text{ with } \boldsymbol{p} = \boldsymbol{u}\boldsymbol{P} \text{ and } \boldsymbol{u} \neq \boldsymbol{0}}{\text{number of all matrices } \boldsymbol{P}} \\ &= \frac{2^{(k-1)(n-k)}}{2^{k(n-k)}} = 2^{-(n-k)}, \end{aligned}$$

since for $\boldsymbol{u} \neq \boldsymbol{0}$, $\boldsymbol{p} = \boldsymbol{uP}$ is a system of $n-k$ equations with $k(n-k)$ unknowns. Its solution space is $k(n-k)-(n-k) = (k-1)(n-k)$-dimensional, therefore $2^{(k-1)(n-k)}$ is the number of matrices $\boldsymbol{P}$ with $\boldsymbol{p} = \boldsymbol{uP}$. Finally, for $r > 0$,

$$E(A_r) = \sum_{\substack{(\boldsymbol{u},\boldsymbol{p})\in\mathbb{F}_2^n \\ w_H((\boldsymbol{u},\boldsymbol{p}))=r}} P(\ (\boldsymbol{u},\boldsymbol{p}) \in \mathcal{C}) = B_r \cdot 2^{-(n-k)}$$

which provides the formula for $E(A_r)$ we were to prove. ■

5.2 The Parity-Check Matrix

5.2.1 Basic Properties

A code can not only be defined by the generator matrix $\boldsymbol{G}$ but also by the parity-check matrix $\boldsymbol{H}$, however, the two matrices can be derived from each other. The parity-check matrix is also used to define the main concept of syndromes, which is discussed in Section 4.6.

Definition 5.2. *A matrix $\boldsymbol{H} \in \mathbb{F}_q^{n-k,n}$ is called a* parity-check matrix *for the linear $(n,k)_q$ code $\mathcal{C}$, if*

$$\mathcal{C} = \Big\{ \boldsymbol{a} \in \mathbb{F}_q^n \ \Big|\ \boldsymbol{aH}^T = \boldsymbol{0} \Big\}, \tag{5.2.1}$$

where the all-zero word is of length $n-k$. Thus $\boldsymbol{aH}^T = \boldsymbol{0}$ for the codewords $\boldsymbol{a} \in \mathcal{C}$ and $\boldsymbol{aH}^T \neq \boldsymbol{0}$ for all other words $\boldsymbol{a} \notin \mathcal{C}$. Therefore $\mathcal{C}$ is also called a null space *of $\boldsymbol{H}$ or* row space *of $\boldsymbol{G}$.*

The parity-check matrix, like the generator matrix, is not uniquely defined by the code.

Theorem 5.3. *The parity-check matrix $\boldsymbol{H}$ has the maximum possible rank $n-k$ and elementary row operations are allowed for $\boldsymbol{H}$, but not for $\boldsymbol{H}^T$.*

Proof (a short version). If the rank of $\boldsymbol{H}$ was smaller than $n-k$, an all-zero row (i.e., an all-zero column in $\boldsymbol{H}^T$) could be created by using elementary row operations. The null space of $\boldsymbol{H}$ would then have at least the dimension $k+1$ which would be greater than the dimension k of the code. Therefore $\boldsymbol{H}$ would not be a parity-check matrix according to Definition 4.2, leading to a contradiction. Hence, the rank of $\boldsymbol{H}$ is $n-k$. ■

Theorem 5.4. *Let the linear $(n,k)_q$ code $\mathcal{C}$ be created by the generator matrix $\boldsymbol{G} \in \mathbb{F}_q^{k,n}$.*

(1) The matrix $\boldsymbol{H} \in \mathbb{F}_q^{n-k,n}$ *is a parity-check matrix for* $\mathcal{C}$ *if and only if*

$$\boldsymbol{H} \neq \boldsymbol{0} \quad \text{and} \quad \boldsymbol{G}\boldsymbol{H}^T = \boldsymbol{0}. \tag{5.2.2}$$

(2) If $\boldsymbol{G} = \left(\boldsymbol{I}_k \;\middle|\; \boldsymbol{P}\right)$ *is a systematic generator matrix with the identity matrix* $\boldsymbol{I}_k \in \mathbb{F}_q^{k,k}$ *and the matrix* $\boldsymbol{P} \in \mathbb{F}_q^{k,n-k}$*, then a parity-check matrix is given by*

$$\boldsymbol{H} = \left(-\boldsymbol{P}^T \;\middle|\; \boldsymbol{I}_{n-k}\right). \tag{5.2.3}$$

Proof. (1 "⇒"): let $\boldsymbol{H}$ be a parity-check matrix. $\boldsymbol{H} \neq \boldsymbol{0}$ is apparent, otherwise $\boldsymbol{H}$ would classify all words as codewords. For every codeword $\boldsymbol{a} = \boldsymbol{u}\boldsymbol{G}$,

$$\boldsymbol{0} = \boldsymbol{a}\boldsymbol{H}^T = (\boldsymbol{u}\boldsymbol{G})\boldsymbol{H}^T = \boldsymbol{u}(\boldsymbol{G}\boldsymbol{H}^T).$$

Since this equation is valid for all vectors $\boldsymbol{u}$, $\boldsymbol{G}\boldsymbol{H}^T = \boldsymbol{0}$.

(1 "⇐"): let $\boldsymbol{G}\boldsymbol{H}^T = \boldsymbol{0}$ and let $\boldsymbol{a} = \boldsymbol{u}\boldsymbol{G}$ be a codeword, then

$$\boldsymbol{a}\boldsymbol{H}^T = (\boldsymbol{u}\boldsymbol{G})\boldsymbol{H}^T = \boldsymbol{u}(\boldsymbol{G}\boldsymbol{H}^T) = \boldsymbol{u}\boldsymbol{0} = \boldsymbol{0},$$

thus $\boldsymbol{H}$ is a parity-check matrix.

(2) Let $\boldsymbol{G} = \left(\boldsymbol{I}_k \;\middle|\; \boldsymbol{P}\right)$ and $\boldsymbol{H} = \left(-\boldsymbol{P}^T \;\middle|\; \boldsymbol{I}_{n-k}\right)$, then

$$\boldsymbol{G}\boldsymbol{H}^T = \left(\boldsymbol{I}_k \;\middle|\; \boldsymbol{P}\right) \cdot \begin{pmatrix} -\boldsymbol{P} \\ \boldsymbol{I}_{n-k} \end{pmatrix} = -\boldsymbol{I}_k \cdot \boldsymbol{P} + \boldsymbol{P} \cdot \boldsymbol{I}_{n-k} = \boldsymbol{0},$$

where $\boldsymbol{0} \in \mathbb{F}_q^{k,n-k}$. According to (1), $\boldsymbol{H}$ is a parity-check matrix. ■

Example 5.3. (1) For the $(7,4)_2$ Hamming code, we already know $\boldsymbol{G}$ from Example 4.1, $\boldsymbol{H}$ is created according to (4.2.3):

$$\boldsymbol{G} = \begin{pmatrix} 1\,0\,0\,0 & 0\,1\,1 \\ 0\,1\,0\,0 & 1\,0\,1 \\ 0\,0\,1\,0 & 1\,1\,0 \\ 0\,0\,0\,1 & 1\,1\,1 \end{pmatrix}, \qquad \boldsymbol{H} = \begin{pmatrix} 0\,1\,1\,1 & 1\,0\,0 \\ 1\,0\,1\,1 & 0\,1\,0 \\ 1\,1\,0\,1 & 0\,0\,1 \end{pmatrix}.$$

(2) For the $(n,1)_2$ repetition code, $\boldsymbol{G}$ is apparent, implying $\boldsymbol{H}$:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \end{pmatrix}, \qquad \boldsymbol{H} = \begin{pmatrix} 1 & 1 & & & & \\ 1 & & 1 & & & \\ \vdots & & & \ddots & & \\ 1 & & & & 1 & \\ 1 & & & & & 1 \end{pmatrix}.$$

(3) For the $(n, n-1)_2$ parity-check code, $\boldsymbol{H}$ is apparent, implying $\boldsymbol{G}$:

$$\boldsymbol{G} = \begin{pmatrix} 1 & 1 & & & \\ 1 & & 1 & & \\ \vdots & & & \ddots & \\ 1 & & & & 1 \\ 1 & & & & & 1 \end{pmatrix}, \qquad \boldsymbol{H} = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \end{pmatrix}.$$

The generator matrix $\boldsymbol{G}$ provides an encoding rule where the parity-check bit is placed in front. The encoding rule can not be recognized from the code itself (see also Example 1.1). ■

5.2.2 Calculation of the Minimum Distance

As already stated the minimum distance can not be easily calculated from the generator matrix. However, there is the following relation, as already used in the proof of the Gilbert-Varshamov bound in Theorem 3.12.

