Coding theory lectures

By Dr. Mohammed M. AL-Ashker
Associated professor
Mathematics Department
E.mail:mashker @ mail.iugaza.edu
Islamic University of Gaza P.O.Box 108, Gaza, Palestine
1 Basic concepts of linear codes

1.1 Three fields

1.2 Linear codes, generators and parity check matrices

1.2.1 Subcode:

1.3 Dual codes

1.4 Weights and distances:

1.5 New codes from old

1.5.1 Puncturing codes

1.5.2 Extending codes

1.5.3 Shortening codes

1.5.4 Direct sums

1.5.5 The (u|u+v) construction

1.6 Permutation equivalent codes:

1.7 Hamming codes

1.8 The Golay codes

1.8.1 The Golay code \( G_{24} \)

1.8.2 The ternary Golay codes

1.9 Reed- Muller codes

1.10 Encoding, decoding, and Shannon’s theorem

1.10.1 Decoding and shannon’s theorem

1.11 Sphere packing bound, covering radius and perfect codes
# 2 Bounds on the size of codes

2.1 $A_q(n, d)$ and $B_q(n, d)$ ............................................................... 51

2.2 Singleton upper bound and MDS codes. ....................................... 53

2.3 Lexicodes ......................................................... 55

# 3 Finite fields

3.1 Introduction to finite fields ......................................................... 58

3.2 Polynomials and the Euclidean algorithm .................................... 59

3.3 Primitive elements ............................................................. 62

3.4 Constructing finite fields ........................................................ 63

3.5 Subfields ........................................................................... 66

3.6 Field automorphisms ............................................................... 68

3.7 Cyclotomic cosets and minimal polynomials .................................. 69

# 4 Cyclic codes

4.1 Factoring $x^n - 1$ ................................................................. 76

4.2 Basic theory of cyclic codes ..................................................... 78

4.3 Idempotent and multipliers ........................................................ 87

4.4 Zeros of a cyclic code ............................................................... 93

4.5 Meggitt decoding of cyclic codes .............................................. 96
Chapter 1

Basic concepts of linear codes

Introduction:

Coding theory began in the late 1948’s with work of C. Shannon, hamming, Golay and others. Historically coding theory originated as the mathematical foundation for transmission of messages over noisy channels. In fact a multitude of diverse applications have been discovered such as the minimization of noise from compact disc recording, the transmission of financial information across telephone lines, data transfer from one computer to another and so on. Coding theory deals with the problem of detecting and correcting transmission errors caused by noise on the channel. The following diagram provides a rough idea of a general information transmission system.

The most important part of the diagram, as far as we are concerned is the noise, for
without it there would be no need for the theory.

1.1 Three fields

A field \( \mathcal{F} \) is an algebraic structure consisting of a set together with two operations, usually called addition (+) and multiplication (\( \cdot \)) which satisfy certain axioms

1) \((\mathbb{F}, +)\) is an abelian group
2) \( \cdot \) is associative \( a \cdot (b \cdot c) = (a \cdot b) \cdot c \) for all \( a, b, c \) in \( \mathbb{F} \)
3) \( \cdot \) is left and right distributive over +, \( a \cdot (b + c) = a \cdot b + a \cdot c \) and \( (b + c) \cdot a = b \cdot c + c \cdot a \).

If \( \mathbb{F} \) is commutative over \( \cdot \) and \( \mathbb{F}^* = \mathbb{F}/0 \) is a group over, then \( \mathbb{F} \) is a field. The three fields that are very common in the study of linear codes are the binary field \( \mathbb{Z}_2 = \mathbb{F}_2 = \{0, 1\} \) with two elements, the ternary field with three elements \( \mathbb{Z}_3 = \mathbb{F}_3 = \{0, 1, 2\} \) and the quaternary field with four elements is \( \mathbb{F}_4 = \{0, 1, \omega, \bar{\omega} = 1 + \omega = \omega^2\} \) the addition and multiplication tables of the above fields are:

The field \( \mathbb{F}_2 \)

<table>
<thead>
<tr>
<th>+</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \cdot )</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The field \( \mathbb{F}_3 \)

<table>
<thead>
<tr>
<th>+</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \cdot )</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

And addition and multiplication table for \( \mathbb{F}_4 \) are
1.2 Linear codes, generators and parity check matrices

Let $\mathbb{F}_q^n$ be the vector space of all $n$-tuples over the finite field $\mathbb{F}_q$. $n$ is the length of the vectors in $\mathbb{F}_q^n$, which is the dimension of $\mathbb{F}_q^n$. An $(n, M)$ code $C$ over $\mathbb{F}_q$ is a subset of $\mathbb{F}_q^n$ of size $M$, that is $|C| = M$ (the number of all codewords in $C$).

Remark 1.2.1. We write the vectors $(a_1, a_2, \ldots, a_n)$ in $\mathbb{F}_q^n$ in the form $a_1a_2\cdots a_n$ and call the vectors in $C$ codewords.

The codes over $\mathbb{Z}_2 = \mathbb{F}_2$ are called binary codes and the codes over $\mathbb{Z}_3 = \mathbb{F}_3$ are called ternary codes and over $\mathbb{F}_4$ are called quaternary codes. Also the term quaternary has also been used to refer to codes over $\mathbb{Z}_4$ the integers modulo 4.

Definition 1.2.1. Linear codes: If $C$ is a $k$-dimensional subspace of $\mathbb{F}_q^n$, then $C$ will be called an $[n, k]$ linear code over $\mathbb{F}_q$, and the linear $[n, k]$ code $C$ has $q^k$ codewords.

Definition 1.2.2. Generator matrices and parity check matrices: A generator matrix for an $[n, m]$ code $C$ is any $k \times n$ matrix $G$ whose rows form a basis for $C$.

Remark 1.2.2. In general for any code $C$ there are many generator matrices of size $k \times n$.

If $C$ is any $[n, k]$-code, with generator matrix $G$, then the codewords in $C$ are the linear combination of the rows of $G$.
Example 1.2.1. : Consider the binary code with generator matrix

\[ G = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}, \]

this matrix has three rows, then the dimension of the code is 3 and has \(2^3\) codewords.

Definition 1.2.3. For any set of \(k\) independent columns of a generator matrix \(G\), the corresponding set of coordinates form an information set of \(C\). and the remaining \(r = n-k\) coordinates are called the redundancy set and \(r\) is called the redundancy of \(C\).

Definition 1.2.4. a generator matrix of the form

\[ G = \begin{bmatrix} I_k \\ A \end{bmatrix}, \]

where \(I_k\) is the \(k \times k\) identity matrix of size \(k\)

is said to be in the standard form.

Example 1.2.2. : The binary \([7,4]\)-code with generator matrix

\[ G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}, \]

in standard form the first 4 coordinates form information set.

Definition 1.2.5. Let \(C\) be a linear \([n,k]\) code. The set \(C^\perp = \{ \mathbf{x} \in \mathbb{F}_q^n : \mathbf{x} \cdot \mathbf{c} = 0 \text{ for all } \mathbf{c} \in C \}\) is called the dual code of \(C\). If \(C\) is a linear code with \(k \times n\) generator matrix

\[ G = \begin{bmatrix} I_k \\ A \end{bmatrix} \]

in standard form, and if \(H\) is the matrix

\[ H = \begin{bmatrix} -A^\top \\ I_{n-k} \end{bmatrix}, \]

where \(A^\top\) is the transpose of \(A\). Then \(HG^\top = -A^\top + A^\top = 0\). Hence the rows of \(H\) are orthogonal to the rows of \(G\) and since \(\text{rank } H = n - k = \dim (C^\perp)\), we deduce that \(H\) is a
generator matrix for the dual code \( C^\perp \). The matrix \( H \) is called the parity check matrix for the \([n, k]−\) code \( C \) of size \((n − k × n)\).

The matrix \( H \) is defined also by

\[
C = \{ x \in \mathbb{F}_q^n : Hx^\top = 0 \}.
\]

The code \( C \) is the kernel of the linear transformation \( L : \mathbf{x} \rightarrow H\mathbf{x}^\top \), \( \dim(\ker L) = \dim C = k \), because \( H \) has rank \( n − k \).

**Theorem 1.2.1.** If \( G = \begin{bmatrix} I_k & A \end{bmatrix} \) is a generator matrix for the \([n, k]\) code \( C \) in standard form then \( H = \begin{bmatrix} -A^\top & I_{n-k} \end{bmatrix} \), is a parity check matrix for \( C \).

**Example 1.2.3.** Binary repetition code:

The repetition code \( C \) is \([n, 1]\) binary linear code if it consisting only the two codewords \( \mathbf{0} = 00 \cdots 0 \) and \( \mathbf{1} = 11 \cdots 1 \). This code can correct up to \( e = \left\lfloor \frac{(n-1)}{2} \right\rfloor \) errors.

The generator matrix of the repetition code is

\[
G = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \end{bmatrix}
\]

in standard form.

The parity check matrix has the form

\[
H = \begin{bmatrix}
1 \\
1 \\
\vdots \\
1 \\
1 \\
\end{bmatrix}
\]

The first coordinate is an information set and the last \( n − 1 \) coordinates form a redundancy set.

**Example 1.2.4.** The matrix

\[
G = \begin{bmatrix} I_k & A \end{bmatrix}, \quad \text{where } G = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1
\end{bmatrix}
\]
is a generator matrix in standard form for a \([7, 4]\) binary code that we denote by \(\mathcal{H}_3\). The parity check matrix for \(\mathcal{H}_3\) is
\[
H = \begin{bmatrix}
A^\top & I_3
\end{bmatrix} = \begin{bmatrix}
0 & 1 & 1 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1
\end{bmatrix}.
\]
This code \(\mathcal{H}_3\) is called the \([7, 4]\) Hamming code.

1.2.1 Subcode:

If \(C\) is not linear code, a subcode of \(C\) is any subset of \(C\). If \(C\) is linear, a sub code will be a subset of \(C\) which must also be linear, in this case a subcode of \(C\) is a subspace of \(C\).

1.3 Dual codes

Inner products: Let \(x = x_1x_2\cdots x_n, ~ y = y_1y_2\cdots y_n \in F_q^n\) be two vectors, then the inner product is denoted by the formula \(x \cdot y = \sum_{i=1}^n x_iy_i = x_1y_1 + \cdots + x_ny_n\). If \(C\) is a code over \(F_q\) then
\[
C^\perp = \{ x \in F_q^n : x \cdot c = 0 \text{ for all } c \in C \},
\]
is the dual code of \(C\). If \(G\) is the generator matrix for \([n, k]\) code \(C\) then \(H\) is the generator matrix for \([n, n-k]\) code \(C^\perp\).

Example 1.3.1. The generator matrix for the repetition code \([n, 1]\) is
\[
G = \begin{bmatrix}
1 & 1 & 1 & \cdots & 1 & 1
\end{bmatrix}
\]
and the generator matrix for the dual code \([n, n-1]\) code \(C^\perp\) is
\[
H = \begin{bmatrix}
1 \\
1 \\
\vdots \\
1
\end{bmatrix}
\]

is
\[
H = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}
\]
$H$ consists of all binary $n$-tuples $a_1a_2\cdots a_{n-1}b$, where $b = a_1 + a_2 + \cdots + a_{n-1}$. (The $n$th coordinate $b$ is an over all parity check for the first $n-1$ coordinates chosen, therefore so that the sum of all the coordinates equal 0). Then $G$ is the parity check matrix for $C^\perp$.

Remark 1.3.1. The code $C^\perp$ has the property that a single transmission error can be detected (since the sum of the coordinates will be 0) but not corrected (since changing any one of the received coordinates will give a vector whose sum of coordinates will be 0.)

**Definition 1.3.1.** A code $C$ is self-orthogonal if $C \subseteq C^\perp$ and is self-dual if $C = C^\perp$.

The length $n$ of a self-dual code is even and the dimension is $\frac{n}{2}$ because $\text{dim}(C) + \text{dim}(C^\perp) = n$ and $C = C^\perp$ so $\text{dim}(C) = \frac{n}{2}$.

The [7, 4] Hamming code $H_3$ is presented by the generator matrix

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

length 7, dimension 4.

Let $\widehat{H}_3$ be the code of length 8 and dimension 4 obtained from $H_3$ by adding an over all parity check coordinate to each vector of $G$ and thus to each codeword $H_3$. Then

$$
\widehat{G} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0
\end{bmatrix}
$$

is a generator matrix for $\widehat{H}_3$, we can easy verify that $\widehat{H}_3$ is self-dual because $n = 8$, dimension of $\widehat{H}_3$ is equal $\frac{n}{2} = 4$ and all codewords are orthogonal in pairs and in there self.

**Example 1.3.2.** The ternary [4, 2] code $H_{3,2}$ is called tetracode has generator matrix in standard form given by

$$
G = \begin{bmatrix}
1 & 0 & 1 & 1 \\
0 & 1 & 1 & 1 \\
0 & 1 & 1 & -1
\end{bmatrix}
$$
This code is also self dual because \( n = 4 \), dimension of \( H_{3,2} \) is equal \( \frac{n}{2} = 2 \) and all codewords are orthogonal in pairs and in there self.

**Definition 1.3.2.** The Hermitian inner product over the quaternary field \( F_4 \) is given by

\[
\langle x, y \rangle = x \cdot \bar{y} = \sum_{i=1}^{n} x_i \bar{y}_i,
\]

where \( \bar{\cdot} \), called conjugation and is given by \( \bar{0} = 0, \bar{1} = 1, \bar{\omega} = \omega, \bar{\omega^2} = 1 + \omega \).

**Definition 1.3.3.** The Hermitian dual of a quaternary code \( C \) by \( C^\perp_H = \{ x \in F_4^n : \langle x, c \rangle = x \cdot \bar{c} = 0 \text{ for all } c \in C \} \).

**Definition 1.3.4.** We define the conjugate of \( C \) to be \( \bar{C} = \{ \bar{c} | c \in C \} \), where \( \bar{c} = c_1 \bar{c}_2 \cdots \bar{c}_n \) when \( c = c_1 c_2 \cdots c_n \).

\( C^\perp_H = \bar{C}^\perp \).

**Remark 1.3.2.**

**Definition 1.3.5.** The code \( C \) is called self-orthogonal if \( C \subseteq C^\perp_H \)

and is called Hermitian self-dual if \( C = C^\perp_H \).

**Example 1.3.3.** The \([6,3] \) quaternary code \( G_6 \) has generator matrix \( G_6 \) in standard form

given by

\[
G_6 = \begin{bmatrix}
1 & 0 & 0 & 1 & \omega & \omega \\
0 & 1 & 0 & \omega & 1 & \omega \\
0 & 0 & 1 & \omega & \omega & \omega
\end{bmatrix}.
\]

This code is called the hexacode. It is Hermitian self-dual.

### 1.4 Weights and distances:

**Definition 1.4.1.** The Hamming distance \( d(x, y) \) between two vectors \( x, y \in F_4^n \) is defined to be the number of coordinates in which \( x \) and \( y \) differ and is denoted by \( d \). For example if \( x = 10112 \), and \( y = 20110 \), then \( d(x, y) = 2 \).
Theorem 1.4.1. The distance function \(d(x, y)\) satisfies the following four properties:

(i) (non-negativity) \(d(x, y) \geq 0\) for all \(x, y \in \mathbb{F}_q^n\).

(ii) \(d(x, y) = 0\), if and only if \(x = y\).

(iii) (Symmetry), \(d(x, y) = d(y, x)\) \(\forall x, y \in \mathbb{F}_q^n\).

(iv) (Triangle inequality), \(d(x, z) \leq d(x, y) + d(y, z)\) for all \(x, y, z \in \mathbb{F}_q^n\).

Proof. (i), (ii) and (iii) are obvious from the definition of the Hamming distance. It is enough to prove (iv) when \(n = 1\). If \(x = z\) then (iv) is obviously true since \(d(x, z) = 0\).

If \(x \neq z\), then either \(y \neq x\) or \(y \neq z\), so (iv) is again true. \(\square\)

Remark 1.4.1. \((C, d)\) is a metric space.

Definition 1.4.2. The minimum distance of a code \(C\) is the smallest Hamming distance \(d(x, y)\), where \(x \neq y\).

Definition 1.4.3. The minimum weight of \(x \in \mathbb{F}_q^n\), \(wt(x)\) is the number of nonzero coordinates in \(x\) which equals \(d(x, 0)\).

Theorem 1.4.2. If \(x, y \in \mathbb{F}_q^n\), then \(d(x, y) = wt(x - y)\). If \(C\) is a linear code, the minimum distance \(d\) is the same as the minimum weight of the nonzero codewords of \(C\).

Proof. \(d(x, y) = d(0, y - x) = wt(y - x)\) or \(wt(x - y)\) where \(y - x \in C\).

So the minimum distance \(\{d(x, y)\}, \text{ where } x \neq y, x, y \in C\) = the minimum weight \(\{wt(x - y)\}, \text{ where } x \neq y, x, y \in C\}.

Then the minimum distance \(d(C)\) = the minimum weight of nonzero codeword of \(C = \text{minimum}\{wt(a): a \neq 0, a \in C\}\). \(\square\)

Remark 1.4.2. If the minimum distance of the \([n, k]\)– code \(C\) is \(d\) then the code will now be defined as \([n, k, d]\) code.

Definition 1.4.4. If \(x = x_1x_2\cdots x_n\) and \(y = y_1y_2\cdots y_n\) are binary words then \(x \cap y = (x_1y_1, x_2y_2, \cdots , x_ny_n)\). Thus \(x \cap y\) has a 1 in the \(i\)th position if and only if both \(x, y\) have a 1 in the \(i\)th position.
Theorem 1.4.3. The following hold:

(i) If \( x, y \in F_2^n \), then \( wt(x + y) = wt(x) + wt(y) - 2wt(x \cap y) \), where \( x \cap y \) is a vector in \( F_2^n \), which has 1’s precisely in those positions where both \( x \) and \( y \) have 1’s.

(ii) If \( x, y \in F_2^n \), then \( wt(x \cap y) \equiv x \cdot y \mod 2 \).

(iii) If \( x \in F_2^n \), then \( wt(x) \equiv x \cdot x \mod 2 \).

(iv) If \( x \in F_3^n \), then \( wt(x) \equiv x \cdot x \mod 3 \).

(v) If \( x \in F_4^n \), then \( wt(x) \equiv \langle x \cdot x \rangle \mod 2 \).

Proof.

(i) If \( x, y \in F_2^n \), then \( wt(x + y) = wt(x - y) = d(x - y, 0) \) is the number of nonzero coordinates of \( x + \) the number of nonzero coordinates of \( y - 2( \text{the number of nonzero coordinates of } x \cap y ) = wt(x) + wt(y) - 2wt(x \cap y) \).

(ii) \( wt(x \cap y) = wt(x_1y_1, x_2y_2, \cdots, x_ny_n) = \) the number of nonzero coordinates of \( (x \cap y) \equiv (x_1y_1 + x_2y_2 + \cdots + x_ny_n) \mod 2 \).

(iii) If \( x \in F_2^n \), then \( wt(x) = \sum_{x_i \neq 0} x_i = \sum_{x_i \neq 0} x_i^2 \equiv x \cdot x \mod 2 \).

(iv) If \( x \in F_3^n \), then \( wt(x) = \sum_{x_i \neq 0} x_i = \sum_{x_i \neq 0} x_i^2 \equiv x \cdot x \mod 3 \).

(v) If \( x \in F_4^n \), then \( wt(x) = \sum_{x_i \neq 0} x_i = \sum_{x_i \neq 0} x_i^2 \equiv 2( x \cdot x \mod 2 ) \equiv \langle x \cdot x \rangle \mod 2 \).

Definition 1.4.5. Let \( A_i \) or \( A_i(C) \) be the number of codewords of weight \( i \) in \( C \). The list \( A_i \) for \( 0 \leq i \leq n \) is called the weight distribution or weight spectrum of \( C \).

Example 1.4.1. Let \( C \) be the binary code with generator matrix

\[
G = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 \\
\end{bmatrix}.
\]
All codewords are

\[000000, \quad 110000, \quad 111100, \quad 110011, \quad 001111, \quad 111111, \quad 001100, \quad 000011.\]

The weight distribution of \(C\) are \(A_0 = 1, \ A_6 = 1, \ A_2 = 3, \ A_4 = 3\)

\textbf{Theorem 1.4.4.} Let \(C\) be an \([n, k, d]\) code over \(\mathbb{F}_q\), then

(i) \(A_0(C) + A_1(C) + \cdots + A_n(C) = q^k.\)

(ii) \(A_0(C) = 1 \text{ and } A_1(C) = A_2(C) = \cdots = A_{d-1}(C) = 0.\)

(iii) If \(C\) is a binary code containing the codeword \(1 = 11\cdots1\), then \(A_i(C) = A_{n-i}(C)\) for \(0 \leq i \leq n.\)

(iv) If \(C\) is a binary self-orthogonal code, then each codeword has even weight, and \(C^\perp\) contains the codeword \(1 = 11\cdots1.\)

(v) If \(C\) is a ternary self-orthogonal code, then the weight of each codeword is divisible by three.