Theorem 5.5. *Let $\boldsymbol{H} \in \mathbb{F}_q^{n-k,n}$ be a parity-check matrix for the $(n, k, d_{\min})_q$ code $\mathcal{C}$. The minimum distance $d_{\min}$ is the minimum number of linearly dependent columns in $\boldsymbol{H}$, i.e., every selection of $d_{\min}-1$ columns is linearly independent and there exists at least one selection of $d_{\min}$ linearly dependent columns.*

Proof. Let $\boldsymbol{H}_0, \dots, \boldsymbol{H}_{n-1}$ denote the columns of $\boldsymbol{H}$.

(1) We are to show that there exists no selection of $d_{\min}-1$ linearly dependent columns in $\boldsymbol{H}$. Assume the opposite: there exists a selection $\boldsymbol{H}_{i_1}, \dots, \boldsymbol{H}_{i_{d_{\min}-1}}$ of $d_{\min}-1$ linearly dependent columns. Hence, there exists a word $\boldsymbol{y} = (y_0, \dots, y_{n-1})$ with $w_H(\boldsymbol{y}) \le d_{\min}-1$ and $\boldsymbol{0} = \sum_{j=1}^{d_{\min}-1} y_{i_j}\boldsymbol{H}_{i_j} = \sum_{\nu=0}^{n-1} y_\nu \boldsymbol{H}_\nu = \boldsymbol{y}\boldsymbol{H}^T$, implying that $\boldsymbol{y}$ is a codeword, which contradicts the definition of the minimum distance. Thus the assumption was wrong.

(2) We are to show that there exists a selection of $d_{\min}$ linearly dependent columns in $\boldsymbol{H}$. There exists a codeword $\boldsymbol{a} = (a_0, \dots, a_{n-1})$ with $w_H(\boldsymbol{a}) = d_{\min}$. Since $\boldsymbol{0} = \boldsymbol{a}\boldsymbol{H}^T = \sum_{\nu=0}^{n-1} a_\nu h_\nu$, the $d_{\min}$ indices ν with $a_\nu \neq 0$ determine a selection of $d_{\min}$ linearly dependent columns in $\boldsymbol{H}$. ■

This theorem, by the way, directly implies the Singleton bound as stated in Theorem 3.7. Since the columns of the parity-check matrix are of length $n-k$, there can only exist a maximum of $n-k$ linearly independent columns. Therefore $d_{\min} - 1 \le n-k$.

Example 5.4. Consider the parity-check matrix of the $(7,4)_2$ Hamming code in Example 4.3(1). No column is a multiple of any other column, therefore each

two columns are linearly independent. Since the first column is the sum of the last two, there are three linearly dependent columns in $\boldsymbol{H}$, therefore $d_{\min} = 3$. (However, the column rank and the rank of $\boldsymbol{H}$ are 3, since there are also three linearly independent columns.) ■

5.3 Dual Codes and MacWilliams Identity

5.3.1 Basic Properties of Dual Codes

Example 4.3(2,3) showed that the repetition code and the parity-check code emerge from each other by swapping $\boldsymbol{G}$ and $\boldsymbol{H}$. Generally, by swapping the generator matrix and the parity-check matrix of the code $\mathcal{C}$, the so-called dual code $\mathcal{C}^\perp$ is defined.

Definition 5.3. *Let $\boldsymbol{G} \in \mathbb{F}_q^{k,n}$ be the generator matrix and $\boldsymbol{H} \in \mathbb{F}_q^{n-k,n}$ be the parity-check matrix of the $(n,k)_q$ code $\mathcal{C}$. Using $\boldsymbol{H}$ as the generator matrix or $\boldsymbol{G}$ as the parity-check matrix, an $(n, n-k)_q$ code is created which is called the* dual code $\mathcal{C}^\perp$.

For arbitrary $\boldsymbol{a} = \boldsymbol{uG} \in \mathcal{C}$ and $\boldsymbol{b} = \boldsymbol{vH} \in \mathcal{C}^\perp$ the scalar product is zero, since $\boldsymbol{ab}^T = \boldsymbol{uGH}^T\boldsymbol{v}^T = \boldsymbol{u0v}^T = 0$. This orthogonality is also denoted $\boldsymbol{a} \perp \boldsymbol{b}$ or $\mathcal{C} \perp \mathcal{C}^\perp$ and dual codes are also called orthogonal codes.

Theorem 5.6. *The following notations apply for dual codes.*

$$\begin{aligned} \mathcal{C}^\perp &= \Big\{ \boldsymbol{b} \in \mathbb{F}_q^n \;\Big|\; \boldsymbol{b} \perp \boldsymbol{a} \text{ for all } \boldsymbol{a} \in \mathcal{C} \Big\} \\ &= \Big\{ (b_0, \ldots, b_{n-1}) \;\Big|\; \sum_{i=0}^{n-1} b_i a_i = 0 \text{ for all } (a_0, \ldots, a_{n-1}) \in \mathcal{C} \Big\}. \end{aligned} \tag{5.3.1}$$

Furthermore $\mathcal{C}^{\perp\perp} = \mathcal{C}$ and $\dim(\mathcal{C}^\perp) = n - \dim(\mathcal{C})$.

Proof. The statement for the dimension of $\mathcal{C}^\perp$ is implied by $\operatorname{rank}(\boldsymbol{G}) = k$ and $\operatorname{rank}(\boldsymbol{H}) = n - k$. According to the definition of dual codes, $\mathcal{C}^{\perp\perp} = \mathcal{C}$.

Let $\mathcal{C}'$ denote the right side of (4.3.1). Since $\boldsymbol{b} \in \mathcal{C}^\perp$ is orthogonal to all $\boldsymbol{a} \in \mathcal{C}$ it follows that $\boldsymbol{b} \in \mathcal{C}'$ or $\mathcal{C}^\perp \subseteq \mathcal{C}'$.

In reverse, let $\boldsymbol{b} \in \mathcal{C}'$, i.e., $\boldsymbol{b}$ is orthogonal to all $\boldsymbol{a} = \boldsymbol{uG} \in \mathcal{C}$, thus $0 = \boldsymbol{ba}^T = \boldsymbol{bG}^T\boldsymbol{u}^T$ for all $\boldsymbol{u}$ implies that $\boldsymbol{bG}^T = \boldsymbol{0}$. Since $\boldsymbol{G}$ is a parity-check matrix for $\mathcal{C}^\perp$, $\boldsymbol{b} \in \mathcal{C}^\perp$ or $\mathcal{C}' \subseteq \mathcal{C}^\perp$. Conclusively, $\mathcal{C}^\perp = \mathcal{C}'$. ■

Example 5.5. (1) Referring to Example 4.3(2,3), the $(n,1)_2$ repetition code and the $(n, n-1)_2$ parity-check code are the dual codes of each other.

(2) The $(7,3)_2$ code

$$\begin{aligned}\mathcal{C}^{\perp} = \{ & 0000\ 000, \quad 0111\ 100, \\ & 1101\ 001, \quad 1010\ 101, \\ & 1011\ 010, \quad 1100\ 110, \\ & 0110\ 011, \quad 0001\ 111 \}\end{aligned}$$

generated by $\boldsymbol{H}$ in Example 4.3(1), is the dual code of the $(7,4)_2$ Hamming code $\mathcal{C}$. Obviously the 16 words in $\mathcal{C}$ and the 8 words in $\mathcal{C}^{\perp}$ are orthogonal to each other. ■

Theorem 5.7. *The dual code for an MDS code is again an MDS code and* $d_{\min} + d_{\min}^{\perp} = n + 2$.

Proof. Let $\mathcal{C}$ be an $(n, k, d_{\min} = n - k + 1)_q$ MDS code with the generator matrix $\boldsymbol{G}$. According to Theorem 3.8, the codeword is uniquely determined by every selection of k code symbols, therefore every selection of k columns of $\boldsymbol{G}$ is linearly independent. Since $\boldsymbol{G}$ is the parity-check matrix of the dual $(n, n-k, d_{\min}^{\perp})_q$ code $\mathcal{C}^{\perp}$, $d_{\min}^{\perp} \geq k + 1$. Using the Singleton bound of Theorem 3.7 on $\mathcal{C}^{\perp}$ results in $d_{\min}^{\perp} \leq n - (n-k) + 1 = k + 1$. Conclusively, $d_{\min}^{\perp} = k + 1$, thus $\mathcal{C}^{\perp}$ is an MDS code. ■

5.3.2 MacWilliams Identity

For some codes the weight distribution is difficult to calculate, whereas the weight distribution of the dual code is simple to calculate. However, there is the following fundamental relation between the two weight distributions.

Theorem 5.8 (MacWilliams Identity). *Let $A(Z)$ be the weight distribution of the $(n,k)_q$ code $\mathcal{C}$ and $A^{\perp}(Z)$ the weight distribution of the dual $(n, n-k)_q$ code $\mathcal{C}^{\perp}$. Then*

$$A^{\perp}(Z) = q^{-k}\Big(1 + (q-1)Z\Big)^n \cdot A\left(\frac{1-Z}{1+(q-1)Z}\right) \tag{5.3.2}$$

or described by the alternative form of the weight enumerator

$$W^{\perp}(X,Y) = q^{-k} \cdot W(X + (q-1)Y, X - Y). \tag{5.3.3}$$

For $q = 2$ this is simplified to

$$A^{\perp}(Z) = 2^{-k}(1+Z)^n \cdot A\left(\frac{1-Z}{1+Z}\right), \tag{5.3.4}$$

$$W^{\perp}(X,Y) = 2^{-k} \cdot W(X+Y, X-Y). \tag{5.3.5}$$

The inverse relations are

$$A(Z) = 2^{-(n-k)}(1+Z)^n \cdot A^\perp\left(\frac{1-Z}{1+Z}\right), \tag{5.3.6}$$

$$W(X,Y) = 2^{-(n-k)} \cdot W^\perp(X+Y, X-Y). \tag{5.3.7}$$

The time-consuming proof can be found in [17, 82, 105, 107, 123], for example. The relations in A and W can be easily derived from each other. Corresponding to $\mathcal{C}^{\perp\perp} = \mathcal{C}$,

$$\begin{aligned}
W^{\perp\perp}(X,Y) &= q^{-(n-k)} \cdot W^\perp(\underbrace{X+(q-1)Y}_{=X'}, \underbrace{X-Y}_{=Y'}) \\
&= q^{-(n-k)} \cdot q^{-k} \cdot W(X'+(q-1)Y', X'-Y') \\
&= q^{-n} \cdot W(X+(q-1)Y+(q-1)(X-Y), X+(q-1)Y-(X-Y)) \\
&= q^{-n} \cdot W(qX, qY) \\
&= W(X,Y).
\end{aligned}$$

Example 5.6. For the $(n,1)_2$ repetition code $\mathcal{C}$ with its weight enumerator $A(Z) = 1 + Z^n$ the MacWilliams identity implies that

$$\begin{aligned}
A^\perp(Z) &= 2^{-1}(1+Z)^n\left(1+\left(\frac{1-Z}{1+Z}\right)^n\right) \\
&= \frac{1}{2}\Big((1+Z)^n + (1-Z)^n\Big) \\
&= \sum_{r=0}^{n}\binom{n}{r}\frac{Z^r + (-Z)^r}{2} \\
&= \sum_{r \text{ even}}\binom{n}{r}Z^r
\end{aligned}$$

for the corresponding dual $(n, n-1)_2$ parity-check code $\mathcal{C}^\perp$. This result is already known from Example 3.8(3). ■

5.3.3 Self-Dual and Self-Orthogonal Codes

Definition 5.4. *The $(n,k)_q$ code $\mathcal{C}$ is called* self-orthogonal, *if $\mathcal{C} \subseteq \mathcal{C}^\perp$ and* self-dual, *if $\mathcal{C} = \mathcal{C}^\perp$.*

Due to the dimensions, a self-orthogonal code dictates $k \leq n-k$ and a code rate of $R \leq 1/2$. A self-dual code implies that $2k = n$ and $R = 1/2$. In a self-orthogonal code all codewords are orthogonal to each other (excluding self-orthogonality to itself).

Theorem 5.9. *Let $\mathcal{C}$ be an $(n,k)_q$ code with a generator matrix $\boldsymbol{G}$ of any form, then*

$$\mathcal{C} \text{ is self-orthogonal} \iff \boldsymbol{G}\boldsymbol{G}^T = \boldsymbol{0}.$$

For the systematic generator matrix $\boldsymbol{G} = \left(\boldsymbol{I}_k \,\middle|\, \boldsymbol{P}\right)$ and $2k = n$,

$$\mathcal{C} \text{ is self-dual} \iff \boldsymbol{P}\boldsymbol{P}^T = -\boldsymbol{I}_k.$$

Proof. "self-orthogonal, $\Rightarrow$": for the row vectors $\boldsymbol{G}_i$ of $\boldsymbol{G}$, $\boldsymbol{G}_i \in \mathcal{C} \subseteq \mathcal{C}^\perp$, of course. Since $\boldsymbol{G}$ is a parity-check matrix for $\mathcal{C}^\perp$, $\boldsymbol{G}_i\boldsymbol{G}^T = \boldsymbol{0}$. Therefore $\boldsymbol{G}\boldsymbol{G}^T = \boldsymbol{0}$.

"self-orthogonal, $\Leftarrow$": two codewords $\boldsymbol{a} = \boldsymbol{u}\boldsymbol{G}$ and $\boldsymbol{b} = \boldsymbol{v}\boldsymbol{G}$ in $\mathcal{C}$ are orthogonal to each other since $\boldsymbol{a}\boldsymbol{b}^T = \boldsymbol{u}\boldsymbol{G}\boldsymbol{G}^T\boldsymbol{v}^T = 0$. So $\boldsymbol{a} \in \mathcal{C}$ is orthogonal to $\mathcal{C}$, therefore $\boldsymbol{a} \in \mathcal{C}^\perp$ or $\mathcal{C} \subseteq \mathcal{C}^\perp$.

"self-dual": Since $\boldsymbol{G}\boldsymbol{G}^T = \left(\boldsymbol{I}_k \,\middle|\, \boldsymbol{P}\right) \cdot \begin{pmatrix} \boldsymbol{I}_k \\ \hline \boldsymbol{P}^T \end{pmatrix} = \boldsymbol{I}_k + \boldsymbol{P}\boldsymbol{P}^T$, the property $\boldsymbol{P}\boldsymbol{P}^T = -\boldsymbol{I}_k$ is equivalent to $\mathcal{C}$ being self-orthogonal. A self-dual code is obviously self-orthogonal. A self-orthogonal code satisfies $\mathcal{C} \subseteq \mathcal{C}^\perp$ and for $k = n-k$, $|\mathcal{C}| = |\mathcal{C}^\perp|$, implying that $\mathcal{C} = \mathcal{C}^\perp$ and thus the self-duality. ■

Further properties of self-orthogonal codes are discussed in [107, 123] and self-dual codes are extensively discussed in [83].