(vi) If \(C\) is a quaternary Hermitian self-orthogonal code, then the weight of each codeword is even.

\textbf{Proof.} (i) Since \(A_i(C)\) ranges over all codewords of \(C\) then \(A_0(C) + A_1(C) + \cdots + A_n(C) = q^k.\)

(ii) If \(d\) is the minimum distance of \(C\) then the minimum weight of \(C\) is \(d\) and no codeword in \(C\) with weight less than \(d\) so \(A_0(C) = 1\) and \(A_1(C) = A_2(C) = \cdots = A_{d-1}(C) = 0.\)

(iii) If \(C\) is binary code and \(1 = 11\cdots1\) is a codeword in \(C\), then if \(x \in C \ni wt(x) = i\) then \(wt(x+1) = wt(y) = n - i,\) where \(y = x + 1\) has weight \(n - i,\) then the number of words of weight \(i\) equals the number of words of weight \(n - i,\) that is \(A_i(C) = A_{n-i}(C).\)
(iv) Let $C \subseteq C^\perp$ and $C$ be binary code, if $x \in C$ then $x \cdot x \equiv 0 \pmod{2}$, then $wt(x)$ is even and $1 = 11 \cdots 1 \in C^\perp$, because it is orthogonal to any codeword $x \in C$ and because $x$ has even weight so $x \cdot 1 \equiv 0 \pmod{2}$.

(v) If $C \subseteq C^\perp$ and $C$ is any ternary code, if $x \in C$, then $wt(x) = x \cdot x \equiv 0 \pmod{3}$, then $wt(x)$ is divisible by 3.

(vi) If $C \subseteq C^\perp_H$ and $C$ is quaternary code then $\langle x, x \rangle \equiv 0 \pmod{2}$ implies $wt(x)$ is even, because $wt(x) = \langle x, x \rangle \equiv 0 \pmod{2}$.

Remark 1.4.3. The subset of codewords of a binary self-orthogonal code $C$ that have weight divisible by four form a subspace of $C$, this is not necessarily the case for non-self-orthogonal codes.

Theorem 1.4.5. Let $C$ be an $[n, k]$ self-orthogonal binary code. Let $C_0$ be the set of codewords in $C$ whose weights are divisible by four then either:

(i) $C = C_0$ or

(ii) $C_0$ is an $[n, k - 1]$ subcode of $C$ and $C = C_0 \cup C_1$, where $C_1 = x + C_0$ for any codeword $x$ whose weight is even but not divisible by four. Furthermore $C_1$ consists of all codewords of $C$ whose weights are not divisible by four.

Proof. Since $C$ is self-orthogonal then all codewords have even weight, then either (i) holds or there exists a codeword $x$ of even weight but not of weight a multiple of four.

Assume the later and let $y$ be another codeword whose weight is even but not a multiple of four, then by previous theorem (3.3) $wt(x + y) = wt(x) + wt(y) - 2wt(x \cap y) \equiv 2 + 2 - 2wt(x \cap y) \pmod{4}$. But $wt(x \cap y) \equiv x \cdot y \pmod{2}$ and $wt(x + y)$ is divisible by 4 then $x + y \in C_0$. This shows that $y \in x + C_0$ and $C = C_0 \cup (x + C_0)$. That is $C_0$ is a subcode of $C$ and that $C_1 = x + C_0$ consists of all codewords of $C$ whose weight are not divisible by 4 follow from a similar argument.

\[ \square \]
Theorem 1.4.6. Let $C$ be an $[n, k]$ binary code. Let $C_e$ be the set of codewords in $C$ whose weights are even. Then either

(i) $C = C_e$ or;

(ii) $C_e$ is an $[n, k]$ subcode of $C$ and $C = C_e \cup C_0$, where $C_0 = x + C_e$ for any codeword $x$ whose weight is odd. Furthermore $C_0$ consists of all codewords of $C$ whose weight are odd.

Theorem 1.4.7. Let $C$ be a binary linear code.

(i) If $C$ is self-orthogonal and has a generator matrix each of whose rows has weight divisible by 4, then every codeword of $C$ has weight divisible by 4.

(ii) If every codeword of $C$ has weight divisible by four, then $C$ is self-orthogonal.

Proof. (i) Let $x, y$ be tow rows of a generator matrix $G$, then $wt(x + y) = wt(x) + wt(y) - 2wt(x \cap y) \equiv 0 + 0 - 2wt(x \cap y) \equiv 0 \mod 4$. We proceed by induction as every codeword is a sum of rows of the generator matrix.

(ii) Let $x, y \in C$, then by previous theorem $wt(x \cap y \equiv x \cdot y \mod 2 \implies 2(x \cdot y) \equiv 2wt(x \cap y) \equiv 2wt(x \cdot y) - wt(x) - wt(y) \equiv -wt(x + y) \equiv 0 \mod 4 \implies x \cdot y \equiv 0 \mod 2 \implies C$ is self-orthogonal.

Example 1.4.2. The dual of the $[n, 1]$ binary repetition code $C$, consists of all the even weight vectors of length $n$ and has generator matrix

$$H = \begin{bmatrix} 1 & I_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$ If $n > 2$ this code is not self-orthogonal because any two different codewords are not orthogonal.
Theorem 1.4.8. Let $\mathcal{C}$ be a code over $\mathbb{F}_q^n$, with $q = 3$ or $4$.

(i) when $q = 3$, every codeword of $\mathcal{C}$ has weight divisible by three if and only if $\mathcal{C}$ is self-orthogonal.

(ii) when $q = 4$, every codeword of $\mathcal{C}$ has weight divisible by two if and only if $\mathcal{C}$ is Hermitian self-orthogonal.

Proof. (i) If $\mathcal{C}$ is self-orthogonal. Then the codewords have weights divisible by three (by theorem (1.3.4)(v).)

For the converse let $x, y \in \mathcal{C}$ and $x, y$ has weight divisible by three. We need to show that $x \cdot y = 0$. We can view the codewords $x$ and $y$ having the following parameters.

$$x : * 0 = \neq 0$$
$$y : 0 * = \neq 0$$
$$: a \ b \ c \ d \ e.$$ 

Where there are $a$ coordinates where $x$ is non zero and $y$ is zero, $b$ coordinates where $y$ is nonzero and $x$ is zero, $c$ coordinates where both agree and are nonzero, $d$ coordinates where both disagree and are nonzero, and $e$ coordinates where both are zero. So $wt(x + y) = a + b + c$ and $wt(x - y) = a + b + d$, but $x \pm y \in \mathcal{C} \implies a + b + c \equiv a + b + d \equiv 0 (mod 3)$. In particular $c \equiv d (mod 3)$ implies $x \cdot y = c + 2d \equiv 0 (mod 3)$ (because $c \equiv d (mod 3)$).

(ii) If $\mathcal{C}$ is Hermitian self-orthogonal, then the codewords have even weights, by the (1.3.4)(vi) $wt(x) = \langle x, x \rangle \equiv 0 (mod 2)$. For the converse, let $x \in \mathcal{C}$. If $x$ has $a$ 0’s $b$ 1’s, $c$’s and $d$’s then $b + c + d$ is even as $wt(x) = a + c + d$. However $\langle x, x \rangle$ also equals $b + c + d$ (as an element of $\mathbb{F}_4$), then $\langle x, x \rangle = 0$ for all $x \in \mathcal{C}$.

Now let $x, y \in \mathcal{C}$. So both $x + y$ and $\omega x + y$ are in $\mathcal{C}$. We have $0 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = 0 + \langle x, y \rangle + \langle y, x \rangle + 0$ .... (1)

Also $0 = \langle \omega x + y, \omega x + y \rangle = \langle x, x \rangle + \omega \langle x, y \rangle + \omega \langle y, x \rangle + \langle y, y \rangle = \omega \langle x, y \rangle + \omega \langle y, x \rangle + \omega \langle x, y \rangle$. 

17
From (1) and (2) \(0 = (1 + \omega)\langle x, y \rangle + (1 + \bar{\omega})\langle y, x \rangle = \bar{\omega}\langle x, y \rangle + \omega\langle y, x \rangle \implies \langle x, y \rangle = 0\) and \(\langle y, x \rangle = 0 \implies C\) is self-orthogonal.

**Theorem 1.4.9.** Let \(C\) be a binary code with a generator matrix each of whose rows has even weight. Then every codeword of \(C\) has even weight.

**Proof.** Let \(x, y\) be rows of the generator matrix \(wt(x + y) = wt(x) + wt(y) - 2wt(x \cap y) \equiv 0 + 0 - 2wt(x \cap y) \equiv 0 (mod\ 2) \implies wt(x + y)\) is even, so we proceed by induction every codeword is a sum of rows of the generator matrix. \(\square\)

**Definition 1.4.6.** Binary codes for which all codewords have weight divisible by 4 are called doubly even and by theorem 1.4.7 doubly even codes are self-orthogonal.

A self-orthogonal code must be even by theorem 1.4.4(iv).

The code that is not doubly even is called singly even.

**Definition 1.4.7.** A vector \(\mathbf{x} = x_1x_2 \cdots x_n \in \mathbb{F}_q^n\) is called even-like if \(\sum_{i=1}^{n} x_i = 0\) and is called odd like otherwise. A binary vector is even like if and only if it has even weight for \(\mathbb{F}_3\) the vector \((1, 1, 1) \in \mathbb{F}_3^3\) is even like but not of even weight. Also \((1, \omega, \bar{\omega}) \in \mathbb{F}_3^3\) is even like but not of even weight.

**Definition 1.4.8.** A code is even like if it has only even like codewords, a code is odd like if it is not even like.

**Theorem 1.4.10.** Let \(C\) be an \([n, k]\) code over \(\mathbb{F}_q\), let \(C_e\) be the set of even like codewords in \(C\), then either.

(i) \(C = C_e\) or,

(ii) \(C_e\) is an \([n, k - 1]\) subcode of \(C\).

**Proof.** Either (i) holds or if \(x, y \in C_e\) and \(\alpha, \gamma \in \mathbb{F}_q\implies \alpha \sum_{i=1}^{n} x_i = 0, \gamma \sum_{i=1}^{n} y_i = 0 \implies \sum_{i=1}^{n} \alpha x_i + \gamma y_i = 0 \implies \alpha x + \gamma y\) is even like \(\implies \alpha x + \gamma y \in C_e \implies C_e\) is a subcode of \(C\).
If \( z \) is odd like codeword, then \( z + C_e \) is a code of odd like codewords. Let \( C_1 = z + C_e \implies C = C_e \cup C_1 \) and \( dimC_e = k - 1 \).

The relationship between the weight of a codeword and a parity check matrix for a linear code.

**Theorem 1.4.11.** Let \( C \) be a linear code with parity check matrix \( H \). If \( c \in C \), the columns of \( H \) corresponding to the nonzero coordinates of \( C \) are linearly dependent. Conversely if a linear dependence relation with nonzero coefficients exists among \( w \) columns of \( H \), then there is a codeword in \( C \) of weight \( w \) whose nonzero coordinates correspond to these columns.

**Proof.** If \( G \) is a generator matrix of size \( k \times n \) for the code \( C \) and \( H \) is the parity check matrix for \( C \), then \( H \) is of size \( n - k \times n \) and \( rank(H) = n - k \). If \( c = c_1c_2 \cdots c_n \in C \) and if \( H_1, H_2, \cdots, H_n \) are the columns of \( H \), then

\[
Hc^\top = \begin{bmatrix} H_1 & H_2 & \cdots & H_n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = 0 \implies c_1H_1 + c_2H_2 + \cdots + c_nH_n = 0.
\]

Since \( c = c_1c_2 \cdots c_n \) is non zero codeword so there exists \( i, \ 1 \leq i < n \ \ni c_i \neq 0 \) then the columns of \( H \) corresponding to the nonzero coordinates of \( C \) are linearly dependent.

Conversely if there exists some \( i \) such that \( c_i \neq 0 \) and \( c_1H_1 + c_2H_2 + \cdots + c_wH_w = 0 \) where \( H_1, H_2, \cdots, H_w \) are the \( w \) columns of \( H \) which are linearly dependent, then there exists \( c = c_1c_2 \cdots c_n \neq 0 \in C \) such that \( Hc^\top = 0 \) and \( wt(c) = w \).

**Corollary 1.4.12.** A linear code has minimum weight \( d \) if and only if its parity check matrix has a set of \( d \) linearly dependent columns but no set of \( d - 1 \) linearly dependent columns.

**Proof.** Let \( H = \begin{bmatrix} H_1 & H_2 & \cdots & H_n \end{bmatrix} \) and \( c = c_1c_2 \cdots c_n \) be a codeword in \( C \) with Hamming weight \( w > 0 \). Let \( J = \{j_1, j_2, \cdots, j_w\} \) be the set of indexes of the nonzero
entries in \( c \). Since \( Hc^\top = 0 \) then \( \sum_{i=1}^{w} c_j H_{ji} = 0 \implies \) the columns of \( H \) indexed by \( J \) are linearly dependent.

Conversely, every set of \( w \) linearly dependent columns of \( H \) correspond to at least one nonzero codeword \( c \in C \) with \( \text{wt}(c) \leq w \).

Let \( d = \) the minimum distance of \( C \) then no nonzero codewords of \( C \) has Hamming weight less than \( d \), but there is at least one codeword in \( C \) whose hamming weight is \( d \).

**Example 1.4.3.** Consider the generator matrix

\[
G = \begin{bmatrix} I_n \mid A \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}
\]

for the Hamming binary \([7, 4]\) code \( H_3 \) its parity check matrix \( H \) is of the form

\[
H = \begin{bmatrix} A^\top \mid I_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.
\]

Every two columns in \( H \) are linearly independent and so, the minimum distance of the Hamming code \( H_3 \) is at least 3. In fact the minimum distance is exactly 3 since there are three dependent columns which are \((001)^\top\), \((0, 1, 0)^\top\), \((011)^\top\).

**Theorem 1.4.13.** If \( C \) is an \([n, k, d]\) code, then every \( n-d+1 \) coordinate position contains an information set. Furthermore, \( d \) is the largest number with this property.

**Proof.** Let \( G \) be a generator matrix for \( C \), consider any set \( X \) of \( s \) coordinate positions. Assume that \( X \) is the set of the last \( s \) positions. Suppose \( X \) does not contain an information set. Let \( G = \begin{bmatrix} A \mid B \end{bmatrix} \), where \( A \) is \( k \times (n-s) \) and \( B \) is \( k \times s \). Then the column rank of \( B \) and the row rank of \( B \), is less than \( k \). Hence there exists a non trivial linear combination of the rows of \( B \) which equals 0, and hence a codeword \( c \) which is 0 in the last \( s \) positions. Since the rows of \( G \) are linearly independent, \( c \neq 0 \) and hence \( d \leq n-s \implies s \leq n-d \). Then the theorem now follows. \( \square \)
Remark 1.4.4. for theorem 1.4.13.
The parity check matrix $H$ of an $[n, k, d]$ linear code $C$ is $(n-k) \times n$ matrix such that every $d-1$ columns of $H$ are linearly independent. Since the columns of $H$ have length $n-k$, we can never have more than $n-k$ independent columns, then $d-1 \leq n-k \implies k \leq n-d+1$.

1.5 New codes from old

1.5.1 Puncturing codes

Definition 1.5.1. Let $C$ be an $[n, k, d]$ code over $\mathbb{F}_q$, we can puncture $C$ by deleting the same coordinate $i$ in each codeword. the puncture code of $C$ denoted by $C^*$ has length $n-1$.

Theorem 1.5.1. Let $C$ be an $[n, k, d]$ code over $\mathbb{F}_q$, and let $C^*$ be the code $C$ punctured on the $i$th coordinate.

(i) If $d > 1$, $C^*$ is an $[n-1, k, d^*]$ code where $d^* = d-1$ if $C$ has a minimum weight codeword with a non zero $i$th coordinate and $d^* = d$ otherwise.

(ii) When $d = 1$, $C^*$ is an $[n-1, k, 1]$ code, if $C$ has no codeword of weight 1 whose nonzero entry is in coordinate $i$; otherwise, if $k > 1$, $C^*$ is an $[n-1, k-1, d^*]$ code with $d^* \geq 1$.

Example 1.5.1. Let $C$ be the $[5, 2, 2]$ binary code with generator matrix

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

Let $C^*_1$ and $C^*_5$ be the code $C$ punctured on coordinate 1 and 5, respectively, they have generator matrices

$$G^*_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}, \quad G^*_5 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$ 

So $C^*_1$ is a $[4, 2, 1]$ code, while $C^*_5$ is a $[4, 2, 2]$ code.
Example 1.5.2. Let $D$ be the $[4,2,1]$ binary code with generator matrix

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

Let $D_1^*$ and $D_4^*$ be the code $D$ punctured on coordinate 1 and 4, respectively, they have generator matrices

$$D_1^* = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad D_4^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$ 

So $D_1^*$ is a $[3,1,3]$ code, while $D_4^*$ is a $[3,2,1]$ code.

In general a code $C$ can be punctured on the coordinate set $T$ by deleting $T$ components in all codewords of $C$. If $T$ has size $t$, the resulting code, which we will often denoted $C^T$ is an $[n-t, k^*, d^*]$ where $k^* \geq k-t$, $d^* \geq d-t$.

1.5.2 Extending codes

We can create longer codes by adding a coordinate.

**Definition 1.5.2.** If $C$ is an $[n, k, d]$ code over $\mathbb{F}_q$, define the extended code $\hat{C}$ to be the code

$$\hat{C} = \{x_1 x_2 \cdots x_{n+1} \in \mathbb{F}_q^{n+1} | x_1 x_2 \cdots x_n \in C \text{ with } x_1 + x_2 + \cdots + x_{n+1} = 0\}.$$ 

In fact $\hat{C}$ is an $[n+1, k, \hat{d}]$ code, where $\hat{d} = d$ or $d+1$. If $G$ and $H$ be generator and parity check matrices for $C$, then a generator matrix $\hat{G}$ of $\hat{C}$ can be obtained from $G$ by adding an extra column to $G$, so the sum of the coordinates of each row of $\hat{G}$ is 0. A parity check matrix for $\hat{C}$ is the matrix

$$\hat{H} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & & & 0 \\ & H & & \\ & & \vdots & \end{bmatrix}.$$ 

This construction is referred to as adding an over all parity check.
Remark 1.5.1. * If $C$ is an $[n, k, d]$ binary code, then the extended code $C$ contains only even weight vectors and is an $[n + 1, k, \hat{d}]$ code, where $\hat{d}$ equals $d$ if $d$ is even and equals $d + 1$ if $d$ is odd.

** If $C$ is an $[n, k, d]$ code over $\mathbb{F}_q$, call the min. weight of the even-like codewords, respectively the odd-like codewords, the min. even-like weight, respectively the min. odd-like weight of the code. Denote the min. even like weight by $d_e$ and the min. odd-like weight by $d_0$. So $\hat{d} = \min[d_e, d_0]$. If $d_e \leq d_0$, then $\hat{C}$ has min. weight $\hat{d} = d_e$. If $d_0 < d_e$, then $\hat{d} = d_0 + 1$.

Example 1.5.3. The tetracode $H_{3,2}$ is a $[4, 2, 3]$ code over $\mathbb{F}_3$ with generator matrix $G$ and parity check matrix $H$ given by

$$G = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix} \text{ and } H = \begin{bmatrix} -1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{bmatrix}.$$ 

The codewords $(1, 0, 1, 1)$ extends to $(1, 0, 1, 1, 0)$ and the codeword $(0, 1, 1, -1)$ extends to $(0, 1, 1, -1, -1)$. Hence $d = d_e = d_0 = 3$ and $\hat{d} = 3$. the generator and parity check matrices for $\hat{H}_{3,2}$ are

$$\hat{G} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & -1 & -1 \end{bmatrix} \text{ and } \hat{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \end{bmatrix}.$$ 

Remark 1.5.2. If we extend a code and then puncture the new coordinate, we obtain the original code, but if we performing the operations in the other order will in general result in a different code.

Example 1.5.4. If we puncture the binary code $C$ with generator matrix

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

on its last coordinate and extend on the right, the resulting code has generator matrix

$$G = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$
1.5.3 Shortening codes

**Definition 1.5.3.** Let $C$ be an $[n, k, d]$ code over $F_q$ and let $T$ be any set of $t$ coordinates. Consider the set $C(T)$ of codewords which are $0$ on $T$; this set is a subcode of $C$, puncturing $C(T)$ on $T$ gives a code over $F_q$ of length $n - t$ called the code shortened on $T$ and denoted $C_T$.