Example 5.7. **(1)** The $(4,2)_2$ code with $\boldsymbol{G} = \begin{pmatrix} 1\,0 & 1\,0 \\ 0\,1 & 0\,1 \end{pmatrix} = \boldsymbol{H}$ produces the self-dual code $\mathcal{C} = \{0000, 1010, 0101, 1111\}$ and one can verify that $\boldsymbol{G}\boldsymbol{G}^T = \boldsymbol{0}$ or $\boldsymbol{P}\boldsymbol{P}^T = -\boldsymbol{I}_2$.

(2) Consider the $(5,2)_2$ code $\mathcal{C}$ with

$$\boldsymbol{G} = \begin{pmatrix} 1\,0 & 1\,0\,0 \\ 0\,1 & 0\,0\,1 \end{pmatrix}, \qquad \boldsymbol{H} = \begin{pmatrix} 1\,0 & 1\,0\,0 \\ 0\,0 & 0\,1\,0 \\ 0\,1 & 0\,0\,1 \end{pmatrix}.$$

These matrices imply the codes

$$\mathcal{C} = \{00000, 10100, 01001, 11101\},$$
$$\mathcal{C}^\perp = \mathcal{C} \cup \{00010, 01011, 10110, 11111\}$$

Since $\mathcal{C} \subseteq \mathcal{C}^\perp$, $\mathcal{C}$ is self-orthogonal, but $\mathcal{C}^\perp$ is not self-orthogonal, of course. ■

5.4 Hamming Codes and Applications

First we introduce the general class of Hamming codes and derive their weight distribution. Then we will take a look at the dual codes of Hamming codes which form the class of simplex codes. Finally the colored hats puzzle is introduced which can be solved by using binary Hamming codes and their perfect property.

5.4.1 Hamming Codes and their Weight Distribution

Until now, only the Hamming code with parameters $(7,4)_2$ was discussed. However, the Hamming codes form a whole class of 1-error-correcting or 2-error-detecting codes with a code rate converging to 1.

Theorem 5.10. *An* $(n,k,d_{\min})_q = (n, n-r, 3)_q$ Hamming code of order r *is defined by*

$$n = \frac{q^r - 1}{q-1} = 1 + q + q^2 + \cdots + q^{r-1}. \tag{5.4.1}$$

Hamming codes exist for all orders and are perfect. The code is unique except for permutations (i.e., equivalences or permutations of columns). The code is an MDS code only for $r = 2$. *In particular for* $q = 2$ *we have a* $(2^r - 1, 2^r - r - 1, 3)_2$ *code, some examples are*

$$(3,1),\ (7,4),\ (15,11),\ (31,26),\ (63,57),\ \ldots.$$

For the binary case the columns of the parity-check matrix are the $2^r - 1$ *different binary words of length* r *(apart from the all-zero word). For the weight enumerator according to Definition 3.7,*

$$\begin{aligned} A(Z) &= \frac{1}{n+1}\Big[(1+Z)^n + n(1+Z)^{(n-1)/2}(1-Z)^{(n+1)/2}\Big] \\ &= \frac{1}{n+1}\Big[(1+Z)^n + n(1-Z)(1-Z^2)^{(n-1)/2}\Big] \end{aligned} \tag{5.4.2}$$

in the case of $q = 2$ *and calculation with respect to rational numbers. Alternatively, the coefficients of the weight distribution can be recursively computed from the recurrence relation*

$$(i+1)A_{i+1} + (n-i+1)A_{i-1} = \binom{n}{i} - A_i \tag{5.4.3}$$

with the initial values $A_0 = 1$ *and* $A_1 = A_2 = 0$.

Proof. Initially, n is an integer. The existence of corresponding parity-check matrices is apparent. Since (for $q = 2$) the columns consist of binary numbers, no column is a multiple of another. The column $1100\ldots0$ is the sum of $1000\ldots0$ and $0100\ldots0$, therefore there are three linearly dependent columns and Theorem 4.4 implies that $d_{\min} = 3$.

An MDS code must satisfy the equation $3 = d_{\min} = n - k + 1 = r + 1$ which can only be solved by $r = 2$.

A perfect code must generally satisfy $\sum_{i=0}^{t} \binom{n}{i}(q-1)^i = q^{n-k}$. For $t = 1$ and $n - k = r$ this reduces to $1 + n(q-1) = q^r$, giving an equation for n.

The weight formula (5.4.2) can be easily computed from the so-called simplex code, which is discussed further down, using the MacWilliams identity.

Obviously $A(0) = A_0 = 1$, $A_1 = A_2 = 0$ and $A(1) = 2^n/(n+1) = 2^{n-r} = 2^k$. It is also possible to derive the weight distribution directly including the proof of the recurrence relation (5.4.3). To do so, we consider the sets of arbitrary words

$$\mathcal{M}_i = \left\{ x \in \mathbb{F}_q^n \;\middle|\; \begin{array}{l} \text{there exists } \boldsymbol{a} \in \mathcal{C} \text{ with } w_H(\boldsymbol{a}) = i+1 \text{ such that} \\ \boldsymbol{x} \text{ is obtained from } \boldsymbol{a} \text{ by replacing one of the } i+1 \\ \text{one-coordinates by zero} \end{array} \right\},$$

$$\mathcal{M}_i' = \left\{ x \in \mathbb{F}_q^n \;\middle|\; \begin{array}{l} \text{there exists } \boldsymbol{a} \in \mathcal{C} \text{ with } w_H(\boldsymbol{a}) = i-1 \text{ such that } \boldsymbol{x} \\ \text{is obtained from } \boldsymbol{a} \text{ by replacing one of the } n-i+1 \\ \text{zero-coordinates by one} \end{array} \right\}.$$

Apparently, $|\mathcal{M}_i| = (i+1)A_{i+1}$, $|\mathcal{M}_i'| = (n-i+1)A_{i-1}$ and $w_H(\boldsymbol{x}) = i$ for $\boldsymbol{x} \in \mathcal{M}_i \cup \mathcal{M}_i'$. Furthermore $\mathcal{M}_i \cap \mathcal{M}_j = \mathcal{M}_i' \cap \mathcal{M}_j' = \mathcal{M}_i \cap \mathcal{M}_j' = \emptyset$ for $i \neq j$. Observing $d_{\min} = 3$, $\mathcal{M}_i \cap \mathcal{M}_i' = \emptyset$. Since $\mathcal{C}$ is a perfect code,

$$\begin{aligned} \biguplus_{i=0}^{n} (\mathcal{M}_i \uplus \mathcal{M}_i') &= \biguplus_{i=0}^{n} \{x \in \mathbb{F}_q^n \mid \text{there exists } \boldsymbol{a} \in \mathcal{C} \text{ with } d_H(\boldsymbol{a}, \boldsymbol{x}) = 1\} \\ &= \mathbb{F}_q^n \setminus \mathcal{C}, \end{aligned}$$

hence (5.4.3) is valid for all $i \in \mathbb{Z}$ (let $A_i = 0$ for $i < 0$ or $i > n$). If we multiply the relation (5.4.3) with Z^i and sum over i, we find

$$\underbrace{\sum_i (i+1)A_{i+1}Z^i}_{A'(Z)} + \underbrace{\sum_i (n-i+1)A_{i-1}Z^i}_{nZA(Z) - Z^2A'(Z)} = \underbrace{\sum_i \binom{n}{i} Z^i}_{(Z+1)^n} - \underbrace{\sum_i A_i Z^i}_{A(Z)}. \tag{5.4.4}$$

Since $A_0 = 1$, this differential equation has a unique solution given by (5.4.2) [82], but it is rather lengthy to check this. However, it is more convenient to derive (5.4.2) from the simplex code using the MacWilliams identity. ■

Example 5.8. (1) The parity-check matrix of the $(7,4,3)_2$ Hamming code of order $r = 3$ in Example 4.3(1) was constructed according to Theorem 4.10. With a little calculation (4.4.2) implies the well-known result of

$$A(Z) = \frac{1}{8}\left[(1+Z)^7 + 7(1+Z)^3(1-Z)^4\right] = 1 + 7Z^3 + 7Z^4 + Z^7$$

for the weight enumerator.