**Example 1.5.5.** Let $C$ be the $[6, 3, 2]$ binary code with generator matrix

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

$C^\perp$ is also $[6, 3, 2]$ code with generator matrix

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

If the coordinates are labeled $1, 2, \cdots, 6$, let $t = \{5, 6\}$, then the generator matrices for the shortened code $C_T$ and the punctured code $C^T$ are

$$G_T = \begin{pmatrix}
1 & 0 & 1 & 0 \\
0 & 1 & 1 & 0
\end{pmatrix} \text{ and } G^T = \begin{pmatrix}
1 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1
\end{pmatrix}.$$  

because the set $C(T)$ of codewords which are zero on $T$ are

$$\begin{align*}
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0.
\end{align*}$$

By puncturing $C(T)$ on $T$ gives the code $C_T$ whose generator matrix is $G_T$.

The shortening and puncturing the dual code gives the code $(C^\perp)_T$ and $(C^\perp)^T$ which have
generator matrices

\[(G^\perp)_T = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}\] and \[(G^\perp)^T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.

From the generator matrices \(G_T\) and \(G^T\), we find that the duals of \(C_T\) and \(C^T\) have generator matrices

\[(G_T)^\perp = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\] and \[(G^T)^\perp = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}.

Notice that these matrices show that \((C^\perp)_T = (C^T)^\perp\) and \((C^\perp)^T = (C_T)^\perp\).

**Theorem 1.5.2.** Let \(C\) be an \([n, k, d]\) code over \(\mathbb{F}_q\). Let \(T\) be a set of \(t\) coordinates. Then:

(i) \((C^\perp)_T = (C^T)^\perp\) and \((C^\perp)^T = (C_T)^\perp\), and

(ii) if \(t < d\), then \(C^T\) and \((C^\perp)_T\) have dimensions \(k\) and \(n - t - k\), respectively;

(iii) if \(t = d\) and \(T\) is the set of coordinates where a minimum weight codeword is non-zero, then \(C^T\) and \((C^\perp)_T\) have dimensions \(k - 1\) and \(n - d - k + 1\), respectively.

**Proof.** (i) Let \(C\) be a codeword of \(C^\perp\) which is \(0\) on \(T\) and \(c^*\) the codeword with coordinates in \(T\) removed. So \(c^* \in (C^\perp)_T\). If \(x \in C\), then \(0 = x \cdot c = x^* \cdot c^*\) where \(x^*\) is the codeword \(x\) punctured on \(T\). Thus \((C^\perp)_T \subseteq (C^T)^\perp\)….\(1\). Any vector \(c \in (C^T)^\perp\) can be extended to the vector \(\hat{c}\) by inserting 0,s in the position of \(T\). If \(x \in C\), puncture \(x\) on \(T\) to obtain \(x^*\). As \(0 = x^* \cdot c = x \cdot \hat{c}\), \(c \in (C^\perp)_T \implies (C^T)^\perp \subseteq (C^\perp)_T\)…..\(2\). Thus from \((1)\) and \((2)\) we have \((C^\perp)^T = (C_T)^\perp\). We complete the proof of \((i)\)

(ii) Assume \(t < d\). Then \(n - d + 1 \leq n - t\), implying any \(n - t\) coordinates of \(C\) contain an information set (By theorem 1.4.13). Therefore \(C^T\) must be \(k\)-dimensional and hence \((C^\perp)_T = (C^T)^\perp\) has dimension \(n - t - k\) this proves \((ii)\).

(iii) As in (ii), (iii) is completed if we show that \(C^T\) has dimension \(k - 1\). If \(S \subset T\) with \(S\) of size \(d - 1\), \(C^S\) has dimension \(k\) by part (ii) clearly \(C^S\) has minimum distance 1 and
is obtained by puncturing \( C^s \) on the nonzero coordinate of a weight 1 codeword in \( C^s \). By theorem 1.5.1 \( C^T \) has dimension \( k - 1 \).

\[ \square \]

1.5.4 Direct sums

For \( i \in \{1, 2\} \) let \( C_i \) be an \([n_i, k_i, d_i]\) code, both over the same finite field \( \mathbb{F}_q \). Then their direct sum is the \([n_1 + n_2, k_1 + k_2, min\{d_1, d_2\}]\) code

\[
C_1 \oplus C_2 = \{(c_1, c_2) | c_1 \in C_1, \ c_2 \in C_2\}.
\]

If \( C_i \) has generator matrix \( G_i \) and parity check matrix \( H_i \), then

\[
G_1 \oplus G_2 = \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix} \quad \text{and} \quad H_1 \oplus H_2 = \begin{pmatrix} H_1 & 0 \\ 0 & H_2 \end{pmatrix}
\]

are a generator matrix and parity check matrix for \( C_1 \oplus C_2 \).

**Example 1.5.6.** Let \( C \) be the binary code with generator matrix

\[
G = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.
\]

Give another generator matrix for \( C \) that shows that \( C \) is a direct sum of two binary codes.
Solution: By elementary row operations $G$ has equivalent matrix in the form

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

which is equivalent to

$$
G_1 \oplus G_2 = \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix},
$$

where

$$
G_1 = \begin{pmatrix}
1 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 \\
0 & 0 & 1 & 1 \\
\end{pmatrix}
$$
is a generator matrix for $[4, 3, 2]$ code $C_1$ and

$$
G_2 = \begin{pmatrix}
0 & 1 & 1 \\
1 & 1 & 0 \\
\end{pmatrix}
$$
is a generator matrix for $[3, 2, 2]$ code $C_2$.

1.5.5 The $(u|u+v)$ construction

Two codes of the same length can be combined to form a third code of twice of length in a way similar to the direct sum construction. Let $C_i$ be an $[n, k_i, d_i]$ code for $i \in \{1, 2\}$ over $F_q$. Then the $(u|u+v)$ construction produces the $[2n, k_1 + k_2, \min\{2d_1, d_2\}]$-code

$$
C = \{ (u|u+v) | u \in C_1, v \in C_2 \}.
$$

If $C_i$ has the generator matrix $G_i$ and the parity check matrix $H_i$, then the generator matrix and the parity check matrices for $C$ are

$$
\begin{pmatrix} G_1 & G_1 \\ 0 & G_2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} H_1 & 0 \\ -H_2 & H_2 \end{pmatrix}.
$$

27
Example 1.5.7. Consider the $[8, 4, 4]$-binary code $C$ with generator matrix
\[
\begin{pmatrix}
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1
\end{pmatrix}.
\]
Then $C$ can be produced from the $[4, 3, 2]$ code $C_1$ and the $[4, 1, 4]$ code $C_2$ with generator matrices
\[
G_1 = \begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}, \quad \text{and} \quad G_2 = \begin{pmatrix}
1 & 1 & 1 \\
1 & 1
\end{pmatrix},
\]
respectively using the $(u|u+v)$ construction. Notice that the code $C_1$ is also constructed using the $(u|u+v)$ construction from the $[2, 2, 1]$ code $C_3$ and the $[2, 1, 2]$ code $C_4$ with generator matrices
\[
G_3 = \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix} \quad \text{and} \quad G_4 = \begin{pmatrix}
1 \\
1
\end{pmatrix}.
\]

Remark 1.5.3. The $(u|u+v)$ construction are used to construct the family of Reed Muller codes.

Example 1.5.8. Prove that the $(u|u+v)$ construction using $[n, k_i, d_i]$ codes $C_i$ produces a code of dimension $k = k_1 + k_2$ and minimum weight $d = \min\{2d_1, d_2\}$.

Proof. Let $x_1 = (u_1|u_1 + v_1)$ and $x_2 = (u_2|u_2 + v_2)$ be distinct two codewords in the $(u|u+v)$ construction.

If $v_1 = v_2$, then
\[
d(x_1, x_2) = 2d(u_1, u_2) \geq 2d_1.
\]
On the other hand, if $v_1 \neq v_2$ then
\[
d(x_1, x_2) = wt(x_1 - x_2) = wt(u_1 - u_2) + wt(u_1 - u_2 + v_1 - v_2) \geq wt(v_1 - v_2) = d(v_1, v_2) \geq d_2.
\]
Hence the minimum distance of the $(u|u+v)$ construction code $d \geq \min\{2d_1, d_2\}$, since equality can hold in all cases, then $d = \min\{2d_1, d_2\}$. The dimension of $C$ is equal $k_1 + k_2$ from the generator matrix of $(u|u+v)$ construction. \qed
1.6 Permutation equivalent codes:

Definition 1.6.1. two linear codes $C_1$ and $C_2$ are called permutation equivalent provided there is a permutation of coordinates which sends $C_1$ to $C_2$. This permutation can be described using a permutation matrix, which is a square matrix with exactly one 1 in each row and column and 0’s elsewhere.

* $C_1$ and $C_2$ are permutation equivalent provided that there is a permutation matrix $P$ such that $G_1$ is a generator matrix of $C_1$ if and only if $G_1P$ is a generator matrix of $C_2$.

** If $P$ is a permutation sending $C_1$ to $C_2$ we will write $C_1P = C_2$ when $C_1P = \{ y | y = xP \text{ for } x \in C_1 \}$.

Example 1.6.1. If $C_1$ and $C_2$ are permutation equivalent codes then

(1) $C_1^\perp P = C_2^\perp$.

(2) If $C_1$ is self dual, so is $C_2$.

Proof. (1) $C_1^\perp = \{ x \in \mathbb{F}_q^n : x \cdot c_1 = 0 \ \forall c_1 \in C_1 \}$, then

$$C_1^\perp P = \{ xP \in \mathbb{F}_q^n : xP \cdot c_1P = 0 \ \forall c_1 \in C_1 \} = \{ y \in \mathbb{F}_q^n : y \cdot c_2 \ \forall c_2 \in C_2 \} = C_2,$$

where $xP = y$ and $c_2 = c_1P$.

(2) Let $C_1 = C_1^\perp \implies C_1^\perp P = C_2^\perp \implies C_1P = C_2^\perp \implies C_2 = C_2^\perp$. so $C_2$ is self dual.

Example 1.6.2. Let $C_1$, $C_2$ and $C_3$ be binary codes with generator matrices

$$G_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \quad G_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad G_3 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

These codes have weight distributions $A_0 = A_6 = 1$ and $A_2 = A_4 = 3$. The permutation switching columns 2 and 6 sends $G_1$ to $G_2$, showing that $C_1$ and $C_2$ are permutation
equivalent. $C_1$ and $C_2$ are self-dual. $C_3$ is not self-dual so $C_1$ and $C_3$ are not permutation equivalent.

**Theorem 1.6.1.** Let $C$ be a linear code.

(i) $C$ is permutation equivalent to a code which has generator matrix in standard form.

(ii) If and $I$ and $R$ are information and redundancy positions for $C$, then $R$ and $I$ are information and redundancy positions respectively for the dual code $C^\perp$.

**Proof.**

(i) If we apply the elementary row operations to any generator matrix of $C$. This will produce a new generator matrix of $C$ which has columns the same as those in $I_k$, but possibly in a different order. Now choose a permutation of the columns of the new generator matrix so that these columns are moved to the order that produces $[I_k|A]$. The code generated by $[I_k|A]$ is equivalent to $C$.

(ii) If $I$ is an information set for $C$, then by row reducing a generator matrix for $C$, we obtain columns in the information positions that are the columns of $I_k$ in some order. As above, choose a permutation matrix $P$ to move the columns so that $CP$ has generator matrix $[I_k|A]$; $P$ has moved $I$ to the first $k$ coordinate positions. By previous theorem (2.1.2) $(CP)^\perp$ has the last $n-k$ coordinates as information positions, but $(CP)^\perp = C^\perp P$, imply that $R$ is a set of information positions for $C^\perp$. 

Let $\text{Sym}_n$ be the set of all permutations of the set of $n$ coordinates. If $\sigma \in \text{Sym}_n$ and $x = x_1x_2 \cdots x_n$, define $x\sigma = y_1y_2 \cdots y_n$ where $y_j = x_{j\sigma^{-1}}$ for $1 \leq j \leq n$.

So $x\sigma = xP$, where $P = [p_{i,j}]$ is the permutation matrix given by

\[
p_{i,j} = \begin{cases} 
1 & \text{if } j = i\sigma, \\
0 & \text{otherwise.}
\end{cases}
\]
Example 1.6.3. Let \( n = 3 \), \( x = x_1x_2x_3 \), \( \sigma = (1, 2, 3) \).

Then \( 1\sigma^{-1} = 3 \), \( 2\sigma^{-1} = 1 \), \( 3\sigma^{-1} = 2 \), so \( x\sigma = x_3x_1x_2 \).

Let \( P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \).

Then \( xP \) also equals \( x_3x_1x_2 \).

Definition 1.6.2. The set of coordinate permutation that map a code \( C \) to itself forms a group. That is a set with an associative binary operation which has an identity and all elements have inverses, called permutation automorphism group of \( C \). This group is denoted by \( \text{PAut}(C) \). So if \( C \) is a code of length \( n \), then \( \text{PAut}(C) \) is a subgroup of the symmetric group \( \text{Sym}_n \).

Example 1.6.4. Show that \( (1, 2)(5, 6) \), \( (1, 2, 3)(5, 6, 7) \) and \( (1, 2, 4, 5, 7, 3, 6) \)
are automorphisms of the \([7, 4]\) binary code \( H_3 \) which has the following generator matrix,

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & 1
\end{pmatrix}
\]

The Automorphism \( (1, 2)(5, 6) \) gives the generator matrix of the form

\[
\begin{pmatrix}
0 & 1 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & 1
\end{pmatrix}
\]

This generator matrix is the same as the above generator matrix if we interchange rows 1 and 2.

Theorem 1.6.2. Let \( C, C_1 \) and \( C_2 \) be codes over \( \mathbb{F}_q \). Then:

(i) \( \text{PAut}(C) = \text{PAut}(C^\perp) \).
(ii) If $q = 4$, $PAut(C) = PAut(C^\perp)$

(iii) If $C_1P = C_2$ for a permutation matrix $P$, then $P^{-1}PAut(C_1)P = PAut(C_2)$.

Proof.

(i) Let $\phi \in PAut(C) \implies \phi(C) = C$, then there exists a permutation matrix $P \ni CP = C$.
But $(CP)^\perp = C^\perp P = C$. Then $P$ is a permutation matrix for $C^\perp$ corresponding to $\phi$, so $\phi \in PAut(C^\perp) \implies PAut(C) \subseteq PAut(C^\perp)........(1)$.

Similarly if $\alpha \in PAut(C^\perp) \implies \alpha C^\perp = C^\perp \implies (\alpha C^\perp)^\perp = (C^\perp)^\perp = C$.

then $\alpha C = C \implies \alpha \in PAut(C) \implies PAut(C^\perp) \subseteq PAut(C)........(2)$.

From (1) and (2) we have $PAut(C) = PAut(C^\perp)$.

(ii) Same as (i)

(iii) If $C_1P = C_2$ for a permutation matrix $P$, then $P^{-1}PAut(C_1)P = PAut(C_2)$, for $P \notin PAut(C_1)$.

If $Q \in PAut(C_1) \implies QC_1 = C_1 \implies P^{-1}QC_1P = P^{-1}C_1P$

$\implies P^{-1}QC_2 = P^{-1}C_1P$.

Then $PP^{-1}QC_2 = PP^{-1}C_1P = I_nC_1P = C_2$

$\implies QC_2 = C_2 \implies Q \in PAut(C_2) \implies P^{-1}PAut(C_1)P \subseteq PAut(C_2)........(1)$.

Let $Q \in PAut(C_2) \implies QC_2 = C_2 \implies QC_1P = C_1P \implies QC_1PP^{-1} = C_1PP^{-1} \implies QC_1 = C_1 \implies Q \in PAut(C_1) \implies PAutC_2 \subseteq P^{-1}PAutC_1P........(2)$ From (1) and (2) we have $P^{-1}PAut(C_1)P = PAut(C_2)$.

\[\square\]

1.7 Hamming codes

We have obtained the parity check matrix of the Hamming code $H_3$, $[7, 4, 3]$ in the form

\[
H = \begin{pmatrix}
0 & 1 & 1 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1
\end{pmatrix}.
\]
Notice that the columns of this parity check matrix are all the distinct nonzero binary columns of length 3. So $H_3$ is equivalent to the code with parity check matrix,

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

Whose columns are the numbers 1 through 7 written as binary numerals (with leading 0’s as necessary to have a 3-tuple in their natural order.

We generalize this form in the following.

Let $n = 2^r - 1$ with $r \geq 2$. Then the $r \times (2^r - 1)$ matrix $H_r$ whose columns are $1, 2, \cdots 2^r - 1$ written as binary numerals is the parity check matrix of the hamming code $[n = 2^r - 1, k = n - r]$ binary code. Any arrangement of columns of $H_r$ gives an equivalent code, these codes denoted by $H_r$ or $H_{3r}$.

The columns of $H_r$ are distinct and nonzero, the minimum distance is at least 3. Since there are three columns linearly dependent, so the minimum distance of the Hamming code is 3. So $H_r$ is a binary $[2^r - 1, 2^r - 1 - r, 3]$ code and these codes are unique.

**Theorem 1.7.1.** Any $[2^r - 1, 2^r - 1 - r, 3]$ binary code is equivalent to the binary hamming code $H_r$.

**Proof.** The parity check matrix is $r \times (2^r - 1)$, so $\text{dim}(C^\perp) = r$ and the maximum number of linearly independent columns is 2, so $d = 3$ and $\text{dim}(C) = 2^r - 1 - r$. \hfill $\square$

Hamming codes $H_{q,r}$ can be defined over an arbitrary finite field $F_q$. For $r \geq 2$, $H_{q,r}$ has parity check matrix $H_{q,r}$ defined by choosing for its columns a non zero vector from each 1-dimensional subspace of $F_q^r$. There are $\frac{q^r-1}{q-1}$ 1-dimensional subspaces. Therefore $H_{q,r}$ has length $n = \frac{(q^r-1)}{q-1}$, dimension $n - r$, redundancy $r$. No two columns are multiple of each other. So $H_{q,r}$ has minimum distance 3.

**Theorem 1.7.2.** Any $[\frac{(q^r-1)}{q-1}, \frac{(q^r-1)}{q-1} - r, 3]$ code over $F_q$ is monomially equivalent to the Hamming code $H_{q,r}$.
Remark 1.7.1. The duals of the Hamming codes are called simplex codes. They are \([\frac{q^r-1}{q-1}, r]\) codes, denoted by \(S_r\) and has minimum weight \(q^{r-1}\).

Theorem 1.7.3. The non zero codewords of the \([\frac{q^r-1}{q-1}, r]\) simplex code over \(\mathbb{F}_q\) all have weights \(q^{r-1}\).

Construction of binary simplex code.

Let \(G_2\) be the matrix

\[
G_2 = \begin{pmatrix}
0 & 1 & 1 \\
1 & 0 & 1 \\
\end{pmatrix},
\]

for \(r \geq 3\), define \(G_r\) inductively by

\[
G_r = \begin{pmatrix}
0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots \\
0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots \\
G_{r-1} & \vdots & G_{r-1} \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\end{pmatrix}.
\]

\(G_r\) has \(n_r = 2^r - 1\) distinct nonzero columns. The weight of the simplex code is \(2^{r-1}\).

We claim the code \(S_r\) generated by \(G_r\) is the dual of the Hamming code \(H_{r-1}\). Since \(G_r\) has one more row than \(G_{r-1}\) and, as \(G_2\) has 2 rows, \(G_r\) has \(r\) rows, let \(G_r\) have \(n_r\) columns. So \(n_2 = 2^2 - 1\) and \(n_r = 2n_{r-1} + 1\); by induction \(n_r = 2^r - 1\). The columns of \(G_2\) are nonzero and distinct; by construction the columns of \(G_r\) are nonzero and distinct if the columns of \(G_{r-1}\) are also nonzero and distinct. So by induction \(G_r\) has \(2^r - 1\) distinct nonzero columns of length \(r\). But there are only \(2^r - 1\) possible distinct nonzero \(r\)-tuples; these are the binary expansions of \(1, 2, \cdots, 2^r - 1\). So \(S_r = H_{r-1}^\perp\).

The nonzero codewords \(S_2\) have weight 2. Assume the nonzero codewords of \(S_{r-1}\) have weight \(2^{r-2}\). Then the nonzero codewords of the subcode generated by the last \(r-1\) rows of \(G_r\) have weight \((a,0,b)\), where \(a,b \in S_{r-1}\). So these codewords have weight \(2 \cdot 2^{r-2} = 2^{r-1}\). Also the top row of \(G_r\) has weight \(1 + 2^{r-1} - 1 = 2^{r-1}\). The remaining non zero codewords of \(S_r\) have the form \((a,1,b+1)\), where \(a,b \in S_{r-1}\). As \(wt(b+1) = 2^{r-2} - 1\), \(wt(a,1,b+1) = 2^{r-2} + 1 + 2^{r-2} - 1 = 2^{r-1}\). Thus by induction \(S_r\) has all nonzero codewords of weight \(2^{r-1}\).
1.8 The Golay codes

The binary codes of length 23 and the ternary code of length 11 were first described by Golay in 1949.

1.8.1 The Golay code $G_{24}$

We let $G_{24}$ be the $[24,12]$ code with generator matrix $G_{24} = [I_{12} | A]$ in standard form where $I_{12}$ is the identity matrix and $A$ is a matrix of size $12 \times 12$ defined by

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

Lable the columns of $A$ by $\infty$, 0, 1, 2, $\cdots$ 10. The first row contains 0 in column $\infty$ and 1 elsewhere. To obtain the second row, a 1 is placed in column $\infty$ and a 1 is placed in columns 0, 1, 3, 4, 5 and 9, these numbers are the squares of the integers modulo 11. That is $0^2 = 0, 1^2 = 10^2 \equiv 1 \pmod{11}$, $2^2 \equiv 9^2 \equiv 4 \pmod{11}$ etc.

The first third row of $A$ is obtained by putting a 1 in column $\infty$ and then shifting the components in the second row one place to the left and wrapping the entry in column 0 around to column 10, and so on all other rows.
Remark 1.8.1. (1) All rows has weight divisible by 4 and \( \dim \mathcal{G}_{24} = 12 \), it is self-dual binary code.

(2) The minimum weight of \( \mathcal{G}_{24} \) is 8

(3) If we puncture in any of the coordinates we obtain a \([23, 12, 7]\) binary code \( \mathcal{G}_{23} \) called binary Golay code has minimum weight 7

(4) The extended code of \( \mathcal{G}_{23} \) is \( \mathcal{G}_{24} \) so \( \mathcal{G}_{24} \) is called extended Golay code.

1.8.2 The ternary Golay codes

The ternary Golay code \( \mathcal{G}_{12} \) is \([12, 6, 6]\) code over \( \mathbb{F}_3 \) with generator matrix \( G_{12} = [I_6|A] \) in standard form where

\[
A = \begin{pmatrix}
0 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & -1 & -1 & 1 \\
1 & 1 & 0 & 1 & -1 & -1 \\
1 & -1 & 1 & 0 & 1 & -1 \\
1 & -1 & -1 & 1 & 0 & 1 \\
1 & 1 & -1 & -1 & 1 & 0
\end{pmatrix}
\]

Remark 1.8.2. (1) The \( \mathcal{G}_{12} \) is self-dual ternary code \([12, 6, 6]\)

(2) The \( \mathcal{G}_{11} \) is a \([11, 6, 5]\) code obtained from \( \mathcal{G}_{12} \) by puncturing

(3) The extended code of \( \mathcal{G}_{11} \) is not \( \mathcal{G}_{12} \) it will give either a \([12, 6, 6]\) code or a \([12, 6, 5]\) code depending upon the coordinate.

1.9 Reed- Muller codes

Definition 1.9.1. Let \( m \) be a positive integer and \( r \) a nonnegative integer with \( r \leq m \). The binary codes we construct will have length \( 2^m \). For each length there will be \( m + 1 \) linear codes denoted by \( R(r, m) \) and called the \( r \)th order Reed-Muller or \( RM \), code of
length $2^m$.

* The codes $R(0, m)$, $R(m, m)$ are trivial codes the first is called 0th order $RM$ code $R(0, m)$ which is a binary repetition code of length $2^m$ with bases $[1]$, and the $m$th order $RM$ code $R(m, m)$ is the entire space $\mathbb{F}_2^{2^m}$.

** For $1 \leq r \leq m$, define $$R(r, m) = \{(u, u + v) | u \in R(r, m-1), v \in R(r-1, m-1)\}.$$  

*** Let $G(0, m) = [1 \cdots 1]$ and $G(m, m) = I_{2^m}$ be the generator matrices for $R(0, m)$ and $R(m, m)$ respectively for $1 \leq r < m$, using $(u|u+v)$ construction, a generator matrix $G(r, m)$ for $R(r, m)$ is

$$G(r, m) = \begin{pmatrix} G(r, m-1) & G(r, m-1) \\ 0 & G(r-1, m-1) \end{pmatrix}$$

Example 1.9.1. The generator matrices for $R(r, m)$ with $1 \leq r \leq 3$ are

$$G(1, 2) = \begin{pmatrix} G(1, 1) \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

and

$$G(1, 3) = \begin{pmatrix} G(1, 2) \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$
and

\[
G(2, 3) = \begin{pmatrix}
G(2, 2) & G(2, 2) \\
0 & G(1, 2)
\end{pmatrix} =
\begin{pmatrix}
1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 1 & 0
\end{pmatrix}
\]

Note: (1) \(R(1, 2)\) and \(R(2, 3)\) are both the set of all even weight vectors in \(\mathbb{F}^4\) and \(\mathbb{F}^8\) respectively.

(2) \(R(1, 3)\) is an \([8, 4, 4]\) self-dual code which must be the extended Hamming code \(\mathcal{H}_3\).

**Theorem 1.9.1.** Let \(r\) be an integer with \(0 \leq r \leq m\). Then the following are hold:

(i) \(R(i, m) \subseteq R(j, m)\) if \(0 \leq i \leq j \leq m\).

(ii) The dimension of \(R(r, m)\) equals

\[
\binom{m}{0} + \binom{m}{1} + \cdots + \binom{m}{r}.
\]

(iii) The minimum weight of \(R(r, m)\) equals \(2^{m-r}\).

(iv) \(R(m, m)^\perp = \{0\}\), and if \(0 \leq r < m\), then \(R(r, m)^\perp = R(m-r-1, m)\).

### 1.10 Encoding, decoding, and Shannon’s theorem

Let \(C\) be an \([n, k]\) linear code over the field \(\mathbb{F}_q\), with generator matrix \(G\). This code has \(q^k\) codewords which will be in one to one correspondence with \(q^k\) messages. The simplest way to view these messages is as \(k\)–tuples \(x \in \mathbb{F}_q^k\).

To encode the message \(x\) as a codeword \(c = xG\), if \(G = [I_k|A]\) in standard form, then the first \(k\) coordinates of the codeword \(C\) are the information symbol \(x\); the remaining \(n-k\)
symbols are the parity check symbols, that is a redundancy added to \( x \) in order to help recover \( x \) if errors occur.

If \( G \) is not in standard form then there exists indices \( i_1, i_2, \ldots, i_k \) such that the \( k \times k \) matrix consisting of these \( k \) columns of \( G \) is the \( k \times k \) identity matrix consisting of those \( k \) columns of \( G \). Then the message is found in the \( k \)– coordinates \( i_1, i_2, \ldots, i_k \) of the codeword scrambled but otherwise unchanged, that is the message symbol \( x_j \) is in the component \( i_j \) of the codeword. This encoder is called systematic.

Encoder

Let \( x \) be a message \( x = x_1 x_2 \cdots x_k \).

Let \( G = [I_k | A] \), \( H = [-A^T | I_{n-k}] \). Suppose \( x = x_1 x_2 \cdots x_k \) is encoded as a codeword \( c = c_1 c_2 \cdots c_n \) as \( G \) in standard form, \( c_1 c_2 \cdots c_k = x_1 x_2 \cdots x_k \). So we need to determine the \( n - k \) parity check symbols (redundancy symbols) \( c_{k+1} c_{k+2} \cdots c_n \).

As \( 0 = HC^T = [-A^T | I_{n-k}] C^T \implies 0 = -A^T x^T + I_{n-k}[c_{k+1} \cdots c_n]^T \implies A^T x^T = [c_{k+1} \cdots c_n]^T \)

**Example 1.10.1.** Let \( G \) be the \([6,3,3]\) binary code with generator matrix and parity check matrices

\[
G = \begin{pmatrix}
1 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1
\end{pmatrix}
\quad \text{and} \quad
H = \begin{pmatrix}
1 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 1
\end{pmatrix}
\]

Let \( x = x_1 x_2 x_3 \) to obtain the codeword \( C = c_1 c_2 \cdots c_6 \) using \( G \) to encode yields \( C = xG = (x_1, x_2, x_3, x_1 + x_2, x_2 + x_3, x_1 + x_3) \). Using \( H \) to encode \( 0 = HC^T \) leads to the system

\[
0 = c_1 + c_2 + c_4 \\
0 = c_2 + c_3 + c_5 \\
0 = c_1 + c_3 + c_6
\]

As \( G \) in standard form \( c_1 c_2 c_3 = x_1 x_2 x_3 \) and solving this system clearly gives the same codeword

\[
c_4 = c_1 + c_2 = x_1 + x_2 \\
c_5 = c_2 + c_3 = x_2 + x_3 \\
c_6 = c_1 + c_3 = x_1 + x_3
\]
\[ C = (x_1, x_2, x_3, x_1 + x_2, x_2 + x_3, x_1 + x_3). \]

If \( G \) is not in standard form, since \( G \) has \( k \) independent rows, so there exists \( n \times k \) matrix \( K \) such that \( GK = I_k \), \( K \) is called a right inverse for \( G \) and is not necessarily unique. As \( c = xG \implies cK = xGK = xI_k = x \).
1.10.1 Decoding and Shannon’s theorem

The process of decoding, that is, determining which codeword was sent when a vector $y$ is received.

Binary symmetric channel is a mathematical model of a channel that transmits binary data. This model is called (BSC) with crossover probability $p$ and illustrated in figure 1.1.

If 0 or 1 is sent, the probability it received without error is $1 - p$, If a 0 (respectively 1) is sent, the probability that a 1 (respectively 0) is received is $p$, where $0 \leq p < \frac{1}{2} < 1$.

\[
\text{Prob}(y_i \text{ was received} | c_i \text{ was sent}) = \begin{cases} 
1 - p, & \text{if } y_i = c_i; \\
p, & \text{if } y_i \neq c_i,
\end{cases}
\]

where $y_i, c_i \in \mathbb{F}_2$.

**Discrete memoryless channel** (or DMC) a channel in which inputs and outputs are discrete and the probability of error in one bit is independent of previous bits. $0 \leq p < \frac{1}{2}$.

**conditional probability**
Assume $c \in \mathbb{F}_2^n$ is sent and $y \in \mathbb{F}_2^n$ is received and decoded as $\hat{c} \in \mathbb{F}_2^n$.

$$prop(c|y) = \frac{prop(y|c)prop(c)}{prop(y)}$$

where $prop(c)$ is the probability that $c$ is sent and $prop(y)$ is received.

There is two choice of the decoder

(1) The decoder could choose $\hat{c} = c$ for the codeword $c$ with $prop(c|y)$ max. Such decoder is called a maximum a posteriori probability or (MAP) decoder, in symbols

MAP decoder makes the decision $\hat{c} = \arg \max_{c \in C} prop(c|y)$. Here $arg \max_{c \in C} prop(c|y)$, is the argument $c$ of the probability function $prop(c|y)$ that maximizes this probability.

(2) The decoder could choose $\hat{c} = c$ for the codeword $c$ with $prop(y|c)$ maximum, such a decoder is called a maximum likelihood (or ML) decoder in symbols, a ML decoder makes the decision $\hat{c} = \arg \max_{c \in C} prop(y|c)$.

Consider ML decoding over a B.S.C. if $y = y_1y_2 \cdots y_n \in \mathbb{F}_2^n$, and $c = c_1c_2 \cdots c_n \in \mathbb{F}_2^n$

$$prob(y|c) = \prod_{i=1}^{n} prob(y_i|c_i) = p^{d(y,c)}(1-p)^{n-d(y,c)} = (1-p)^n \left( \frac{p}{1-p} \right)^{d(y,c)},$$

where $0 < p < \frac{1}{2}$, $0 < \frac{p}{1-p} < 1$.

So $prob(y|c)$ is maximum when $d(y,c)$ is minimum. That is finding the codeword $c$ is closest to the received vector $y$ in Hamming distance, this is called nearest neighbor decoding. Hence A B.S.C, maximum likelihood and nearest neighbor decoding are the same.

Let $e = y - c$ so that $y = c + e$, where $e$ is an error vector added to the codeword $c$ on the effect of noise in the communication channel.

The goal of decoding is to determine $e$.

Nearest neighbor decoding is equivalent to finding a vector $e$ of smallest weight such that $y - e$ is in the code.

$e$ need not be unique since there may be more than one codeword closest to $y$. 

42
When we have a decoder capable of finding all codewords nearest to the received vector \( y \), then we have a complete decoder.

**Sphere of radius** \( r \) A set of words in \( \mathbb{F}_q^n \) at Hamming distance \( r \) or less from a given codeword \( u \in \mathbb{F}_q^n \) is called a sphere of radius \( r \).

\[
S_r(u) = \{ v \in \mathbb{F}_q^n | d(u, v) \leq r \}.
\]

The number of words or vectors in \( S_r(u) \) or the volume of \( S_r(u) \) is equal

\[
\sum_{i=0}^{r} \binom{n}{i} (q - 1)^i.
\]

**Theorem 1.10.1.** If \( d \) is the minimum distance of a code \( c \) (linear or non linear) and \( t = \lfloor \frac{d - 1}{2} \rfloor \), then spheres of radius \( t \) about distinct codewords are disjoint.

**Proof.** If \( z \in S_t(c_1) \cap S_t(c_2) \), where \( c_1, c_2 \) are codewords, then by triangle inequality

\[
d(c_1, c_2) = d(c_1 - z, c_2 - z) \leq d(c_1, z) + d(c_2, z) \leq 2t \leq d - 1 < d \text{ contradiction so } c_1 = c_2.
\]

**Corollary 1.10.2.** with the notation of previous theorem, if a codeword \( c \) is sent and \( y \) is received where \( t \) or fewer errors have occurred, then \( c \) is the unique codeword closest to \( y \). In particular, nearest neighbor decoding uniquely and correctly decodes any received vector in which at most \( t \) errors occurred in transmission.

**Proof.** Let \( c \) and \( y \) be the transmitted codeword and received word, respectively, and write \( y = c + e \) where \( \text{wt}(e) \leq \frac{d - 1}{2} \), suppose that \( c' \neq c \) and \( c' \) is closest to \( y \) in \( C \) also,

\[
d(c', y) \leq d(c, y) \leq \frac{d - 1}{2}.
\]

By triangle inequality

\[
d \leq d(c', c) \leq d(c, y) + d(c', y) \leq d - 1
\]

which is a contradiction, so \( c' = c \). \(\square\)

**Remark 1.10.1.** (1) Given \( n \) and \( d \), a code with largest number of codewords, with highest dimension and high minimum weight is an efficient code for decoding
(2) Since the minimum distance of $C$ is $d$, there exist two distinct codewords such that the spheres of radius $t + 1$ about them are not disjoint. Therefore, if more than $t$ errors occur, nearest neighbor decoding may yield more than one nearest codeword. Thus $C$ is a $t$-error correcting code but not $(t + 1)$-error correcting code.

(3) The packing radius of a code is the largest radius of spheres centered at codewords so that the spheres are pairwise disjoint.

**Theorem 1.10.3.** Let $C$ be an $[n, k, d]$ code over $\mathbb{F}_q$. The following hold:

(i) The packing radius of $C$ equals $t = \left\lfloor \frac{d-1}{2} \right\rfloor$.

(ii) The packing radius $t$ of $C$ is characterized by the property that nearest neighbor decoding always decode correctly a received vector in which $t$ or fewer errors have occurred but will not always decode correctly a received vector in which $t + 1$ errors have occurred.

The decoding problem becomes one of finding an efficient algorithm that will correct up to $t$ errors one of the most obvious decoding algorithm is to examine all codewords until one is found with distance $t$ or less from the received vector. But this is efficient for codes of number of codewords.

Another obvious algorithm is to make a table consisting of a nearest codeword for each of the $q^n$ vectors in $\mathbb{F}_q^n$ and then look up a received vector in the table in order to decode it. This is impractical if $q^n$ is very large.

**Syndrome decoding for** $[n, k, d]$ **linear code** $C$

We can devise an algorithm using a table with $q^{n-k}$ rather than $q^n$ entries where one can find the nearest codeword by looking up one of those $q^{n-k}$ entries.

The code $C$ is abelian subgroup of the additive group $\mathbb{F}_q^n$.

If $\mathbf{x} \in \mathbb{F}_q^n$, then $\mathbf{x} + C$ is a coset of $C$.

The cosets of $C$ form a partition of $\mathbb{F}_q^n$ into $q^{n-k}$ sets, each of size $q^n$.

Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$ belong to the same coset if and only if $\mathbf{y} - \mathbf{x} \in C$.

The weight of a coset is the smallest weight of a vector in the coset.
A coset leader is the vector in the coset of smallest weight. The zero vector is the unique coset leader of the code \( C \).

In general every coset of weight at most \( t = \lfloor \frac{d-1}{2} \rfloor \) has a unique coset leader.

**Definition 1.10.1.** Let \( H \) be the parity check matrix for \( C \). The syndrome of a vector \( x \) in \( \mathbb{F}_q^n \) with respect to the parity check matrix \( H \) is the vector in \( \mathbb{F}_q^{n-k} \) defined by
\[
\text{syn}(x) = Hx^\top.
\]

The code \( C \) consists of all vectors whose syndrome equal 0. As rank \( H = n - k \) every vector in \( \mathbb{F}_q^{n-k} \) is a syndrome.

**Theorem 1.10.4.** Two vectors belong to the same coset if and only if they have the same syndrome.

**Proof.** If \( x_1, x_2 \in \mathbb{F}_q^n \) are in the same coset of \( C \), then \( x_1 - x_2 = c \in C \implies x_1 = x_2 + c \). Therefore \( \text{syn}(x_1) = H(x_2 + c)^\top = Hx_2^\top + Hc^\top = Hx_2^\top = \text{syn}(x_2) \), then \( x_1, x_2 \) have the same syndrome and then lie on the same coset of \( C \).

If \( \text{syn}(x_1) = \text{syn}(x_2) \implies H(x_1 - x_2)^\top = 0 \implies x_2 - x_1 \in C \implies x_2 \in x_1 + C \). So \( x_1, x_2 \) lie on the same coset of \( C \).

There is one to one correspondence between cosets of \( C \) and syndromes. We denote by \( C_s \) the coset of \( C \) consisting of all vectors in \( \mathbb{F}_q^n \) with syndrome \( s \). So there is a one to one correspondence between the \( q^{n-k} \) cosets of \( C \) in \( \mathbb{F}_q^n \) and the \( q^{n-k} \) possible values of the syndromes.

The trivial coset \( C \) corresponding to the syndrome 0.

Nearest codeword decoding can thus be performed by the following steps:

1. Let \( y \) be a received vector, we seek an error vector \( e \) of smallest weight such that \( c = y - e \in C \). We find the syndrome of (the coset of) the received vector \( y \in \mathbb{F}_q^n \).

That is we compute \( s = \text{syn}(y) = Hy^\top \).
Finding a coset leader $e$ in the coset of the received vector $y$. Find a minimum weight vector $e \in \mathbb{F}_q^n$ such that

$$ s = \text{syn}(y) = H(c + e)^\top = Hc^\top + He^\top = He^\top. $$

Create a table pairing the syndrome with the coset leader, $y$ is decoded as the code word $y - e$. The table is used to look up the syndrome and find the coset leader.

**How do we construct a table of syndromes in step (1).**

Let $C$ be $t$ correcting code of $[n, k, d]$, with parity check matrix $H$, we construct the syndromes as follows.

1. The coset of weight 0 has coset leader 0.

2. Consider the $n$ cosets of weight 1.

3. Choose an $n$-tuple with a 1 in position $i$ and 0’s elsewhere, the coset leader is the $n$-tuple and the associated syndrome is column $i$ of $H$.

4. For the $\binom{n}{2}$ cosets of weight 2, choose an $n$-tuple with two 1’s in positions $i$ and $j$, with $i < j$ and the rest 0’s, the coset leader is the $n$-tuple and the associated syndrome is the sum of columns $i$ and $j$ of $H$.

5. continue in this manner through the cosets of weight $t$.

6. We could choose to stop. if we do we can decode any received vector with $t$ or fewer errors.

Remark 1.10.2. To find the syndrome $s = He^\top = \text{syn}(y)$ is equivalent to finding a smallest set of columns in $H$ whose linear span contains the vector $s$.

The syndrome decoding for binary Hamming codes $[2^r - 1, 2^r - 1 - r, 3]$ takes the form.

(i) After receiving a vector $y$, compute its syndrome $s$ using the parity check matrix $H_r$ of the Hamming code $\mathcal{H}_r$.

(ii) If $s = 0$, then $y$ is in the code and $y$ is decoded as $y$; otherwise, $s$ is the binary
numeral for some positive integer \( i \) and \( y \) is decoded as the codeword obtained from \( y \) by adding 1 to its \( i \)th bit.

**Example 1.10.2.** Construct the parity check matrix of the binary Hamming code \( \mathcal{H}_4 \) of length 15 where the columns are the binary numbers 1, 2, \( \cdots \), 15 in that order. Using this parity check matrix decode the following vectors, and then check that your decoded vectors are actually codewords.

(a) \( y_1 = 001000001100100 \)

(b) \( y_2 = 101001110101100 \)

(c) \( y_3 = 000100100011000 \).