(2) For the $(15,11,3)_2$ Hamming code of order $r = 4$, $\boldsymbol{H}$ can be chosen as

$$\boldsymbol{H} = \begin{pmatrix} 0\,0\,0\,0\,1\,1\,1\,1\,1\,1\,1 & 1\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0\,0\,1\,1\,1\,1 & 0\,1\,0\,0 \\ 1\,0\,1\,1\,0\,1\,1\,0\,0\,1\,1 & 0\,0\,1\,0 \\ 1\,1\,0\,1\,1\,0\,1\,0\,1\,0\,1 & 0\,0\,0\,1 \end{pmatrix}.$$

There are fifteen binary columns of length 4, the all-zero column not counted. The four unit vectors are placed together to form the identity matrix to the right. The other eleven columns are sorted as binary numbers. ■

5.4.2 Simplex Codes

Definition 5.5. *The dual code* $\mathcal{C}^\perp$ *of the form* $(2^r-1, r, 2^{r-1})_2$ *of the binary* $(2^r-1, 2^r-r-1, 3)_2$ *Hamming code* $\mathcal{C}$ *is called a* simplex code.

The generator matrix of the simplex code is the parity-check matrix of the Hamming code, so according to Theorem 4.10 it consists of the binary numbers as columns. Thus every row and every codeword (apart from the all-zero word) are of weight 2^{r-1} which can be easily proven by mathematical induction. Therefore $d_{\min} = 2^{r-1}$ and

$$A^\perp(Z) = 1 + (2^r-1)Z^{2^{r-1}} = 1 + nZ^{(n+1)/2}. \tag{5.4.5}$$

Since the difference of two codewords is again a codeword all pairs of codewords have a constant Hamming distance of 2^{r-1}. In geometry such a construction is called a *simplex*.

The weight distribution (4.4.3) of the $(n, n-k)$ simplex code implies the weight distribution (4.4.2) of the dual (n, k) Hamming code with $n-k=r$ and $2^r = n+1$ as follows:

$$\begin{aligned} A(Z) &= 2^{-(n-k)}(1+Z)^n A^\perp\left(\frac{1-Z}{1+Z}\right) \\ &= 2^{-r}(1+Z)^n\left[1 + n\left(\frac{1-Z}{1+Z}\right)^{(n+1)/2}\right] \\ &= \frac{1}{n+1}\Big[(1+Z)^n + n(1+Z)^{(n-1)/2}(1-Z)^{(n+1)/2}\Big]. \end{aligned}$$

Obviously the low-rate simplex code satisfies the Plotkin bound with equality whereas the high-rate Hamming code satisfies the Hamming bound with equality. Asymptotically for the simplex code $k/n \to 0$ and $d_{\min}/n \to 1/2$ as $r \to \infty$.

5.4.3 The Colored Hats Puzzle

The colored hats puzzle is a nice example for the successful application of Hamming codes to problems which do not seem to be connected to communications or coding. This puzzle was first published as an article which appeared in the Science Times section of the New York Times of April 10th, 2001.

The puzzle is stated as follows: Each player of a team is randomly and independently assigned to wear a colored hat (either red=0 or blue=1). Each player views the colors of his other teammates (but can not see his own color), and then tries to guess the color of his own hat. No communication is allowed between the players except for a strategy session before the game begins. It is allowed that some players do not guess and remain neutral. The team wins a prize if at least one player guesses his own color correctly and no player guesses

incorrectly. Vice versa, the team loses if there are no guesses or at least one player guesses incorrectly.

On the first view, the team seems to have a chance of winning of only 50%. However, with a smart strategy, the chance of winning is almost 100%. More precisely, if the number of players has the form $n = 2^r - 1$, then the chance of winning is $n/(n+1)$.

The smart strategy is defined as follows. Let $\boldsymbol{y} = (y_0, \ldots, y_{n-1})$ be a vector representing the colors of the n hats. Let $\mathcal{C}$ be the $(2^r-1, 2^r-r-1, 3)_2$ Hamming code. By viewing his teammates, the i-th player knows the two vectors

$$\begin{aligned}\boldsymbol{a}_i &= (y_0, \ldots, y_{i-1}, 0, y_{i+1}, \ldots, y_{n-1}),\\ \boldsymbol{b}_i &= (y_0, \ldots, y_{i-1}, 1, y_{i+1}, \ldots, y_{n-1})\end{aligned}$$

and guesses the color of his own hat as follows:

$$\begin{aligned}\boldsymbol{a}_i \notin \mathcal{C} \text{ and } \boldsymbol{b}_i \notin \mathcal{C} \quad &\Rightarrow \quad \text{neutral}\\ \boldsymbol{a}_i \in \mathcal{C} \text{ and } \boldsymbol{b}_i \notin \mathcal{C} \quad &\Rightarrow \quad \text{guess 1}\\ \boldsymbol{a}_i \notin \mathcal{C} \text{ and } \boldsymbol{b}_i \in \mathcal{C} \quad &\Rightarrow \quad \text{guess 0}\end{aligned}$$

The fourth case $\boldsymbol{a}_i \in \mathcal{C}$ and $\boldsymbol{b}_i \in \mathcal{C}$ is not possible since $d_H(\boldsymbol{a}_i, \boldsymbol{b}_i) = 1$, however, $\mathcal{C}$ has a minimum Hamming distance of 3.

Now we compute the chance of winning. Two cases have to be distinguished. Firstly, if $\boldsymbol{y} \in \mathcal{C}$, then all players guess incorrectly and the team loses. Secondly, we consider the case of $\boldsymbol{y} \notin \mathcal{C}$. Since $\mathcal{C}$ is perfect, there exists exactly one $\boldsymbol{c} \in \mathcal{C}$ with $d_H(\boldsymbol{y}, \boldsymbol{c}) = 1$. Let l be the position where the two vectors differ.

The l-th player guesses as follows: if $\boldsymbol{a}_l \in \mathcal{C}$, then he guesses 1 and $\boldsymbol{a}_l \neq \boldsymbol{y}$ since $\boldsymbol{y} \notin \mathcal{C}$. Hence, $\boldsymbol{a}_l = \boldsymbol{c} \in \mathcal{C}$ and $\boldsymbol{b}_l = \boldsymbol{y}$ and so his guess is correct. The same arguments also show a correct guess in case of $\boldsymbol{b}_l \in \mathcal{C}$. All other s-th players with $s \neq l$ remain neutral since $\boldsymbol{a}_s \notin \mathcal{C}$ and $\boldsymbol{b}_s \notin \mathcal{C}$.

In summary, in case of $\boldsymbol{y} \notin \mathcal{C}$ one player guesses correctly and all other players remain neutral. Hence

$$P(\boldsymbol{y} \notin \mathcal{C}) = \frac{2^n - 2^k}{2^n} = 1 - 2^{-(n-k)} = 1 - 2^{-r} = 1 - (n+1)^{-1} = \frac{n}{n+1}$$

is the teams's winning chance.

5.5 Simple Modifications to a Linear Code

The following modifications generating codes with modified parameters are of less theoretical but of great practical importance.

Definition 5.6. *An $(n, k, d_{\min})_q$ code can be modified to an $(n', k', d'_{\min})_q$ code in many different ways:*

(1) By expanding *a code, additional parity-check bits are attached:*

$$n' > n, \quad k' = k, \quad R' < R, \quad d'_{\min} \geq d_{\min}. \tag{5.5.1}$$

(2) By puncturing *a code, parity-check bits are deleted:*

$$n' < n, \quad k' = k, \quad R' > R, \quad d'_{\min} \leq d_{\min}. \tag{5.5.2}$$

(3) By lengthening *a code, additional information bits are attached:*

$$n' > n, \quad k' > k, \quad n' - k' = n - k, \quad R' > R, \quad d'_{\min} \leq d_{\min}. \tag{5.5.3}$$

(4) By shortening *a code, information bits are deleted:*

$$n' < n, \quad k' < k, \quad n' - k' = n - k, \quad R' < R, \quad d'_{\min} \geq d_{\min}. \tag{5.5.4}$$

The modifications 1 and 2 as well as 3 and 4 are inverse to each other. The following theorem describes a method for expansion, which can only be applied once, but not repeatedly.