**Solution:**

\[
H = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1
\end{pmatrix}
\]

(a) Let \( y_1 = 001000001100100 \) Then

\[
H y_1^\top = \begin{pmatrix}
1 \\
1 \\
0 \\
1
\end{pmatrix} = \text{col}_{13}.
\]

So \( e_1 = 00000000000100 \)

\[
\therefore c_1 = y_1 - e_1 = 00100000110000.
\]

(b)

\[
H y_2^\top = \begin{pmatrix}
0 \\
0 \\
1 \\
0
\end{pmatrix} = \text{col}_2
\]

47
So \( e_2 = 010000000000000 \)

\[ \therefore c_2 = y_2 - e_2 = 111001110101100. \]

**Example 1.10.3.** Let \( C \) be a linear \([5, 2, 3]\) code over \( \mathbb{F}_2 \) with generator matrix

\[
G = \begin{pmatrix}
1 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 1
\end{pmatrix}.
\]

The cosets of the code \( C \) are shown in the following table.

<table>
<thead>
<tr>
<th></th>
<th>10110</th>
<th>01011</th>
<th>11101</th>
</tr>
</thead>
<tbody>
<tr>
<td>00000</td>
<td>10110</td>
<td>01011</td>
<td>11101</td>
</tr>
<tr>
<td>00001</td>
<td>10111</td>
<td>01010</td>
<td>11100</td>
</tr>
<tr>
<td>00010</td>
<td>10100</td>
<td>01001</td>
<td>11111</td>
</tr>
<tr>
<td>00100</td>
<td>100010</td>
<td>01111</td>
<td>11001</td>
</tr>
<tr>
<td>01000</td>
<td>11110</td>
<td>00011</td>
<td>10101</td>
</tr>
<tr>
<td>10000</td>
<td>00110</td>
<td>11011</td>
<td>01101</td>
</tr>
<tr>
<td>00101</td>
<td>10011</td>
<td>01110</td>
<td>11000</td>
</tr>
<tr>
<td>10001</td>
<td>00111</td>
<td>11010</td>
<td>01100</td>
</tr>
<tr>
<td>00101</td>
<td>10011</td>
<td>01110</td>
<td>11000</td>
</tr>
<tr>
<td>10001</td>
<td>00111</td>
<td>11010</td>
<td>01100</td>
</tr>
</tbody>
</table>

Each row in the table is a coset of \( C \) and the first vector in each row is the coset leader of minimum weight. The last two rows could start with any of the words 00101, 11000, 10001, or 01100. Suppose that the received word is \( y = 01111 \). This word appears in the fourth row and the third column. The coset leader of the fourth row is 00100, and the decoded codeword is 01011, which is the first entry in the third column.

We can use the syndrome decoding to decode the receive word \( y \) by using the parity check matrix

\[
H = \begin{pmatrix}
1 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1
\end{pmatrix}
\]

\[
H y^\top = H (01111)^\top = (100)^\top = \text{col}_3.
\]

\( e = (00100) \) so \( c = y - e = (01111) - (00100) = (01011). \)
1.11 Sphere packing bound, covering radius and perfect codes

Definition 1.11.1. Given $n$ and $d$, the maximum number of codewords in a code over $\mathbb{F}_q$ of length $n$ and minimum distance $d$ is denoted by $A_q(n,d)$. For binary field it is denoted by $A_2(n,d)$ or $A(n,d)$.

Definition 1.11.2. For linear codes the maximum number of codewords is denoted by $B_q(n,d)$ and $B(n,d)$ for binary case.

Theorem 1.11.1. (Sphere packing bound)

$$B_q(n,d) \leq A_q(n,d) \leq q^n \sum_{i=0}^{t} \binom{n}{i} (q-1)^i$$

where $t = \lfloor \frac{d-1}{2} \rfloor$.

Proof. Let $C$ be a code over $\mathbb{F}_q$ of length $n$ and minimum distance $d$ (possibly nonlinear). Suppose that $C$ contains $M$ codewords. By previous theorem 1.10.1, the spheres of radius $t$ about distinct codewords are disjoint.

Since the volume of any sphere contains $\alpha = \sum_{i=0}^{t} \binom{n}{i} (q-1)^i$ total vectors, and the spheres are disjoint $M\alpha$ cannot exceed the number $q^n$ of vectors in $\mathbb{F}_q^n$, then

$$M\alpha \leq q^n \implies M \leq \frac{q^n}{\alpha} = \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i} (q-1)^i}$$

Definition 1.11.3. If

$$M = \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i} (q-1)^i}$$

then the code is called perfect code. That is $\mathbb{F}_q^n$ is filled by disjoint spheres of radius $t$. 

49
Example 1.11.1. Consider the Hamming code $H_{q,r}$ over $\mathbb{F}_q$ with parameters $[n = \frac{q^r - 1}{q - 1}, k = n - r, d = 3]$, then $t = 1$. Since $n(q - 1) = q^r - 1$, then $q^r = n(q - 1) + 1 = \frac{n}{1 + n(q - 1)} = \frac{q^n}{q^r} = q^n - 1 = q^k$ but $M = q^k$ so $H_{q,r}$ is a perfect code.

Example 1.11.2. Prove that the $[23, 12, 7]$ binary and the $[11, 6, 5]$ ternary Golay codes are perfect.

As $n = 23$, $k = 12$, $t = \left\lfloor \frac{7-1}{2} \right\rfloor = 3$. Then $\sum_{i=0}^{1} \binom{n}{i} (q-1)^i = 2\cdot 12 = 2^k$ so $[23, 12, 7]$ is a perfect code.

Definition 1.11.4. The covering radius $\rho = \rho(C)$ is the smallest integer $s$ such that $\mathbb{F}_q^n$ is the union of the spheres of radius $s$ centered at the codewords of $C$.

$\rho(C) = \max_{x \in \mathbb{F}_q^n} \min_{c \in C} d(x, c)$.

$t \leq \rho(C)$, and $t = \rho(C)$ if and only if $C$ is a perfect code. So the code is perfect if the packing radius equals the covering radius.

If $C$ is a code with packing radius $t$ and covering radius $t + 1$, $C$ is called quasi-perfect.

There is no general classification of quasi-perfect.
Chapter 2

Bounds on the size of codes

2.1 \( A_q(n, d) \) and \( B_q(n, d) \)

An \((n, M, d)\) code \(C\) over \(\mathbb{F}_q\) is a code of length \(n\) with \(M\) codewords whose minimum distance is \(d\).

If \(C\) is linear it is an \([n, k, d]\) code, where \(k = \log_q M\).

Definition 2.1.1. A code of length \(n\) over \(\mathbb{F}_q\) and minimum distance at least \(d\) will called optimal if it has \(A_q(n, d)\) codewords or \((B_q(n, d)\) codewords in the case that \(C\) is linear).

Theorem 2.1.1. \(B_q(n, d) \leq A_q(n, d)\) and \(B_q(n, d)\) is a nonnegative integer of \(q\).

\(B_q(n, d)\) is a lower bound for \(A_q(n, d)\) and \(A_n(n, d)\) is an upper bound for \(B_q(n, d)\). The sphere packing bound is an upper bound on \(A_q(n, d)\) and hence on \(B_q(n, d)\).

Theorem 2.1.2. Let \(d > 1\). Then

(i) \(A_q(n, d) \leq A_q(n - 1, d - 1)\) and \(B_q(n, d) \leq B_q(n - 1, d - 1)\) and

(ii) If \(d\) is even, \(A_2(n, d) = A_2(n - 1, d - 1)\) and \(B_2(n, d) = B_2(n - 1, d - 1)\), Furthermore

(iii) If \(d\) is even and \(M = A_2(n, d)\) then there exists a binary \((n, M, d)\) code such that all codewords have even weight and the distance between all pairs of codewords is also even.
Proof. (i) Let \( C \) be a code (linear or non linear) with \( M \) codewords and minimum distance \( d \). Puncturing on any coordinate gives a code \( C^* \), also with \( M \) codewords; otherwise if \( C^* \) has fewer codewords, there would exist two codewords of \( C \) which differ in one position implying \( d = 1 \). Contradiction because \( d > 1 \). Then \( A_q(n, d) \leq A_q(n - 1, d - 1) \) and \( B_q(n, d) \leq B_q(n - 1, d - 1) \).

(ii) To complete (ii) we only need to show that \( A_2(n, d) \geq A_2(n - 1, d - 1) \) or \( B_2(n, d) \geq B_2(n - 1, d - 1) \) where \( C \) is linear.

Let \( C \) be a binary code with \( M \) codewords, length \( n - 1 \), minimum distance \( d - 1 \). Extend \( C \) by adding an overall parity check to obtain a code \( \hat{C} \) of length \( n \) and minimum distance \( d \), since \( d - 1 \) is odd, then \( d \) is even. Because \( \hat{C} \) has \( M \) codewords, \( A_2(n, d) \geq A_2(n - 1, d - 1) \) or \( B_2(n, d) \geq B_2(n - 1, d - 1) \) then \( B_2(n, d) = B_2(n - 1, d - 1) \) and \( A_2(n, d) = A_2(n - 1, d - 1) \).

(iii) If \( C \) is a binary \((n, M, d)\) code with \( d \) even, the punctured code \( C^* \) is an \((n - 1, M, d - 1)\) code, extending \( C^* \) produces an \((n, M, d)\) code \( \hat{C} \), since \( d - 1 \) is odd.

Further this code has only even weight codewords, since \( d(x, y) = wt(x + y) = wt(x) + wt(y) - 2wt(x \cap y) \), the distance between codewords of even weight is even.

Example 2.1.1. Let \( n = 7, d = 4, t = \lfloor \frac{d - 1}{2} \rfloor \implies A_2(7, 4) \leq 16 \).

If \( n = 6, d = 3, t = 1 \) then the Sphere Packing bound yields \( \frac{64}{7} \) implying that \( A_2(6, 3) \leq 9 \).

But by theorem 2.1.2(ii) \( A_2(n, d) \leq A_2(n - 1, d - 1) \) so 9 is an upper bound for both \( A_2(7, 4), A_2(6, 3) \).

Theorem 2.1.3. \( A_2(n, 2) = B_2(n, 2) = 2^{n - 1} \).

Proof. By theorem 2.1.2(ii) \( A_2(n, 2) = A_2(n - 1, 1) \), but \( A(n - 1, 1) \leq 2^{n - 1} \), and the entire space \( \mathbb{F}_2^{n - 1} \) is a code of length \( n - 1 \) and minimum distance 1, implying \( A_2(n - 1, 1) = 2^{n - 1} \), as \( \mathbb{F}_2^{n - 1} \) is linear \( 2^{n - 1} = B_2(n - 1, 1) = B_2(n, 2) \).
Proof. The linear code of size $q$ consisting of all multiple of all one vector of length $n$ (that is the repetition code) has minimum distance $n$. So $q \leq B_q(n,n) \leq A_q(n,n)$. If $A_q(n,n) > q$, there exists a code with more than $q$ codewords and minimum distance $n$. Hence at least two of the codewords agree on some coordinate, but then these two codewords are less than distance $n$ apart, a contradiction. So $A_q(n,n) = B_q(n,n) = q$. □

Theorem 2.1.5. $A_q(n,d) \leq qA_q(n-1,d)$ and $B_q(n,d) \leq qB_q(n-1,d)$.

Proof. Let $C$ be a (possibly nonlinear) code over $\mathbb{F}_q$ of length $n$ and minimum distance at least $d$ with $M = A_q(n,d)$ codewords. Let $C(\alpha)$ be the subcode of $C$ in which every codeword has $\alpha$ in coordinate $n$. Then, for some $\alpha$, $C(\alpha)$ contains at least $M_q$ codewords. Puncturing this code on coordinate $n$ produces a code of length $n-1$ and minimum distance $d$. Therefore $\frac{M}{q} \leq A_q(n-1,d)$ giving $A_q(n,d) = M \leq qA_q(n-1,d)$. □

Theorem 2.1.6. Let $C$ be either a code over $\mathbb{F}_q$ with $A_q(n,d)$ codewords or a linear code over $\mathbb{F}_q$ with $B_q(n,d)$ codewords. Then $C$ has covering radius $d-1$ or less.

Proof. $t = \lfloor \frac{d-1}{2} \rfloor \leq \rho(C)$ where $\rho(C)$ is the covering radius which is greater than $t$.

$A_q(n,d) \leq A_q(n-1,d-1) \leq \frac{q^n}{\alpha} \text{ where } \alpha = \sum_{i=0}^{t} \binom{n}{i} (q-1)^i \implies \alpha A_q(n-1,d-1) \leq q^n \implies \rho(C) = d-1$ or less. □

2.2 Singleton upper bound and MDS codes.

Theorem 2.2.1. Singleton bound. for $d \leq n$, $A_q(n,d) \leq q^{n-d+1}$. Furthermore if an $[n,k,d]$ linear code over $\mathbb{F}_q$ exists then $k \leq n - d + 1$.

Proof. When $d = n$, then $A_q(n,n) = q$. Assume that $d \leq n$, by theorem 2.1.5 $A_q(n,d) \leq qA_q(n-1,d)$. Inductively we have that $A_q(n,d) \leq q^{n-d}A_q(d,d)$. Since $A_q(d,d) = d$ then $A_q(n,d) \leq q^{n-d+1}$. □

Example 2.2.1. For the code $[6,3,4]$ over $\mathbb{F}_4$, $k = 3 = 6 - 4 + 1 = n - d + 1$ and Singleton bound is met so $A_4(6,4) = 4^3$. 53
Definition 2.2.1. A code for which equality holds in the singleton bound is called maximum distance separable (MDS). That is if \( k = n - d + 1 \) holds then a code is called (MDS).

Remark 2.2.1. No code of length \( n \) and minimum distance \( d \) has more codewords than (MDS) with \( n, d \).

Theorem 2.2.2. Let \( C \) be an \([n, k] \) code over \( \mathbb{F}_q \) with \( k \geq 1 \). Then the following are equivalent:

(i) \( C \) is MDS.

(ii) Every set of \( k \) coordinates is an information set for \( C \).

(iii) \( C^\perp \) is MDS.

(iv) Every set of \( n - k \) coordinates is an information set for \( C^\perp \).

Proof. (i) \( \Rightarrow \) (ii) as \([n, k] \) is MDS if and only if \( k = n - d + 1 \), so any set \( n - d + 1 \) contains an information set if and only if every \( k \) columns of the generator matrix are linearly independent if and only if every \( n - k \) columns of the parity check matrix are linearly independent.

Similarly the last two items are equivalent, (ii) \( \Rightarrow \) (iv).

if \( \mathcal{I} \) and \( \mathcal{R} \) are the information and the redundancy positions for the code \( C \) then \( \mathcal{R} \) and \( \mathcal{I} \) are the information and redundancy positions for \( C^\perp \). So every set of \( n - k \) coordinates is an information set for \( C^\perp \).

(ii) \( \Rightarrow \) (iii). Since any \((n - k) \times (n - k)\) sub-matrix of \( H \) is invertible and all nontrivial linear combinations of the rows of \( H \) have at most \( n - k - 1 \) zero coordinates, therefore all non trivial linear combinations of rows of \( H \) have weight at least \( n - (n - k - 1) = k + 1 \). Since \( d \) for \( C \) is equal \( d = n - k + 1 \), for \( C^\perp \), \( d = n - (n - k) + 1 = n - \dim (C^\perp) + 1 = k + 1 \). So \( C^\perp \) is MDS.

\( \square \)

Definition 2.2.2. We say that \( C \) is a trivial MDS code over \( \mathbb{F}_q \) if and only if \( C = \mathbb{F}_q^n \) or \( C \) is monomially equivalent to the code generated by \( 1 \) or its dual.
Theorem 2.2.3. Let $C$ be an $[n,k,d]$ binary code.

(i) If $C$ is MDS, then $C$ is trivial.

(ii) If $3 \leq d$ and $5 \leq k$, then $k \leq n - d - 1$.

Proof. (i) Let $C$ be MDS, then $B_2(n,d) = 2^{n-d+1}$ but $B_2(n,d) = 2^n$ or $2^1$ or $2^{n-1}$.
Then $n - d + 1 = n \implies d = 1$ or $n - d + 1 = 1 \implies d = n$ or $n - d + 1 = n - 1 \implies d = 2$. The $C = \mathbb{F}_2^n$ or $C = [n,1,n]$ is repetition code. So if $C$ is MDS then $[n,n,1]$, $[n,1,n]$, $[n,n-1,2]$ are binary trivial codes.

(ii) \hfill $\Box$

2.3 Lexicodes

The algorithm for constructing lexicodes of length $n$ and minimum distance $d$ proceeds as follows:

(I) Order all $n-$ tuples in lexicographic order

\begin{align*}
0 & \ 0 \ \cdots \ 0 \ 0 \ 0 \\
0 & \ 0 \ \cdots \ 0 \ 0 \ 1 \\
0 & \ 0 \ \cdots \ 0 \ 1 \ 0 \\
0 & \ 0 \ \cdots \ 0 \ 1 \ 1 \\
0 & \ 0 \ \cdots \ 1 \ 0 \ 0 \\
\vdots
\end{align*}

(ii) Construct the class $L$ of vectors of length $n$ as follows:

(a) Put the zero vector $0$ in $L$.

(b) Look for the first vector $x$ of weight $d$ in the lexicographic ordering. Put $x$ in $L$. 

55
(c) Look for the next vector in the lexicographic ordering whose distance from each vector in \( L \) is \( d \) or more and add this to \( L \).

(d) Repeat (c) until there are no more vectors in the lexicographic list to look at.

The set \( L \) is a linear code, called the lexicode of length \( n \) and minimum distance \( d \).

Note: If \( u \) and \( v \) are binary vectors of length \( n \), we say \( u < v \) provided that \( u \) comes before \( v \) in the lexicographic order.

**Lemma 2.3.1.** If \( u, v \) and \( w \) are binary vectors of length \( n \) with \( u < v + w \) and \( v < u + w \), then \( u + v < w \).

**Proof.** If \( u < v + w \) then \( u \) and \( v + w \) have the same leading entries after which \( u \) has a 0 and \( v + w \) has a 1. We can represent this as follows:

\[
\begin{align*}
    u &= a_0 \cdots \\
    v + w &= a_1 \cdots 
\end{align*}
\]

Similarly as \( v < u + w \), we have

\[
\begin{align*}
    v &= b_0 \cdots \\
    u + w &= b_1 \cdots,
\end{align*}
\]

However, we do not know that the length \( i \) of \( a \) and the length \( j \) of \( b \) are the same. Assume they are different and by symmetry, that \( i < j \). Then we have

\[
\begin{align*}
    v &= b' x \cdots \\
    u + w &= b' x \cdots,
\end{align*}
\]

where \( b' \) is the first \( i \) entries of \( b \). Computing \( w \) in two ways, we obtain

\[
\begin{align*}
    w &= u + (u + w) = (a + b') x \cdots, \\
    w &= v + (v + w) = (a + b')(1 + x) \cdots,
\end{align*}
\]

a contradiction. So \( a \) and \( b \) are the same length, giving

\[
\begin{align*}
    u + v &= (a + b)0 \cdots, \\
    w &= (a + b)1 \cdots,
\end{align*}
\]

showing \( u + v < w \). \( \square \)
Theorem 2.3.2. Label the vectors in the lexicode in the order in which they are generated so that $c_0$ is the zero vector.

(i) $L$ is a linear code and the vectors $c_{2^i}$ are a basis of $L$.

(ii) After $c_{2^i}$ is chosen, the next $2^i - 1$ vectors generated are $c_1 + c_{2^i}$, $c_2 + c_{2^i}$, $\ldots$, $c_{2^i - 1} + c_{2^i}$.

(iii) Let $L_i = \{c_0, c_1, \ldots, c_{2^i - 1}\}$. Then $L_i$ is an $[n, i, d]$ linear code.

Proof. If we prove (ii), we have (i) and (iii). Then proof of (ii) is by induction on $i$.

Clearly it is true for $i = 1$. Assume the first $2^i$ vectors generated are as claimed. The $L_i$ is linear with basis $\{c_1, c_2, \ldots, c_{2^i - 1}\}$. We show that $L_{i+1}$ is linear with next $2^i$ vectors generated in order by adding $c_{2^i}$ to each of the previously chosen vectors in $L_i$. If not, there is an $r < 2^i$ such that $c_{2^i + r} \neq c_{2^i} + c_r$. Choose $r$ to be the smallest such value. As $d(c_{2^i} + c_r, c_{2^i} + c_j) = d(c_r, c_j) \geq d$ for $j < r$, $c_{2^i} + c_r$ was a possible vector to be chosen. Since it was not, it must have come too late in the lexicographic order, so $c_{2^i + r} < c_{2^i} + c_r$.

As $d(c_r + c_{2^i + r}, c_j) = d(c_{2^i + r}, c_r + c_j) \geq d$ for $j < 2^i$ by linearity of $L_i$, $c_r + c_{2^i + r}$ could have been chosen to be in the code instead of $c_{2^i}$ (which it cannot equal). So it must be that

$c_{2^i} < c_r + c_{2^i + r}$. (2)

If $j < r$, then $c_{2^i + j} = c_{2^i} + c_j$ by the assumption on $r$. Hence $d(c_{2^i + r} + c_{2^i}, c_j) = d(c_{2^i + r}, c_{2^i + j}) \geq d$ for $i < r$. So $c_{2^i + r} + c_{2^i}$ could have chosen to be a codeword instead of $c_r$. The fact that it was not implies that

$c_r < c_{2^i + r} + c_{2^i}$. (3)

But then (2) and (3) with previous lemma 2.3.1 imply $c_{2^i} + c_r < c_{2^i + r}$ contradicting (1).

Note: The codes $L_i$ satisfy the inclusions $L_1 < L_2 < \cdots < L_k = L$, where $k$ is the dimension of $L$. 

57
Chapter 3

Finite fields

3.1 Introduction to finite fields

Definition 3.1.1. A Field is a set $\mathbb{F}$ together with two operations $+$ and $\cdot$, which satisfy the following axioms:

1. $\mathbb{F}$ is an abelian group under $+$ with additive identity called zero denoted by $0$.
2. $\mathbb{F}^* = \mathbb{F}/\{0\}$ is an abelian group under multiplication with multiplicative identity called one and denoted by $1$;
3. multiplication distributes over addition.

If $p$ is a prime we let $GF(p)$ (Galois field with $p$ elements) denote the integer ring modulo $p$.

$\mathbb{Z}_p = \{0, 1, 2, \cdots, p-1\}$, group under $+$ mod $p$, the order of $\mathbb{Z}_p$ is $p$.
$\mathbb{Z}^* = \{1, 2, \cdots, p-1\} = \mathbb{Z}_p/\{0\}$ is cyclic group under multiplication mod $p$ and order $\mathbb{Z}_p/\{0\} = p - 1$.
$\mathbb{Z}_p = \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} = GF(p)$.
p is called the characteristic of $\mathbb{F}_p$.

Fields of order $q = p^m$: 

58
Let $\mathbb{F}_p$ be a field, the set $\mathbb{F}[x]$ of all polynomials over $\mathbb{F}_p$, is a P.I.D (principal ideal domain).

**Definition 3.1.2.** A polynomial $f(x)$ is called irreducible in $\mathbb{F}_p[x]$ if $f$ cannot be written as the product of two non-constant polynomial in $\mathbb{F}_p[x]$.

Let $f(x)$ be irreducible polynomial over $\mathbb{F}_p$ of degree $m$, then $\langle f(x) \rangle$ is a maximal ideal in $\mathbb{F}_p[x]$ generated by $f(x)$.

$$\langle f(x) \rangle = \{q(x)f(x) : q(x) \in \mathbb{F}_p[x] \} \implies \mathbb{F}_p[x]/\langle f(x) \rangle \text{ is a field } = \{(f(x)) + g(x) : g(x) \in \mathbb{F}_p[x] \}.$$

Where $g(x) = q(x)f(x) + r(x)$ where $r(x) = 0$ or $\deg r(x) < \deg f(x)$

$$\mathbb{F}_p[x]/\langle f(x) \rangle = \{(f(x)) + r(x) : r(x) = 0 \text{ or } \deg r(x) < m \} = \{a_0 + a_1 x + \cdots + a_{m-1} x^{m-1} : a_i \in \mathbb{F}_p[x] \}.$$

We denote this field by $\mathbb{F}_q$ where $q = p^m$ or $\mathbb{F}_q = GF(p^m)$ the Galois field of order $q = p^m$ which is the extension of the prime field $\mathbb{F}_p$. The characteristic of $\mathbb{F}_{p^m}$ is $p$.

The field $\mathbb{F}_{p^m}$ is a vector space over $\mathbb{F}_p$ of dim $m$ and $\mathbb{F}_p$ is a prime subfield of $\mathbb{F}_q$.

**Theorem 3.1.1.** Let $\mathbb{F}_q$ be a finite field with $q$ elements. Then

(i) $q = p^m$ for some prime $p$,

(ii) $\mathbb{F}_q$ contains the subfield $\mathbb{F}_p$,

(iii) $\mathbb{F}_q$ is a vector space over $\mathbb{F}_p$ of dimension $m$,

(iv) $p \alpha = 0$ for all $\alpha \in \mathbb{F}_q$, and

(v) $\mathbb{F}_q$ is unique up to isomorphism.

### 3.2 Polynomials and the Euclidean algorithm

The polynomial $f(x) = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{F}_q$, where $a_i \in \mathbb{F}_q$ is called a polynomial of degree $n$ over $\mathbb{F}_q$. 

59
The coefficient of the highest degree term is called the leading coefficient.
The polynomial is called monic if its leading coefficient is 1.

Let \( f(x), g(x) \in \mathbb{F}_q[x] \) we say that \( f(x)|g(x) \) if there exists a polynomial \( h(x) \in \mathbb{F}_q[x] \) such that \( g(x) = f(x)h(x) \).

If \( g(x) \neq 0 \) then \( \gcd(f(x), g(x)) \) is the monic polynomial in \( \mathbb{F}_q[x] \) of largest degree dividing both \( f(x) \) and \( g(x) \).

If \( \gcd(f(x), g(x)) = 1 \), then \( f(x), g(x) \) are called relatively prime, and \( \deg \gcd(f(x), g(x)) = 0 \).

**Theorem 3.2.1.** Let \( f(x), g(x) \) be in \( \mathbb{F}_q[x] \) with \( g(x) \) non-zero.

(i) (Division algorithm) There exist unique polynomials \( h(x), r(x) \in \mathbb{F}_q[x] \) such that

\[
f(x) = g(x)h(x) + r(x), \quad \text{where} \quad \deg r(x) < \deg g(x) \text{ or } r(x) = 0.
\]

(ii) If \( f(x) = g(x)h(x) + r(x) \), then \( \gcd(f(x), g(x)) = \gcd(g(x), r(x)) \).

**Theorem 3.2.2.** (Euclidean algorithm)

Let \( f(x), g(x) \) be polynomials in \( \mathbb{F}_q[x] \) with \( g(x) \) nonzero

(i) Perform the following sequence of steps until \( r_n(x) = 0 \) for some \( n \):

\[
\begin{align*}
    f(x) &= g(x)h_1(x) + r_1(x), \quad \text{where} \quad \deg r_1(x) < \deg g(x), \\
g(x) &= r_1(x)h_2(x) + r_2(x), \quad \text{where} \quad \deg r_2(x) < \deg r_1(x), \\
r_1(x) &= r_2(x)h_3(x) + r_3(x), \quad \text{where} \quad \deg r_3(x) < \deg r_2(x), \\
    \vdots \\
r_{n-3}(x) &= r_{n-2}(x)h_{n-1}(x) + r_{n-1}(x), \quad \text{where} \quad \deg r_{n-1}(x) < \deg r_{n-2}(x), \\
r_{n-2} &= r_{n-1}(x)h_n(x) + r_n(x), \quad \text{where} \quad r_n(x) = 0.
\end{align*}
\]

Then \( \gcd(f(x), g(x)) = cr_{n-1}(x) \), where \( c \in \mathbb{F}_q[x] \) is chosen so that \( cr_{n-1}(x) \) is monic.

(ii) There exist polynomials \( a(x), b(x) \in \mathbb{F}_q[x] \) such that \( a(x)f(x) + b(x)g(x) = \gcd(f(x), g(x)) \).
By applying theorem 3.2.1(ii) repeatedly, we have that 
\[ cr_{n-1}(x) = \gcd(r_{n-2}(x), r_{n-3}(x)) = \gcd(r_{n-3}(x), r_{n-4}(x)) = \cdots = \gcd(f(x), g(x)). \]
Part (ii) is obtained by beginning with the next to last equation \( r_{n-3}(x) = r_{n-2}h_{n-1}(x) + r_{n-1}(x) \) and solving for \( r_{n-1}(x) \) in terms of \( r_{n-2}(x) \) and \( r_{n-3}(x) \) using the previous equations, solve \( r_{n-2}(x) \) and substitute into the equation for \( r_{n-1}(x) \) to obtain \( r_{n-1}(x) \) as a combination of \( r_{n-3}(x) \) and \( r_{n-4}(x) \). Continue up through the sequence until we obtain \( r_{n-1}(x) \) as a combination of \( f(x) \) and \( g(x) \).

**Example 3.2.1.** Let \( f(x) = x^4 + x^2 + x + 1 \), and \( g(x) = x^3 + 1 \) in \( \mathbb{F}_2[x] \). Compute \( \gcd(f(x), g(x)) \) and find \( a(x), b(x) \in \mathbb{F}_2[x] \) \( \ni a(x)f(x) + b(x)g(x) = \gcd(f(x), g(x)) \).

**Solution:** Divide \( x^4 + x^2 + x + 1 \) by \( x^3 + 1 \) we get

\[ x^4 + x^2 + x + 1 = x(x^3 + 1) + (x^2 + 1). \]

Now divide \( x^3 + 1 \) by the remainder \( x^2 + 1 \) we get

\[ x^3 + 1 = x(x^2 + 1) + x + 1. \]

Now divide \( x^2 + 1 \) by \( x + 1 \) we get

\[ x^2 + 1 = (x + 1)(x + 1). \]

\[ \gcd(x^4 + x^2 + x + 1, x^3 + 1) = x + 1. \]

To find \( a(x) \) and \( b(x) \), by reversing the above steps we get,

\[ \gcd(f(x), g(x)) = x + 1 = x^3 + 1 + x(x^2 + 1) \]

\[ = (x^3 + 1) + x[x^4 + x^2 + x + 1 + x(x^3 + 1)] \]

\[ = x(x^4 + x^2 + x + 1) + (x^3 + 1)(x + 2) \]

\[ = a(x)f(x) + b(x)g(x). \]

Where \( a(x) = x \) and \( b(x) = x^2 + 1 \).
Remark 3.2.1. \((x - \alpha)|f(x)\) if and only if \(\alpha\) is a root of \(f(x)\) \(\implies\) there exist \(k(x) \in \mathbb{F}_q[x] \ni f(x) = (x - \alpha)k(x)\).

3.3 Primitive elements

The set \(\mathbb{F}_q^* = \mathbb{F}_q/\{0\}\) is the set of non zero elements in \(\mathbb{F}_q\) and it is a multiplicative group.

Theorem 3.3.1.

(i) The group \(\mathbb{F}_q^*\) is cyclic of order \(q - 1\) under the multiplication of \(\mathbb{F}_q\).

(ii) If \(\gamma\) is a generator of this cyclic group, then \(\mathbb{F}_q = \{0, 1 = \gamma^0, \gamma, \gamma^2, \ldots, \gamma^{q-2}\}\) and \(\gamma^i = 1\) if and only if \((q - 1)|i\).

Definition 3.3.1. A generator \(\gamma \in \mathbb{F}_q^*\) is called a primitive element of \(\mathbb{F}_q\).

Theorem 3.3.2. Let \(\mathbb{F}_q\) be a finite field. For every \(a \in \mathbb{F}_q\), \(a^{|\mathbb{F}_q|} = a^q = a\).

Proof. \(0^{|\mathbb{F}_q|} = 0\), and by Lagrange theorem applied to the multiplicative group \(\mathbb{F}_q^*\), we have \(a^{|\mathbb{F}_q|} = a^{q-1} = 1 \implies a^q = a\). \(\square\)

If \(\gamma\) is a primitive element of \(\mathbb{F}_q\). Then \(\gamma^{q-1} = 1\), hence \((\gamma^i)^{q-1} = 1\) for \(0 \leq i \leq q - 1\), then any element of \(\mathbb{F}_q^*\) are roots of the polynomial \(x^{q-1} - 1 \in \mathbb{F}_q[x]\) and hence \(x^q - x\), and \(0\) is a root of \(x^q - x\), so all elements of \(\mathbb{F}_q\) are roots of \(x^q - x\).

Theorem 3.3.3. The elements of \(\mathbb{F}_q\) are precisely the roots of \(x^q - x\).

Definition 3.3.2. The field \(\mathbb{F}_q\) with \(q = p^m\) elements is the smallest field containing \(\mathbb{F}_q\) and all the roots of \(x^q - x\). Such a field is called a splitting field of the polynomial \(x^q - x\) over \(\mathbb{F}_q\) containing all the roots of the polynomial \(x^q - x\). In general splitting fields of a fixed polynomial over a given field are isomorphic.

Definition 3.3.3. Let \(\phi(n)\) be the number of integers \(i\) with \(1 \leq i \leq n\) such that the \(\gcd(i, n) = 1\). \(\phi\) is called the Euler \(\phi\)-function.
Definition 3.3.4. If \( G \) is any cyclic group of order \( n \) with generator \( g \), then the generators of \( G \) are the elements \( g^i \) where \( \gcd(i, n) = 1 \). So \( \phi(n) \) = the number of generators of \( G \).

Definition 3.3.5. If \( \alpha \in G \) where \( G \) is a group of order \( n \). Then the order of \( \alpha \), \( o(\alpha) \) is the smallest positive integer \( i \) such that \( \alpha^i = 1 \) and an element of \( G \) has order \( d \) if and only if \( d|n \). If \( g \) is a generator of \( G \) then \( g^i \) has order \( d = \frac{n}{\gcd(i,n)} \) and there are \( \phi(d) \) elements of order \( d \).

Theorem 3.3.4. Let \( \gamma \) be a primitive element of \( \mathbb{F}_q \).

(i) There are \( \phi(q-1) \) primitive elements in \( \mathbb{F}_q \); these are the elements \( \gamma^i \) where \( \gcd(i, q-1) = 1 \).

(ii) For any \( d \) where \( d|(q-1) \), there are \( \phi(d) \) elements in \( \mathbb{F}_q \) of order \( d \); these are the elements \( \gamma^{(q-1)i/d} \) where \( \gcd(i, d) = 1 \).

Definition 3.3.6. An element \( \xi \in \mathbb{F}_q \) is an nth root of unity provided \( \xi^n = 1 \), and is a primitive nth root of unity if in addition \( \xi^s \neq 1 \) for \( 0 < s < n \).

A primitive element \( \gamma \in \mathbb{F}_q \) is a primitive \( (q-1) \)st root of unity. By theorem 3.3.1 the field \( \mathbb{F}_q \) contains primitive nth root of unity if and only if \( n|(q-1) \), in which case \( \gamma^{(q-1)/n} \) is a primitive nth root of unity.

3.4 Constructing finite fields

Definition 3.4.1. An integral domain is a commutative ring \( \mathcal{R} \) with unity such that the product of two nonzero elements in the ring is also nonzero. In fact \( \mathbb{F}_q[x] \) is an integral domain.

Definition 3.4.2. A non constant polynomial \( f(x) \in \mathbb{F}_q[x] \) is irreducible over \( \mathbb{F}_q \) provided it dose not factor into a product of two polynomials in \( \mathbb{F}_q[x] \) of smaller degree.

Definition 3.4.3. The ring \( \mathbb{F}_q[x] \) is a unique factorization domain and every polynomial in \( \mathbb{F}_q[x] \) is factored into a product of irreducible polynomials in unique form.
Theorem 3.4.1. Let \( f(x) \) be a non constant polynomial. Then

\[ f(x) = p_1(x)^{a_1} p_2(x)^{a_2} \cdots p_k(x)^{a_k}, \]

where each \( p_i(x) \) is irreducible, the \( p_i(x) \)'s are unique up to scalar multiplication and the \( a_i \)'s are unique.

Definition 3.4.4. An ideal \( I \) in a commutative ring \( R \) is a nonempty subset of the ring that is closed under subtraction such that the product of an element in \( I \) with an element in \( R \) is always in \( I \).

Definition 3.4.5. The ideal \( I \) is a principal ideal in a ring \( R \) if there exists \( a \in R \), such that \( I = \{ ra | r \in R \} \), this is denoted by \((a)\). A principal ideal domain is an integral domain in which each ideal is principal. In fact \( \mathbb{F}_q[x] \) is a principal ideal domain.

Theorem 3.4.2. Let \( \mathbb{F} \) be a field. The ideal \((f(x))\) is maximal in \( \mathbb{F}_q[x] \) if and only if \( f(x) \) is irreducible polynomial.

Theorem 3.4.3. Every PID is UFD.

Theorem 3.4.4. If \( f(x) \) is irreducible of degree \( m \) in \( \mathbb{F}_p[x] \), then the residue class ring \( \mathbb{F}_p[x]/(f(x)) \) is a field contains \( q = p^m \) elements.

Every element in the field \( \mathbb{F}_p[x]/(f(x)) \) is a coset \( g(x) + (f(x)) \) where \( \deg g(x) \leq m - 1 \).

We can write each coset or each element in \( \mathbb{F}_q[x] = \mathbb{F}_p[x]/(f(x)) \) as a vector in \( \mathbb{F}_p^m \) with the correspondence,

\[ g_{m-1}x^{m-1} + g_{m-2}x^{m-2} + \cdots g_1x + g_0 + (f(x)) \leftrightarrow g_{m-1}g_{m-2} \cdots g_1g_0. \]

Example 3.4.1. We construct the field \( GF(2^4) \) as a ring of polynomials over \( \mathbb{F}_2 = \{0, 1\} \) modulo the polynomial \( f(x) = x^4 + x + 1 \). \( f(x) \) is irreducible polynomial over \( \mathbb{F}_2 \), so \( \mathbb{F}_{16} = \mathbb{F}_2[x]/(f(x)) \), the elements of \( \mathbb{F}_{16} \) are given the following table, where \( x = \alpha = 0\alpha^3 + 0\alpha^2 + 1\alpha + 0 \cdot 1. \)
<table>
<thead>
<tr>
<th>Cosets</th>
<th>elements</th>
<th>Vectors in $F^4_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0 + (f(x))$</td>
<td>0</td>
<td>0000</td>
</tr>
<tr>
<td>$1 + (f(x))$</td>
<td>$\alpha^0 = 1$</td>
<td>0001</td>
</tr>
<tr>
<td>$x + (f(x))$</td>
<td>$\alpha$</td>
<td>0010</td>
</tr>
<tr>
<td>$x^2 + (f(x))$</td>
<td>$\alpha^2$</td>
<td>0100</td>
</tr>
<tr>
<td>$x^3 + (f(x))$</td>
<td>$\alpha^3$</td>
<td>1000</td>
</tr>
<tr>
<td>$x + 1 + (f(x))$</td>
<td>$\alpha + 1 = \alpha^4$</td>
<td>1000</td>
</tr>
<tr>
<td>$x^2 + 1 + (f(x))$</td>
<td>$\alpha^2 + 1 = \alpha^8$</td>
<td>0101</td>
</tr>
<tr>
<td>$x^3 + 1 + (f(x))$</td>
<td>$\alpha^3 + 1 = \alpha^{14}$</td>
<td>1001</td>
</tr>
<tr>
<td>$x^2 + x + (f(x))$</td>
<td>$\alpha^2 + \alpha = \alpha^5$</td>
<td>0110</td>
</tr>
<tr>
<td>$x^3 + x + (f(x))$</td>
<td>$\alpha^3 + \alpha = \alpha^9$</td>
<td>1010</td>
</tr>
<tr>
<td>$x^3 + x^2 + (f(x))$</td>
<td>$\alpha^3 + \alpha^2 = \alpha^6$</td>
<td>1100</td>
</tr>
<tr>
<td>$x^3 + x + 1 + (f(x))$</td>
<td>$\alpha^3 + \alpha + 1 = \alpha^7$</td>
<td>1011</td>
</tr>
<tr>
<td>$x^3 + x^2 + 1 + (f(x))$</td>
<td>$\alpha^3 + \alpha^2 + 1 = \alpha^{13}$</td>
<td>1101</td>
</tr>
<tr>
<td>$x^2 + x + 1 + (f(x))$</td>
<td>$\alpha^2 + \alpha + 1 = \alpha^{10}$</td>
<td>0111</td>
</tr>
<tr>
<td>$x^3 + x^2 + x + (f(x))$</td>
<td>$\alpha^3 + \alpha^2 + \alpha = \alpha^{11}$</td>
<td>1110</td>
</tr>
<tr>
<td>$x^3 + x^2 + x + (f(x))$</td>
<td>$\alpha^3 + \alpha^2 + \alpha + 1 = \alpha^{12}$</td>
<td>1111</td>
</tr>
</tbody>
</table>

Where $\alpha^4 = \alpha + 1$, $\alpha^5 = \alpha \alpha^4 = \alpha (\alpha + 1) = \alpha^2 + \alpha$, $\alpha^9 = \alpha^5 \alpha^4 = (\alpha^2 + \alpha)(\alpha + 1) = \alpha^3 + \alpha^2 + \alpha^2 + \alpha = \alpha^3 + \alpha$.

Adding $x + (f(x))$ to $x^3 + x + 1 + (f(x))$ gives $x^3 + 1 + (f(x))$, which corresponds to adding 0010 to 1011 and obtaining 1001 $\in F^4_2$.