Theorem 5.11. *Every binary* $(n, k, d_{\min})_2$ *code with an odd minimum distance* $d_{\min}$ *can be expanded to an* $(n+1, k, d_{\min}+1)_2$ *code.*

Proof. Let there be an arbitrary codeword. If the Hamming weight is even, a 0 otherwise a 1 is attached as a parity-check bit. The extended code only contains codewords of even weight, therefore $d'_{\min} = d_{\min} + 1$. Linearity is preserved. ■

Every $(2^r - 1, 2^r - r - 1, 3)_2$ Hamming code can be extended to a $(2^r, 2^r - r - 1, 4)_2$ code, therefore in addition to correcting one error a further error can be detected. The weight enumerator with $n = 2^r$ (see Problem 4.11) is

$$A(Z) = \frac{1}{2n}\Big[(1+Z)^n + (1-Z)^n + 2(n-1)(1-Z^2)^{n/2}\Big]. \tag{5.5.5}$$

For describing the expansion by matrices, let $\boldsymbol{G} \in \mathbb{F}_2^{k,n}$ be the generator matrix and $\boldsymbol{H} \in \mathbb{F}_2^{n-k,n}$ the parity-check matrix of the $(n, k)_2$ code. The generator matrix of the extended $(n+1, k)_2$ code is $\boldsymbol{G'} \in \mathbb{F}_2^{k,n+1}$ and its parity-check matrix is $\boldsymbol{H'} \in \mathbb{F}_2^{n+1-k,n+1}$. Let $s_i = g_{i,0} + \cdots + g_{i,n-1} \in \mathbb{F}_2$ $(0 \leq i \leq k-1)$ be the row sums of $\boldsymbol{G}$. The extended code is then generated by

$$\boldsymbol{G'} = \left(\begin{array}{c|c} & s_0 \\ \boldsymbol{G} & \vdots \\ & s_{k-1} \end{array}\right). \tag{5.5.6}$$

For the unit vectors as information words the construction method of the proof of Theorem 4.11 is used. Due to linearity, this construction is also valid for all the other codewords.

The parity-check matrix can be derived as follows. The first $n-k$ check conditions are unchanged and the sum over all code bits is zero, therefore

$$\boldsymbol{H}' = \left(\begin{array}{ccc|c} & & & 0 \\ & \boldsymbol{H} & & \vdots \\ & & & 0 \\ \hline 1 & 1 & \ldots\ldots\ldots\ldots & 1 \quad 1 \end{array}\right). \tag{5.5.7}$$

Example 5.9. The expansion of the $(7,4,3)_2$ Hamming code leads to a $(8,4,4)_2$ code with

$$\boldsymbol{G}' = \left(\begin{array}{ccccccc|c} 1&0&0&0&0&1&1&1 \\ 0&1&0&0&1&0&1&1 \\ 0&0&1&0&1&1&0&1 \\ 0&0&0&1&1&1&1&0 \end{array}\right) \quad , \quad \boldsymbol{H}' = \left(\begin{array}{ccccccc|c} 0&1&1&1&1&0&0&0 \\ 1&0&1&1&0&1&0&0 \\ 1&1&0&1&0&0&1&0 \\ \hline 1&1&1&1&1&1&1&1 \end{array}\right).$$

Obviously, $\boldsymbol{G}'\boldsymbol{H}'^T = \boldsymbol{0}$. The minimum row weight in $\boldsymbol{G}'$ is 4. Every selection of three columns in $\boldsymbol{H}'$ is linearly independent and the last column is the sum of the first three columns in $\boldsymbol{H}'$, therefore $d'_{\min} = 4$. The extended Hamming code is self-dual only for $r = 3$. ■

5.6 Simple Decoding Techniques

In this section we will restrict our considerations to hard-decision decoding only, i.e., for the received word we will try to find the codeword with minimum Hamming distance. Just trying out all the codewords might work in theory but not in practice since this is usually too time-consuming. The decoding techniques discussed in this section will simplify the matter a great deal and can be easily derived with few algebraic considerations.

5.6.1 Standard Array

The standard array in the space $\mathbb{F}_q^n$ of all possible received words is the basis for the decoding methods discussed in the next subsection.

Definition 5.7. *Given an $(n,k)_q$ code $\mathcal{C}$ with the parity-check matrix $\boldsymbol{H} \in \mathbb{F}_q^{n-k,n}$, the* syndrome *of the received word $\boldsymbol{y}$ is defined as*

$$\boldsymbol{s} = \boldsymbol{y}\boldsymbol{H}^T. \tag{5.6.1}$$

Therefore the syndrome is of length $n-k$. For the representation $\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{e}$ where $\boldsymbol{a}$ is a codeword and $\boldsymbol{e}$ is an error word,

$$\boldsymbol{s} = (\boldsymbol{a} + \boldsymbol{e})\boldsymbol{H}^T = \boldsymbol{a}\boldsymbol{H}^T + \boldsymbol{e}\boldsymbol{H}^T = \boldsymbol{e}\boldsymbol{H}^T. \tag{5.6.2}$$

The syndrome of a word is the all-zero word if and only if the word is a codeword. Therefore the syndrome of the received word is independent of the transmitted word and only dependent on the overlayed error pattern. There are $q^n - q^k$ different error patterns which are not codewords and there are q^{n-k} different syndromes, thus an error pattern is not uniquely determined by its syndrome.

The syndromes are enumerated as $\boldsymbol{s}_\mu$ with $0 \leq \mu \leq q^{n-k}-1$ and $\boldsymbol{s}_0 = \boldsymbol{0}$. For each syndrome the set of error patterns is defined leading only to this syndrome:

$$\mathcal{M}_\mu = \{\boldsymbol{e} \in \mathbb{F}_q^n \mid \boldsymbol{e}\boldsymbol{H}^T = \boldsymbol{s}_\mu\}. \tag{5.6.3}$$

Since all codewords have the syndrome zero, $\mathcal{M}_0 = \mathcal{C}$ and all other $\mathcal{M}_\mu$ contain no codewords. Furthermore, the sets are all disjoint since an error pattern can not have various syndromes. Let there be $\boldsymbol{e}, \boldsymbol{e}' \in \mathcal{M}_\mu$. Since $\boldsymbol{e}\boldsymbol{H}^T = \boldsymbol{e}'\boldsymbol{H}^T$ it follows that $(\boldsymbol{e}' - \boldsymbol{e})\boldsymbol{H}^T = \boldsymbol{e}'\boldsymbol{H}^T - \boldsymbol{e}\boldsymbol{H}^T = \boldsymbol{0}$. Thus the difference between two words in $\mathcal{M}_\mu$ is always a codeword. Hence, for an arbitrary $\boldsymbol{e} \in \mathcal{M}_\mu$,

$$\boldsymbol{e} + \mathcal{C} = \{\boldsymbol{e} + \boldsymbol{a} \mid \boldsymbol{a} \in \mathcal{C}\} = \mathcal{M}_\mu. \tag{5.6.4}$$

So the set $\mathcal{M}_\mu$ can be described by the sum of an arbitrary element in $\mathcal{M}_\mu$ and the code. Therefore each set $\mathcal{M}_\mu$ has the same cardinality

$$\left|\mathcal{M}_\mu\right| = \left|\mathcal{C}\right| = q^k = \frac{q^n}{q^{n-k}} = \frac{\text{number of all words}}{\text{number of syndromes}}. \tag{5.6.5}$$