How we multiply? To multiply $g_1(x) + (f(x))$ times $g_2(x) + (f(x))$, first use the division algorithm to write,

$g_1(x) g_2(x) = f(x) h(x) + r(x)$ when $\deg r(x) \leq m - 1$ or $r(x) = 0$ ($g_1(x) + (f(x)) (g_2(x) + (f(x))) = r(x) + (f(x))$).

We can simplify if we replace $x$ by $\alpha$ and let $f(\alpha) = 0$, $g_1(\alpha) g_2(\alpha) = r(\alpha)$ and we extend our correspondence to

$$g_{m-1} g_{m-2} \cdots g_1 g_0 = g_{m-1} \alpha^{m-1} + g_{m-2} \alpha^{m-2} \cdots g_1 \alpha + g_0.$$
So to multiply in $\mathbb{F}_q$ we simply multiply polynomials in $\alpha$ in the ordinary way and use the equation $f(\alpha) = 0$ to reduce powers of $\alpha$ higher than $m - 1$ to polynomials in $\alpha$ of degree less than $m$.

The subset $\{0\alpha^{m-1} + 0\alpha^{m-2} + \cdots + 0\alpha + a_0 : a_0 \in \mathbb{F}_p\} = \{a_0 : a_0 \in \mathbb{F}_p\} = \mathbb{F}_p$.

We describe the construction of obtaining $\mathbb{F}_q$ from $\mathbb{F}_p$ by adjoining a root $\alpha$ of $f(x)$ to $\mathbb{F}_p$, this root is given by $\alpha = x + (f(x))$ in $\mathbb{F}_q = \mathbb{F}_p[x]/(f(x))$ therefore $g(x) + (f(x)) = g(\alpha)$ and $f(\alpha) = f(x + (f(x))) = f(x) + (f(x)) = 0 + (f(x))$.

**Definition 3.4.6.** An irreducible polynomial over $\mathbb{F}_p$ of degree $m$ is called primitive provided if it has a root that is primitive element of $\mathbb{F}_q = \mathbb{F}_p^m$.

**Theorem 3.4.5.** For any prime $p$ and any positive integer $m$, there exists a finite field, unique up to isomorphism, with $q = p^m$ elements.

**Remark 3.4.1.** In the construction of $\mathbb{F}_q$ by adjoining a root of an irreducible polynomial $f(x)$ to $\mathbb{F}_p$, the field $\mathbb{F}_p$ can be replace by any finite field $\mathbb{F}_r$, where $r$ is a power of $p$ and $f(x)$ be an irreducible polynomial of degree $m$ in $\mathbb{F}_r[x]$ for some positive integer $m$. The field constructed contains $\mathbb{F}_r$ as a subfield is of order $r^m$.

### 3.5 Subfields

Let $\mathbb{F}_q$ be a field with primitive element of order $q - 1 = p^m - 1$. If $\mathbb{F}_s$ is a subfield of $\mathbb{F}_q$ then $\mathbb{F}_s$ has a primitive element of order $s - 1$ where $(s - 1)|(q - 1)$ because the identity element 1 is the same for both $\mathbb{F}_q$ and $\mathbb{F}_s$. $\mathbb{F}_s$ has characteristic $p$ and $s = p^r$. So $(p^r - 1)|(p^m - 1)$.

**Lemma 3.5.1.** Let $a > 1$ be an integer. Then $(a^r - 1)|(a^m - 1)$ if and only if $r|m$.

**Proof.** If $r|m$, then $m = rh$ and $a^m - 1 = (a^r - 1)\sum_{i=0}^{h-1} a^{ir}$ because $\sum_{i=0}^{h-1} a^{ir} = \frac{a^{rh} - 1}{a^r - 1} \Rightarrow a^m - 1 = (a^r - 1)\frac{a^{rh} - 1}{a^r - 1}$.
Conversely, by division algorithm, let \( m = rh + n, \ 0 \leq u < r \), then

\[
a^m - 1 = \frac{a^{rh+u} - 1}{a^r - 1} = \frac{a^{rh} a^u - a^u}{a^r - 1} = \frac{a^r(a^{rh} - 1)}{a^r - 1} + \frac{a^u - 1}{a^r - 1}.
\]

Since \( r \mid rh \), \( \implies \frac{a^{rh} - 1}{a^r - 1} \) is integer, also

\[
a^{u-1} < 1 \text{ a contradiction, since } \frac{a^{u-1}}{a^r - 1} \text{ is integer, then } \frac{a^{u-1}}{a^r - 1} = 0 \implies u = 0 \implies r \mid m.
\]

**Remark 3.5.1.** \((x^{s-1} - 1)(x^{q-1} - 1)\) if and only if \((s - 1)(q - 1)\) thus \((x^s - x)(x^q - x)\) if and only if \((s - 1)(q - 1)\). So if \( s = p^r \), \( q = p^m \), \((x^s - x)(x^q - x)\) if an only if \( r \mid m \).

**Lemma 3.5.2.** Let \( s = p^r \), \( q = p^m \). Then \((x^s - x)(x^q - x)\) if an only if \( r \mid m \).

**Proof.** \((x^s - x)(x^q - x)\) if and only if \( x^{s-1} - 1 \mid x^{q-1} - 1 \) if and only if \( s - 1 \mid q - 1 \) if and only if \( p^r - 1 \mid p^m - 1 \) if and only if \( r \mid m \).

**Theorem 3.5.3.** When \( q = p^m \),

(i) \( F_q \) has a subfield of order \( s = p^r \) if and only if \( r \mid m \),

(ii) the elements of the subfield \( F_s \) are exactly the elements of \( F_q \) that are roots of \( x^s - x \), and

(iii) for each \( r \mid m \) there is only one subfield \( F_{p^r} \) of \( F_q \).

**Proof.** Let \( s = p^r \) and \( q = p^m \). Then by previous lemma 3.5.2 \((x^s - x)(x^q - x)\) if and only if \( r \mid m \) if and only if \( p^r - 1 \mid p^m - 1 \).

Since \( x^q - x \) siplits into linear factors over \( F_q = F_{p^m} \), so does \( x^s - x \) siplits into linear factors over \( F_s = F_{p^r} \). So \( F_q \) contains all the roots of \( x^r - x \) in \( F_s \) implies \( F_s \) is a subfield of \( F_q \) and \( F_s \) contains \( s = p^r \) elements which are the roots of \( x^s - x \).
3.6 Field automorphisms

Definition 3.6.1. An automorphism $\sigma$ of $\mathbb{F}_q$ is a bijection (isomorphism) $\sigma : \mathbb{F}_q \rightarrow \mathbb{F}_q$ such that $\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)$ and $\sigma(\alpha \beta) = \sigma(\alpha) \sigma(\beta)$ for all $\alpha, \beta \in \mathbb{F}_q$.

Definition 3.6.2. (Frobenius automorphism) is defined by $\sigma_p(\alpha) = \alpha^p$ for all $\alpha \in \mathbb{F}_q$. $\sigma_p(\alpha \beta) = \sigma_p(\alpha) \sigma_p(\beta)$ and $\sigma_p(\alpha + \beta) = \sigma_p(\alpha) + \sigma_p(\beta)$, ker($\sigma_p$) = $\{0\}$. So $\sigma_p$ is an automorphism of $\mathbb{F}_q$, called the Frobenius automorphism.

Similiarly $\sigma_{p^r}(\alpha) = \alpha^{p^r}$ is also a Frobenius automorphism.

Definition 3.6.3. The set of all automorphism of a field $\mathbb{F}_q$ onto itself is a group under composition of functions denoted by $\text{Gal}(\mathbb{F}_q)$ called the Galois group of $\mathbb{F}_q$.

Theorem 3.6.1.

(i) $\text{Gal}(\mathbb{F}_q)$ is cyclic of order $m$ and is generated by the Frobenius automorphism $\sigma_p$.

(ii) The prime subfield of $\mathbb{F}_q$ is precisely the set of elements in $\mathbb{F}_q$ such that $\sigma_p(\alpha) = \alpha$.

(iii) The subfield $\mathbb{F}_q$ of $\mathbb{F}_{q^t}$ is precisely the set of elements in $\mathbb{F}_{q^t}$ such that $\sigma_q(\alpha) = \alpha$.

Proof. (i) Let $\gamma$ be a primitive element of $\mathbb{F}_q = \mathbb{F}_{p^m}$. Thus $\gamma$ has order $q - 1 = p^m - 1$, and its minimal polynomial $m(x)$ has roots, $\gamma, \gamma^p, \gamma^{p^2}, \ldots, \gamma^{p^{m-1}}$.

Now let $f(x)$ be a polynomial over $\mathbb{F}_p$. Since an automorphism $\tau$ of $\mathbb{F}_{p^m}$ over $\mathbb{F}_p$ fixes the coefficients of $f(x)$, we see that $f(\alpha) = 0$ if and only if $f(\tau(\alpha)) = 0$. In otherwords, $\tau$ permutes the roots of $f(x)$ that lie in $\mathbb{F}_{p^m}$. In particular, $\tau \sigma$ must send the root $\gamma$ of $m(x)$ to another root, say $\tau(\gamma) = \gamma^{p^i}$. But since $\gamma$ is a primitive element of $\mathbb{F}_{p^m}$, $\tau$ is completely determend by its value on $\gamma$, and since $\sigma_p^i(\gamma) = \gamma^{p^i} = \tau(\gamma)$. We denote $\tau = \sigma_p^i$. Hence all automorphism of $\mathbb{F}_q = \mathbb{F}_{p^n}$ over $\mathbb{F}_p$ have the form $\sigma_p^i$ for some $i$.

(ii) Let $\mathbb{F} = \{\alpha \in \mathbb{F}_q : \sigma_p(\alpha) = \alpha\}$. Let $a, b \in \mathbb{F}$, then $\sigma_p(a - b) = \sigma_p(a) - \sigma_p(b) = a - b$. $\sigma_p(ab) = \sigma_p(a) \sigma_p(b)$. So $a - b \in \mathbb{F}, ab \in \mathbb{F}$.
Let \( F_q = \{ \alpha \in F_q : \sigma_q(\alpha) = \alpha \}. \) Let \( a', b' \in F_q \) then \( \sigma_q(a' - b') = \sigma_q(a') - \sigma_q(b') = a' - b'. \)

Also \( \sigma_q(a'b') = \sigma_q(a') \sigma_q(b') = a'b' \in F_q \)

\( \therefore F_q \) is a subfield of \( F_p^t. \)

Remark 3.6.1. \( Gal(F_q) \) is a cyclic group (denote automorphism of \( F_q = F_p^m \) with \( F_p \) fixed).

The generator of \( Gal(F_q) \) is the Frobenious automorphism.

Remark 3.6.2. We use \( \sigma_p \) to denote the Forbenius automorphism of any field of characteristic \( p. \) If \( E \) and \( F \) are fields of characteristic \( p \) with \( E \) an extension field of \( F, \) then the Forbenius automorphism of \( E \) when restricted to \( F \) is the Frobenius automorphism of \( F. \)

An element \( \alpha \in F \) is fixed by an automorphism \( \sigma \) of \( F \) provided \( \sigma(\alpha) = \alpha. \) Let \( r|m. \) Then \( \sigma_r^r \) generates a cyclic subgroup of \( Gal(F_q) \) of order \( m/r. \) The elements of \( F_q \) fixed by this subgroup are precisely the elements of the subfield of \( F_p^r. \) We let \( Gal(F_q : F_p^r) \) denote automorphisms of \( F_q \) which fix \( F_p^r. \)

Theorem 3.6.2. \( Gal(F_q : F_p^r) \) is the cyclic group generated by \( \sigma_p^r. \)

Proof. The Forbinous automorphism is a generator of the cyclic group \( Gal(F_q : F_p). \)

\( Gal(F_q) = Gal(F_q : F_p) = \{ \sigma^i : 0 \leq i \leq m - 1 \}. \) The group \( Gal(F_q : F_p^r) \) is a subgroup of \( Gal(F_q) \) because \( \sigma_r^i(\alpha) = \alpha^{p^i} = \sigma_p(\alpha) \in Gal(F_q : F_p^r) \implies \sigma_r^i \in Gal(F_q) \implies Gal(F_q : F_p^r) \subseteq Gal(F_q). \) The subgroups of \( Gal(F_q) \) are cyclic, so \( \sigma(\sigma_r^r) = \sigma(\sigma_p^r) = m/r. \) Since \( \sigma_r^m(\alpha) = \alpha^{p^m} = \alpha, \) so \( \sigma_r^m = \sigma_r^r \) is a generator of \( Gal(F_q : F_p^r) \implies |Gal(F_q : F_p^r)| = m/r. \)

### 3.7 Cyclotomic cosets and minimal polynomials

**Definition 3.7.1.** If \( E \) is the extension field \( F_q, \) then \( E \) is a vector space over \( F_q, \) and \( E = F_{q^t} \) for some positive \( t. \) Each element \( \alpha \in E \) is a root of the polynomial \( x^{q^t} - x. \) Thus there is a monic polynomial \( M_\alpha(x) \) in \( F_q[x] \) of smallest degree which has \( \alpha \) as a root, this polynomial is called the minimal polynomial of \( \alpha \) over \( F_q. \)
Remark. Let $M = \langle \alpha \rangle$ be an ideal of $R$. Assume that $\alpha$ is a root of a non-zero polynomial over $R$. Let $g(x) = \langle x \rangle \not\in R[x]$ be an element of $R[x]$ of smallest degree, that is there is only one monic polynomial in $R[x]$ of smallest degree which has $\alpha$ as a root.

\textbf{Proof.} (i) If $M = \langle \alpha \rangle$ is reducible then $M = u(x)v(x)$, with $\deg u(x) < \deg M$ and $\deg v(x) < \deg M$. Then $u(\alpha)v(\alpha) = 0$. So either $u(\alpha) = 0$ or $v(\alpha) = 0$. Contradiction, because $R[x]$ is an integral domain and $M = \langle \alpha \rangle$ is of minimal degree over $R[x]$ having $\alpha$ as a root.

(ii) Let $I = \{g(x) \in R[x] : g(\alpha) = 0\}$. $I$ is an ideal of $R[x]$ because if $g_1(x), g_2(x) \in I$ then $g_1(x) - g_2(x) = 0 \in I$ and if $g(x) \in I$ and $f(x) \in R[x] \implies g(x)f(x) \in I$ then $I$ is an ideal of $R[x]$. Also $I \neq 0$ because $\alpha$ is a root of a non-zero polynomial over $R$, where $\alpha \in R$, with minimal polynomial $M = \langle \alpha \rangle$. But $R[x]$ is P.I.D, hence there exist a generator $h(x)$ for $I$. $I = \langle h(x) \rangle$. Let $p_1(x)$ be the polynomial of minimal degree that generates $I$. If $p_1(x) = a_0 + a_1 x + \cdots + a_t x^t$ then $M = \langle \alpha \rangle = \frac{1}{a_t} (p_1(x))$ is a monic polynomial that generates $I$ having minimal degree. $I = \langle M = \langle \alpha \rangle \rangle$ and $M = \langle \alpha \rangle$ is of minimal degree. Let $g(x) \in I$, since $M = \langle \alpha \rangle \neq 0$, so by division algorithm there exists $q(x), r(x) \in R[x] \ni g(x) = q(x)M + r(x)$, where $r(x) = 0$ or $\deg r(x) < \deg M$, but $g(\alpha) = g(\alpha)M + r(\alpha) \implies r(\alpha) = 0$. If $r(\alpha) \neq 0$, then we get a contradiction, because $M = \langle \alpha \rangle$ is of minimal degree, hence $r(x) = 0$ and so $g(x) = q(x)M = \langle \alpha \rangle$, so $M = \langle \alpha \rangle$.

(iii) Assume that $\langle M = \langle \alpha \rangle \rangle = \langle M' = \langle \alpha \rangle \rangle$ where $M = \langle \alpha \rangle$ and $M' = \langle \alpha \rangle$ are monic irreducible polynomial. Then $M = M' = \langle \alpha \rangle$ since both are monic generated the same ideal, so $M = \langle \alpha \rangle$ is unique.

\[ \square \]
Theorem 3.7.2. Let $f(x)$ be a monic irreducible polynomial over $\mathbb{F}_q$ of degree $r$. Then

(i) all the roots of $f(x)$ lie in $\mathbb{F}_{q^r}$ and any field containing $\mathbb{F}_q$ along with one root of $f(x)$,

(ii) $f(x) = \prod_{i=1}^{r} (x - \alpha_i)$, where $\alpha_i \in \mathbb{F}_{q^r}$ for $1 \leq i \leq r$, and

(iii) $f(x)|x^{q^r} - x$.

Proof. (i) Let $\alpha$ be a root of $f(x)$ which we adjoin to $\mathbb{F}_q$ to form a field $\mathbb{E}_\alpha$ with $q^r$ elements. If $\beta$ is another root of $f(x)$, not in $\mathbb{E}_\alpha$ it is a root of some irreducible factor, over $\mathbb{E}_\alpha$ of $f(x)$. Adjoining $\beta$ to $\mathbb{E}_\alpha$ forms an extension field $\mathbb{E}$ of $\mathbb{E}_\alpha$. However, inside $\mathbb{E}$, there is a subfield $\mathbb{E}_\beta$ obtained by adjoining $\beta$ to $\mathbb{F}_q$. $\mathbb{E}_\beta$ must have $q^r$ elements because $f(x)$ is irreducible of degree $r$ over $\mathbb{F}_q$. Since $\mathbb{E}_\alpha$ and $\mathbb{E}_\beta$ are subfields of $\mathbb{E}$ of same size, then by theorem 3.5.3(iii) $\mathbb{E}_\alpha = \mathbb{E}_\beta$ proving that all roots of $f(x)$ lie in $\mathbb{F}_{q^r}$. Since any field containing $\mathbb{F}_q$ and one root of $f(x)$ contains $\mathbb{F}_{q^r}$, so part (i) follows.

(ii) If $\alpha_i \in \mathbb{F}_{q^r}$ is a root of $f(x)$ for $1 \leq i \leq r$ then $(x - \alpha_i)$ is a factor of $f(x)$, then $f(x) = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_r) = \prod_{i=1}^{r} (x - \alpha_i)$.

(iii) Every element of $\mathbb{F}_{q^r}$ is a root of the polynomial $x^{q^r} - x$, therefore $x - \alpha$ is a factor of $x^{q^r} - x \forall \alpha \in \mathbb{F}_{q^r} \implies x^{q^r} - x = \prod_{\alpha \in \mathbb{F}_{q^r}} (x - \alpha)$. Since $f(x) = \prod_{i=1}^{r} (x - \alpha_i)$, then $f(x)|x^{q^r} - x$.

\[ \square \]

Theorem 3.7.3. Let $\mathbb{F}_{q^t}$ be an extension field of $\mathbb{F}_q$ and let $\alpha$ be an element of $\mathbb{F}_{q^t}$ with minimal polynomial $M_\alpha(x)$ in $\mathbb{F}_q[x]$. The following hold:

(i) $M_\alpha(x)|x^{q^t} - x$.

(ii) $M_\alpha(x)$ has distinct roots all lying in $\mathbb{F}_{q^t}$.

(iii) The degree of $M_\alpha(x)$ divides $t$. 

71
(iv) \( x^q - x = \prod_{\alpha} M_\alpha(x) \) where \( \alpha \) runs through some subset of \( \mathbb{F}_q \) which enumerates the minimal polynomial of all elements of \( \mathbb{F}_q \) exactly once.

(v) \( x^q - x = \prod_{f} f(x) \), where \( f \) runs through all monic irreducible polynomials over \( \mathbb{F}_q \) whose degree divides \( t \).

**Proof.**

(i) If \( \alpha \in \mathbb{F}_q \) and if \( g(x) \) is any polynomial in \( \mathbb{F}_q[x] \) satisfying \( g(\alpha) = 0 \), then \( M_\alpha(x)|g(x) \). But \( g(x)|x^q - x \implies M_\alpha(x)|x^q - x \).

(ii) Since the roots of \( x^q - x \) are the \( q^t \) elements of \( \mathbb{F}_q \), \( x^q - x \) has distinct roots and so by part (i) and part (i) of theorem 3.7.2 implies \( M_\alpha(x) \) has distinct roots all lying in \( \mathbb{F}_q \).

(iii) If \( M_\alpha(x) \) has degree \( r \), adjoining \( \alpha \) to \( \mathbb{F}_q \) gives the subfield \( \mathbb{F}_{q^r} = \mathbb{F}_{p^{mr}} \) of \( \mathbb{F}_{q^t} = \mathbb{F}_{p^{mt}} \) then by theorem 3.5.3 \( mr|mt \) then \( r|t \).