The q^{n-k} sets $\mathcal{M}_\mu$ form a unique disjoint decomposition of $\mathbb{F}_q^n$,

$$\mathbb{F}_q^n = \biguplus_{\mu=0}^{q^{n-k}-1} \mathcal{M}_\mu. \tag{5.6.6}$$

Definition 5.8. *The sets $\mathcal{M}_\mu$ are called* cosets *and the decomposition (4.6.6) is called a* standard array *(or coset decomposition). Every $\boldsymbol{e} \in \mathcal{M}_\mu$ can be the* coset leader *in the representation $\mathcal{M}_\mu = \boldsymbol{e} + \mathcal{C}$.*

In Section A.4 the principles of standard arrays are abstractly explained without referring to the syndromes. The group $\mathcal{G}$ now corresponds to $\mathbb{F}_q^n$ and the subgroup $\mathcal{U}$ corresponds to $\mathcal{C}$. There is an equivalence relation between the two words $\boldsymbol{y}$ and $\boldsymbol{y}'$, if their syndromes are equal or, equivalently, if their difference $\boldsymbol{y} - \boldsymbol{y}'$ is a codeword. For all $\boldsymbol{y} \in \mathcal{M}_\mu$, $[\boldsymbol{y}] = \mathcal{M}_\mu = \boldsymbol{y} + \mathcal{C}$ are the corresponding equivalence classes or cosets.

Now, in each set $\mathcal{M}_\mu$ a leader $\boldsymbol{e}_\mu$ of minimum Hamming weight is selected.

$$\mathcal{M}_\mu = \boldsymbol{e}_\mu + \mathcal{C} \qquad \text{with} \qquad w_H(\boldsymbol{e}_\mu) \leq w_H(\boldsymbol{e}) \text{ for all } \boldsymbol{e} \in \mathcal{M}_\mu. \tag{5.6.7}$$

This leader of minimum weight is not necessarily unique. However, in $\mathcal{M}_0 = \mathcal{C}$, $\boldsymbol{e}_0 = \boldsymbol{0}$ is, of course, unique.

Example 5.10. Consider the $(5,2)_2$ code $\mathcal{C} = \{00000, 10110, 01011, 11101\}$ with

$$\boldsymbol{G} = \begin{pmatrix} 1\,0 & 1\,1\,0 \\ 0\,1 & 0\,1\,1 \end{pmatrix} \quad , \quad \boldsymbol{H} = \begin{pmatrix} 1\,0 & 1\,0\,0 \\ 1\,1 & 0\,1\,0 \\ 0\,1 & 0\,0\,1 \end{pmatrix}.$$

This code implies the standard array in a unique way:

μ	$\boldsymbol{e}_\mu$	$\mathcal{M}_\mu$				$\boldsymbol{s}_\mu$
0	00000	00000	10110	01011	11101	000
1	00001	00001	10111	01010	11100	001
2	00010	00010	10100	01001	11111	010
3	00100	00100	10010	01111	11001	100
4	01000	01000	11110	00011	10101	011
5	10000	10000	00110	11011	01101	110
6	11000	11000	01110	10011	00101	101
7	01100	01100	11010	00111	10001	111

$\mathcal{M}_0$ is the code itself. In $\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \mathcal{M}_4, \mathcal{M}_5$ the leaders $\boldsymbol{e}_\mu$ of minimum weight are unique. In $\mathcal{M}_6$ as well as $\mathcal{M}_7$ there are two words of minimum weight, therefore the leaders can be chosen arbitrarily. For example

$$\begin{aligned} 01100 + \mathcal{C} &= \{01100, 11010, 00111, 10001\} \\ &= \{10001, 00111, 11010, 01100\} = 10001 + \mathcal{C}. \end{aligned}$$

It is easy to see that for all error patterns in $\mathcal{M}_\mu$ the syndrome is $\boldsymbol{s}_\mu$ and that $\mathcal{M}_\mu = \boldsymbol{e}_\mu + \mathcal{C}$. For this example $d_{\min} = 3$, $t = 1$ and $L_t = 1 + n = 6$ (for an explanation see Theorem 4.13). ■

5.6.2 Syndrome Decoding

In this section the syndrome decoding will be discussed as a simple decoding technique for hard decisions. There are further methods as for example the majority and the threshold decoding schemes, however, these will not be discussed here. These methods might deliver a more practical decoding scheme than the primitive trial-and-error method of whether or not the codewords are within the minimum Hamming distance from the received word. But for high-performance codes (i.e., huge cardinality and large coding gain) the simple decoding methods described below are still too time-consuming and too slow in decoding speed. Really powerful decoding methods are not developed until cyclic codes are introduced.

Theorem 5.12. *Let there be an $(n,k)_q$ code with the coset leaders $\boldsymbol{e}_\mu$ of minimum weight in the standard array. The maximum-likelihood decoder is realized by the following method:*

If the received word $\boldsymbol{y}$ is in the same coset $\mathcal{M}_\mu$ as the leader $\boldsymbol{e}_\mu$, then $\hat{\boldsymbol{a}} = \boldsymbol{y} - \boldsymbol{e}_\mu$ is chosen as an estimation for the transmitted codeword.

Proof. If $\boldsymbol{y}$ is in $\mathcal{M}_\mu$, then there exists a representation $\boldsymbol{y} = \boldsymbol{e}_\mu + \boldsymbol{a}'$ with $\boldsymbol{a}' \in \mathcal{C}$. Thus $\hat{\boldsymbol{a}} = \boldsymbol{y} - \boldsymbol{e}_\mu = \boldsymbol{a}'$, therefore $\hat{\boldsymbol{a}} = \boldsymbol{a}'$ is proven to be a codeword.

Let $\boldsymbol{b}$ be an arbitrary codeword. We are to show that $d_H(\boldsymbol{y}, \hat{\boldsymbol{a}}) \leq d_H(\boldsymbol{y}, \boldsymbol{b})$. Since $\hat{\boldsymbol{a}}, \boldsymbol{b} \in \mathcal{C}$, $\hat{\boldsymbol{a}} - \boldsymbol{b} \in \mathcal{C}$, thus $\boldsymbol{e}_\mu + (\hat{\boldsymbol{a}} - \boldsymbol{b}) \in \mathcal{M}_\mu$. Since $\boldsymbol{e}_\mu$ is of minimum weight in $\mathcal{M}_\mu$, $w_H(\boldsymbol{e}_\mu) \leq w_H(\boldsymbol{e}_\mu + \hat{\boldsymbol{a}} - \boldsymbol{b})$. Since $\boldsymbol{e}_\mu = \boldsymbol{y} - \hat{\boldsymbol{a}}$,

$$d_H(\boldsymbol{y}, \hat{\boldsymbol{a}}) = w_H(\boldsymbol{y} - \hat{\boldsymbol{a}}) \leq w_H(\boldsymbol{y} - \boldsymbol{b}) = d_H(\boldsymbol{y}, \boldsymbol{b}).$$

Therefore $\hat{\boldsymbol{a}}$ has a distance smaller than or equal to the distance of every other codeword from $\boldsymbol{y}$, thus the ML rule is realized. ■

The so-called *coset decoding* works as follows. In the standard array $\boldsymbol{y}$ is sought, so μ and $\boldsymbol{e}_\mu$ are known. Thus $\hat{\boldsymbol{a}} = \boldsymbol{y} - \boldsymbol{e}_\mu$ is the ML estimation.

This method can be further simplified, which is then called the *syndrome decoding*. A table with the q^{n-k} rows $(\boldsymbol{s}_\mu, \boldsymbol{e}_\mu)$ is stored. The syndrome $\boldsymbol{s} = \boldsymbol{y}\boldsymbol{H}^T$ of the received word $\boldsymbol{y}$ is computed and $\boldsymbol{s}_\mu$ with $\boldsymbol{s} = \boldsymbol{s}_\mu$ is sought in the table. An additional simplification is explained by the following example.

Example 5.11. (Continuation of Example 4.10) Two tables are given from which $\boldsymbol{e}_\mu$ for $\boldsymbol{s}_\mu$ can be read. The table on the left is only a part of the table in Example 4.10 whereas the table on the right is a rearrangement of the table on the left.

μ	$\boldsymbol{s}_\mu$	$\boldsymbol{e}_\mu$
0	000	00000
1	001	00001
2	010	00010
3	100	00100
4	011	01000
5	110	10000
6	101	11000
7	111	01100

ν	$\boldsymbol{s}_\nu$	$\boldsymbol{e}_\nu$
0	000	00000
1	001	00001
2	010	00010
3	011	01000
4	100	00100
5	101	11000
6	110	10000
7	111	01100

In the table on the right the syndromes are listed as binary numbers, therefore the syndromes can be used as a direct address for a memory which only contains the leaders $\boldsymbol{e}_\nu$. Thus we need not search for $\boldsymbol{s} = \boldsymbol{s}_\mu$, but only compute $\boldsymbol{s}_\nu$, determine ν, consult the table and decide on $\hat{\boldsymbol{a}} = \boldsymbol{y} - \boldsymbol{e}_\nu$. ■

In practice this method is still infeasible for large block lengths, since for the relatively simple $(511, 259, 61)_2$ BCH code, $2^{511-259} \approx 10^{76}$ error patterns of length 511 are to be stored in the table.

The possible ambiguities for the choice of the leader in the cosets correspond to the possible ambiguities for the ML decoding. For the BMD decoding these ambiguities do not exist since the decoder only needs to work properly for a maximum of t errors.

Theorem 5.13. *The BMD decoding for an $(n, k, d_{\min})_q$ code with $2t+1 \leq d_{\min}$, as already discussed, corrects all error patterns up to the weight of t. The number of these error patterns is*

$$L_t = \left|K_t(\boldsymbol{0})\right| = \sum_{r=0}^{t} \binom{n}{r} (q-1)^r. \tag{5.6.8}$$

Among the q^{n-k} cosets there are at least L_t cosets in which the leader of minimum Hamming weight is unique. These L_t leaders are identical to the words of weight $\leq t$.

Proof. According to the Hamming bound, $L_t \leq q^{n-k}$. We are to show that the words of weight $\leq t$ are in different cosets.

For this let $\boldsymbol{e} \neq \boldsymbol{e}'$ be arbitrary with $w_H(\boldsymbol{e}) \leq t$ and $w_H(\boldsymbol{e}') \leq t$. If both the corresponding cosets were not disjoint, there would exist a $\boldsymbol{y}$ with $\boldsymbol{y} = \boldsymbol{e} + \boldsymbol{a} = \boldsymbol{e}' + \boldsymbol{a}'$ and $\boldsymbol{a}, \boldsymbol{a}' \in \mathcal{C}$. However, since the difference $\boldsymbol{a} - \boldsymbol{a}' = \boldsymbol{e}' - \boldsymbol{e} \neq \boldsymbol{0}$ is a codeword,

$$d_{\min} \leq w_H(\boldsymbol{a} - \boldsymbol{a}') = w_H(\boldsymbol{e}' - \boldsymbol{e}) \leq w_H(\boldsymbol{e}') + w_H(\boldsymbol{e}) \leq t + t < d_{\min}.$$

This is a contradiction, therefore both cosets are disjoint. ■

5.7 Problems

5.1. Determine the row-echelon form and the rank of

$$\boldsymbol{G} = \begin{pmatrix} 1\,0\,1\,0\,1\,0\,1 \\ 1\,1\,0\,0\,1\,1\,0 \\ 0\,1\,1\,0\,0\,1\,1 \end{pmatrix}$$

over $\mathbb{F}_2$ and $\mathbb{F}_3$. Comment on the results.

5.2. Is $\boldsymbol{a} = 1011010$ a codeword for

$$\boldsymbol{G} = \begin{pmatrix} 0\,0\,0\,1\,1\,1\,1 \\ 0\,1\,1\,0\,0\,1\,1 \\ 1\,0\,1\,0\,1\,0\,1 \end{pmatrix}?$$

What is the corresponding information word?

5.3. Transform the generator matrices of the parity-check code with parity-check bits in front or behind into each other by using elementary row operations for $k = 3$. Interpret the result.

5.4 Prove for an $(n, k, d_{\min})_q$ code $\mathcal{C}$ the equivalence of the following statements.

(1) $\mathcal{C}$ is an MDS code, i.e., $d_{\min} = n - k + 1$.

(2) Every selection of $n-k$ columns in the parity-check matrix is linearly independent.

(3) Every selection of k columns in the generator matrix is linearly independent.

5.5. Determine the asymptotic coding gain of the binary Hamming codes and the simplex codes as $r \to \infty$. Where are these two code families located in the representation of asymptotic bounds in Figure 3.4?

5.6. Prove that the binary Hamming codes have a symmetrical weight distribution according to (3.5.7).

5.7. Prove for the $(n, k)_2$ code $\mathcal{C}$ that the probability of undetected errors can be calculated from the weight distribution of the dual code according to

$$P_{ue} = 2^{-(n-k)} A^{\perp}(1 - 2p_e) - (1 - p_e)^n. \tag{5.7.1}$$

5.8. Make the inductive step for the derivation of (4.4.3).

5.9. Let a $(5, 2)_2$ block code contain the codewords 01111 and 11100. Determine a systematic generator matrix (parity-check bits at the end) and a parity-check matrix. Determine the matrix description of the extended code (two alternatives for $\boldsymbol{H}$, are they equivalent?) as well as the code (set of codewords) and the minimum distance.

5.10. Prove that the generator matrix and the parity-check matrix of the extended self-dual $(8, 4)_2$ Hamming code can attain an identical form (see Example 4.9). Verify that $A(Z) = A^{\perp}(Z)$ for the weight distributions.

5.11. Let $A(Z)$ be the weight distribution of an $(n, k)_2$ code $\mathcal{C}$ and let $A'(Z)$ be the weight distribution of the extended $(n + 1, k)$ code $\mathcal{C}'$ according to Theorem 4.11. Prove that

$$A'(Z) = \frac{1}{2}\Big[(1 + Z)A(Z) + (1 - Z)A(-Z)\Big]. \tag{5.7.2}$$

From this derive the weight distribution (4.5.1) of the extended Hamming code as well as the weight distribution (3.5.8) of the parity-check code.

5.12. For the weight enumerator of the $(n, n-1)_2$ parity-check code prove that

$$A(Z) = \frac{1}{2} \cdot \Big[(1 + Z)^n + (1 - Z)^n\Big]. \tag{5.7.3}$$

Derive this result in three different ways: directly from (4.5.8), from Theorem 5.8 and from (5.7.2).

5.13. Using the binary encoder

$$(u_0, u_1, u_2) \mapsto (u_0, u_1, u_2, u_0 + u_2, u_0 + u_1, u_1 + u_2)$$

determine the parameters $(n, k, d_{\min})$, the matrices $\boldsymbol{G}$, $\boldsymbol{H}$ and the standard array (coset decomposition). Decode 011011 and 000111.

5.14. Can the $(7, 4)_2$ Hamming code be changed such that the decoding can do without tables? Transform the result to the general Hamming code. Decode 0011100.

5.15. As in Problem 3.26, let $\mathcal{E} \subseteq \mathbb{F}_q^n$ be an arbitrary set of error patterns with $\boldsymbol{0} \in \mathcal{E}$. Prove the equivalence of (1) and (2):

(1) $\mathcal{C}$ corrects all error patterns in $\mathcal{E}$, i.e., $\boldsymbol{a}+\boldsymbol{e} \neq \boldsymbol{a}'+\boldsymbol{e}'$ for all $\boldsymbol{a}, \boldsymbol{a}' \in \mathcal{C}$ with $\boldsymbol{a} \neq \boldsymbol{a}'$ and all $\boldsymbol{e}, \boldsymbol{e}' \in \mathcal{E}$.

(2) The syndromes of all error patterns in $\mathcal{E}$ are different.