(iv) There exist irreducible polynomial \( p_i(x) \) over \( \mathbb{F}_q \) such that \( x^q - x = \prod_{i=1}^{n} p_i(x) \). As \( x^q - x \) has distinct roots, the factors \( p_i(x) \) are distinct. By scaling them, we assume that each is monic as \( x^q - x \) is monic. So \( p_i(x) = M_\alpha(x) \) for any \( \alpha \in \mathbb{F}_q \), with \( p_i(\alpha) = 0 \implies x^q - x = \prod_{\alpha} M_\alpha(x) \).

(v) Follows from (iv) if we show that every monic irreducible polynomial over \( \mathbb{F}_q \) of degree \( r \) dividing \( t \) is a factor of \( x^q - x \). But \( f(x)|x^q - x \) by theorem 3.7.2 (iii). Since \( mr|mt \), \( x^q - x|x^q - x \implies x^q - x = \prod_{f} f(x) \).

\[ \square \]

**Definition 3.7.2.** Two elements of \( \mathbb{F}_q \) which have the same minimal polynomial in \( \mathbb{F}_q[x] \) are called conjugate over \( \mathbb{F}_q \).

**Remark 3.7.1.** All the roots of \( M_\alpha(x) \) are distinct and lie in \( \mathbb{F}_q \), so they are conjugates, we can find them by the following theorem.

**Theorem 3.7.4.** Let \( f(x) \) be a polynomial in \( \mathbb{F}_q[x] \) and let \( \alpha \) be a root of \( f(x) \) in some extension field \( \mathbb{F}_q \). Then:
(i) \( f(x^q) = f(x)^q \), and

(ii) \( \alpha^q \) is also a root of \( f(x) \) in \( \mathbb{F}_q \).

**Proof.** (i) Let \( f(x) = \sum_{i=0}^{n} a_i x^i \). Since \( q = p^m \), where \( p \) is the characteristic of \( \mathbb{F}_q \), then

\[
f(x^q) = \left( \sum_{i=0}^{n} a_i x^i \right)^q = \sum_{i=0}^{n} a_i^q x^{iq}, \text{ but } a_i^q = a_i \forall a_i \in \mathbb{F}_q \text{ because the elements in } \mathbb{F}_q \text{ are the roots of } x^q - x = 0 = f(x)^q = f(x^q).
\]

(ii) \( f(\alpha^q) = \sum_{i=0}^{n} a_i \alpha^{iq} = \sum_{i=0}^{n} a_i \alpha^i = f(\alpha) = 0 = f(\alpha)^q \implies \alpha^q \text{ is a root of } f(x) \text{ in } \mathbb{F}_q. \)

**Remark 3.7.2.** By theorem 3.7.4 \( \alpha, \alpha^q, \alpha^{q^2}, \cdots \) are all roots of \( M_{\alpha}(x) \). This sequence will stop after \( r \) terms, where \( \alpha^{q^r} = \alpha \).

If \( \gamma \) is a primitive element of \( \mathbb{F}_{q^t} \). Then \( \alpha = \gamma^s \) for some \( s \). Hence \( \alpha^{q^r} = \alpha \) if and only if \( \gamma^{sq^r} = \gamma^s \) if and only if \( \gamma^{sq^r-s} = 1 \). By theorem 3.3.1 \( \gamma^i = 1 \) if and only if \( q - 1 | i \). So \( q^i - 1 | sq^r - s \implies sq^r \equiv s \text{ (mod } q^t - 1) \).

**Definition 3.7.3.** The set of \( q^t \)-cyclotomic cosets of \( s \) modulo \( q^t - 1 \) is the set \( C_s = \{s, sq, \cdots, sq^{r-1}\} \text{ (mod } q^t - 1\text{), where } r \text{ is the smallest positive integer such that } sq^r \equiv s \text{ (mod } q^t - 1)\text{. the set } C_s \text{ partition the set } \{0, 1, 2, \cdots, q^t - 2\} \text{ of integers into disjoint sets.} \)

**Example 3.7.1.** The 2- cyclotomic cosets modulo 15 are \( C_0 = \{0\}, C_1 = \{1, 2, 4, 8\}, C_3 = \{3, 6, 12, 9\}, C_5 = \{5, 10\} \text{ and } C_7 = \{7, 14, 13, 11\}. \)

**Remark 3.7.3.** The roots of \( M_{\alpha}(x) = M_{\gamma^s}(x) \) include \( \{\gamma^i | i \in C_s\} \). In fact these are all the roots of \( m_{\alpha}(x) \). So if we know the size of \( C_s \), we know the degree of \( M_{\gamma^s}(x) \), as these are the same.

**Theorem 3.7.5.** If \( \gamma \) is a primitive element of \( \mathbb{F}_{q^t} \), then the minimal polynomial of \( \gamma^s = \alpha \) over \( \mathbb{F}_q \) is

\[
M_{\gamma^s}(x) = \prod_{i \in C_s} (x - \gamma^i).
\]
Proof. We will claim that,

\[ f(x) = \prod_{i \in C_s} (x - \gamma^i) = \sum_j f_j x^j \]

is a polynomial in \( \mathbb{F}_q[x] \) not in \( \mathbb{F}_q^t[x] \).

Let \( g(x) = f(x)^q \). Then \( g(x) = \prod_{i \in C_s} (x - \gamma^i)^q = \prod_{i \in C_s} (x^q - \gamma^{iq})( \text{Because if } q = p^r, (x - \gamma^i)^{p^r} = x^{p^r} - \gamma^{ip^r} = x^q - \gamma^{iq}) \).

As \( C_s \) is a \( q \)- cyclotomic coset, \( qi \) runs through \( C_s \) as \( i \) does. Then \( g(x) = f(x^q) = \sum_j f_j x^{jq}, \) but \( g(x) = f(x^q) = f(x)^q = (\sum_j f_j x^j)^q = \sum_j f_j^q x^{jq} \). Equating coefficients, we have \( f_j^q = f_j \) and hence by theorem 3.6.1 (iii)(The subfield \( \mathbb{F}_q \) of \( \mathbb{F}_q^t \) is precisely the set of elements in \( \mathbb{F}_q^t \) such that \( \sigma_q(\alpha) = \alpha \), then \( f(x) \in \mathbb{F}_q[x] \). \qed

Example 3.7.2. We are constructed \( \mathbb{F}_{16} \) previously using the irreducible polynomial \( x^4 + x + 1 \) over \( \mathbb{F}_2 \). With \( \alpha \) as a root of this polynomial, we give the minimal polynomial over \( \mathbb{F}_2 \) of each element of \( \mathbb{F}_{16} \) and the associated \( 2- \) cyclotomic coset modulo 15 in the table below.

<table>
<thead>
<tr>
<th>Roots</th>
<th>Minimal polynomial</th>
<th>2-cyclotomic coset</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( x )</td>
<td>{0}</td>
</tr>
<tr>
<td>1</td>
<td>( x + 1 )</td>
<td>{1, 2, 4, 8}</td>
</tr>
<tr>
<td>( \alpha, \alpha^2, \alpha^4, \alpha^8 )</td>
<td>( x^4 + x + 1 )</td>
<td>{3, 6, 9, 12}</td>
</tr>
<tr>
<td>( \alpha^3, \alpha^6, \alpha^9, \alpha^{12} )</td>
<td>( x^4 + x^3 + x^2 + x + 1 )</td>
<td>{5, 10}</td>
</tr>
<tr>
<td>( \alpha^5, \alpha^{10} )</td>
<td>( x^2 + x + 1 )</td>
<td>{7, 11, 13, 14}</td>
</tr>
<tr>
<td>( \alpha^7, \alpha^{11}, \alpha^{13}, \alpha^{14} )</td>
<td>( x^4 + x^3 + 1 )</td>
<td></td>
</tr>
</tbody>
</table>

The factorization of \( x^{15} - 1 \) into irreducible polynomials in \( \mathbb{F}_2[x] \) is

\[ (x + 1)(x^4 + x + 1)(x^4 + x^3 + x^2 + x + 1)(x^2 + x + 1)(x^4 + x^3 + 1). \]
Chapter 4

Cyclic codes

Introduction

A linear code $C$ of length $n$ over $\mathbb{F}_q$ is cyclic provided that for each vector $c = c_0c_1 \cdots c_{n-2}c_{n-1}$ in $C$ the vector $c_{n-1}c_0c_1 \cdots c_{n-2}$ obtained from $c$ by the cyclic shift of coordinate $i \mapsto i + 1 \pmod n$, is also in $C$. So a cyclic code contains all $n$ cyclic shifts of any codeword.

We will represent the codewords in a polynomial form. If $c = c_0c_1 \cdots c_{n-2}c_{n-1} \in \mathbb{F}_q^n$ then
$$c(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1} \in \mathbb{F}_q[x]$$
of degree at most $n - 1$. We order the terms of our polynomial from smallest to largest degree.

If $c(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1}$, then $xc(x) = c_{n-1}x^nc_0x + c_1x^2 + \cdots + c_{n-2}x^{n-1}$ which would represent the codeword $c$ cyclically shifted one to right, if $x^n = 1$.

Cyclic code $C$ is invariant under a cyclic shift implies that if $c(x)$ is in $C$, then so is $xc(x)$ provided we multiply modulo $x^n - 1$.

Remark :

1. A linear code $C$ is cyclic if and only if $c(x) \in C \implies x \cdot c(x) \pmod {x^n - 1} \in C$.

   It follows that when $c(x)$ is a codeword in a cyclic code, so are the words $x^i \cdot c(x) \pmod {x^n - 1}$ for $i \geq 0$.

2. By linearity we conclude that in a cyclic code $C$, $c(x) \in C \implies u(x)c(x) \pmod {x^n - 1} \in C$ for every $u(x) \in \mathbb{F}_q[x]$, hence $C$ is an ideal in $\mathcal{R}_n = \mathbb{F}_q[x]/(x^n - 1)$.
4.1 Factoring $x^n - 1$

For $f(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{F}_q[x]$ define the formal derivative of $f(x)$ to be the polynomial $f'(x) = a_1 + 2a_2 x + \cdots + na_n x^{n-1} \in \mathbb{F}_q[x]$. From this definition, show that the following rules hold:

(a) $(f + g)'(x) = f'(x) + g'(x)$

(b) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.

(c) $f(x)^m)' = m(f(x)^{m-1})f'(x)$ for all positive integers $m$.

(d) If $f_1(x)^{a_1} f_2(x)^{a_2} \cdots f_n(x)^{a_n}$, where $a_1, \cdots, a_n$ are positive integers and $f_1(x), \cdots, f_2(x)$ are distinct and irreducible over $\mathbb{F}_q$, the

$$\frac{f(x)}{\gcd(f(x), f'(x))} = f(x) \cdots f_n(x).$$

(e) Show that $f(x)$ has no repeated irreducible factors, if and only if $f(x)$ and $f'(x)$ are relatively prime.

(f) Show that $x^n - 1$ has no repeated irreducible factors, if and only if $q$ and $n$ are relatively prime.

Proof. (f)
Let $c$ be a root of $x^n - 1$ in an extension field $\mathbb{E}$ of $\mathbb{F}_q$. Then $x^n - 1 = (x - c)(x^{n-1} + cx^{n-2} + \cdots + c^{n-2}x + c^{n-1})$.

Let $f(x) = x^{n-1} + cx^{n-2} + \cdots + c^{n-1}, f(c) = c^{n-1} + c \cdot c^{n-2} + \cdots + c^{n-1} = nc^{n-1}$. Since $q \nmid n \implies f(c) \neq 0 \implies c$ is not a root of $f(x) \implies x^n - 1$ has no repeated roots.

$\Longleftrightarrow$ if $f(c) = nc^{n-1}$ and $q|n \implies f(c) = 0 \implies x^n - 1$ has repeated root. A contradiction. So $q$ and $n$ are relatively prime.

A method of factoring $x^n - 1$: Let $\mathbb{F}_{q^t}$ be a field extension of $\mathbb{F}_q$ that contains all of the roots of $x^n - 1$. $\mathbb{F}_{q^t}$ contains a primitive $n$th root of unity when $n|q^t - 1$. 76
Definition 4.1.1. The order ordₙ(q) of q modulo n is the smallest positive integer a such that \( q^a \equiv 1 \pmod{n} \). That is if \( q^a \equiv 1 \pmod{n} \implies \text{ord}_n(q) = a \).

If \( t = \text{ord}_n(q) \), then \( \mathbb{F}_{q^t} \) contains a primitive nth root of unity \( \alpha \), but no smaller extension field of \( \mathbb{F}_q \) contains such a primitive root.

As \( \alpha^i \) are distinct for \( 0 \leq i < n \) and \( (\alpha^i)^n = 1 \), \( \mathbb{F}_{q^t} \) contains all the roots of \( x^n - 1 \). So \( \mathbb{F}_{q^t} \) is called a splitting field of \( x^n - 1 \) over \( \mathbb{F}_q \). So the irreducible factors of \( x^n - 1 \) over \( \mathbb{F}_q \) must be the product of the distinct minimal polynomials of the roots of unity in \( \mathbb{F}_{q^t} \).

Suppose \( \gamma \) is a primitive element of \( \mathbb{F}_{q^t} \). Then \( \alpha = \gamma^d \) is a primitive nth root of unity where \( d = \frac{q^t - 1}{n} \). The roots of \( m_{\alpha}(x) \) are \( \{\gamma^{ds}, \gamma^{dsq}, \gamma^{dsq^2}, \ldots, \gamma^{dsq^{r-1}}\} = \{\alpha^s, \alpha^{sq}, \alpha^{sq^2}, \ldots, \alpha^{sq^{r-1}}\} \), where \( r \) is the smallest positive integer such that \( dsq^r \equiv ds \pmod{q^t - 1} \) (by theorem 3.7.6).

But \( dsq^r \equiv ds \pmod{q^t - 1} \) if and only if \( sq^r \equiv s \pmod{n} \), because \( (dsq^r \equiv ds \pmod{dn} \implies sq^r \equiv s \pmod{n}) \).

This leads us to extend the notion of \( q \)-cyclotomic cosets.

Let \( s \) be an integer with \( 0 \leq s < n \). The \( q \)-cyclotomic cosets of \( s \) modulo \( n \) is the set \( C_s = \{s, sq, \cdots, sq^{r-1}\} \pmod{n} \), where \( r \) is the smallest positive integer such that \( sq^r \equiv s \pmod{n} \). It follows that \( C_s \) is the orbit of the permutation \( i \mapsto iq \pmod{n} \) that contains \( s \). The distinct \( q \)-cyclotomic cosets modulo \( n \) partition the set of integers \( \{0, 1, 2, \cdots, n - 1\} \) into disjoint sets.

Theorem 4.1.1. Let \( n \) be a positive integer relatively prime to \( q \). Let \( t = \text{ord}_n(q) \). Let \( \alpha \) be a primitive nth root of unity in \( \mathbb{F}_{q^t} \).

(i) For each integer \( s \) with \( 0 \leq s < n \), the minimal polynomial of \( \alpha^s \) over \( \mathbb{F}_q \) is \( M_{\alpha^s}(x) = \prod_{i \in C_s}(x - \alpha^i) \), where \( C_s \) is the \( q \)-cyclotomic cosets of \( s \) modulo \( n \).

(ii) The conjugates of \( \alpha^s \) are the elements \( \alpha^i \) with \( i \in C_s \).

(iii) Furthermore \( x^n - 1 = \prod_s M_{\alpha^s}(x) \) in the factorization of \( x^n - 1 \) into irreducible factors over \( \mathbb{F}_q \), where \( s \) runs through a set of representation of the \( q \)-cyclotomic cosets modulo \( n \).
Example 4.1.1. Consider the polynomial $x^9 - 1$ over $\mathbb{F}_2$. Since $q = 2$ and $n = 9$, then the 2-cyclotomic cosets of 9 are $C_0 = \{0\}$, $C_1 = \{1, 2, 4, 8, 7, 5\}$, $C_3 = \{3, 6\}$. So $\text{ord}_n(2) = 6$ and the primitive ninth root of unity lie in $\mathbb{F}_{64}$ but no smaller extension field of $\mathbb{F}_2$. Hence $x^9 - 1$ factors into irreducible factors as

$$x^9 - 1 = (x^3)^3 - 1 = (x^3 - 1)(x^6 + x^3 + 1) = (x - 1)(x + x + 1)(x^6 + x^3 + 1).$$

The polynomial $m_\alpha(x) = M_1(x) = x + 1$, $M_\alpha(x) = x^6 + x^3 + 1$, $M_{\alpha^3}(x) = x^2 + x + 1$, where $\alpha$ is a primitive ninth root of unity in $\mathbb{F}_{64}$. The only irreducible polynomial of degree 2 over $\mathbb{F}_2$ is $x^2 + x + 1 = M_{\alpha^3}(x)$ ($\alpha^3$ is a primitive third root of unity in $\mathbb{F}_{64}$).

Theorem 4.1.2. The size of each $q$-cyclotomic coset is a divisor of $\text{ord}_n(q)$. Furthermore the size of $C_1$ is $\text{ord}_n(q)$.

Proof. Let $t = \text{ord}_n(q)$ and let $m$ be the size of $C_1$. Then $M_{\alpha^s}(x)$ has degree $m$ where $\alpha$ is a primitive $n$th root of unity, so $m|t$ by theorem 3.7.3 there exists a subfield $\mathbb{F}_{q^m} = \mathbb{F}_{q^r}$ of $\mathbb{F}_q$ such that $mr|mt \implies m|t$. The fact that the size of $C_1$ is $\text{ord}_n(q)$ follows directly from the definitions of $q$-cyclotomic cosets and $\text{ord}_n(q)$ as mentioned prior to theorem 4.1.1, because $C_1 = \{1, q, q^2, \ldots, q^{r-1}\} \pmod{q^t - 1}$, where $r$ is the smallest positive integer such that $q^r \equiv 1 \pmod{q^t - 1} \implies r = t$. \hfill $\Box$

### 4.2 Basic theory of cyclic codes

The cyclic codes of length $n$ over $\mathbb{F}_q$ are precisely the ideals of $\mathcal{R}_n = \mathbb{F}_q[x]/(x^n - 1)$, $\mathbb{F}_q[x]$ is a principal ideal domain also the ring $\mathcal{R}_n$ is principal, hence cyclic codes are the principal ideal of $\mathcal{R}_n$, when writing a codeword of a cyclic code as $c(x)$, we technically mean the coset $c(x) + (x^n - 1) \in \mathcal{R}_n$. We think the elements of $\mathcal{R}_n$ as the polynomial in $\mathbb{F}_q[x]$ of degree less than $n$ with multiplication being carried out modulo $x^n - 1$.

To multiply two polynomials we multiply them as we would in $\mathbb{F}_q[x]$ and then replace any term of the form $ax^{ni+j}$ where $0 \leq j < n$ by $ax^i$.

Theorem 4.2.1. Let $C$ be a nonzero cyclic code in $\mathcal{R}_n$. There exists a polynomial $g(x) \in C$ with the following properties:
(i) $g(x)$ is the unique monic polynomial of minimum degree in $\mathcal{C}$.

(ii) $\mathcal{C} = \langle g(x) \rangle$, and

(iii) $g(x)|(x^n - 1)$.

Let $k = n - \deg g(x)$, and let $g(x) = \sum_{i=0}^{n-k} g_i x^i$, where $g_{n-k} = 1$. Then:

(iv) The dimension of $\mathcal{C}$ is $k$ and $\{g(x), xg(x), \cdots, x^{k-1}g(x)\}$ is a basis for $\mathcal{C}$,

(v) every element of $\mathcal{C}$ is uniquely expressible as a product $g(x)f(x)$, where $f(x) = 0$ or $\deg f(x) < k$,

(vi)

$$G = \begin{pmatrix}
g_0 & g_1 & g_2 & \cdots & g_{n-k} & 0 \\
0 & g_0 & g_1 & \cdots & g_{n-k-1} & g_{n-k} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & g_0 & \cdots & g_{n-k} \\
\end{pmatrix} \leftrightarrow \begin{pmatrix}
g(x) \\
xg(x) \\
\cdots \\
x^{k-1}g(x) \\
\end{pmatrix}$$

is a generator matrix for $\mathcal{C}$, and

(vii) if $\alpha$ is a primitive $n$th root of unity in some extension field of $\mathbb{F}_q$, then

$$g(x) = \prod_s M_{\alpha^s}(x).$$

Where the product is over a subset of representatives of the $q$-cyclotomic cosets modulo $n$.

Proof. Let $g(x)$ be a monic polynomial of minimum degree in $\mathcal{C}$. Since $\mathcal{C}$ is nonzero, such a polynomial exists. If $c(x) \in \mathcal{C}$, then by division algorithm in $\mathbb{F}_q[x]$, $c(x) = g(x)h(x) + r(x)$, where either $r(x) = 0$ or $\deg r(x) < \deg g(x)$. As $\mathcal{C}$ is an ideal in $\mathcal{R}_n$, $r(x) \in \mathcal{C}$ and the minimality of degree of $g(x)$ implies $r(x) = 0 \implies \mathcal{C} = \langle g(x) \rangle$.

To proof uniqueness: let $g_1$, $g_2$ be two distinct monic polynomials of minimum degree $m$ then $g_1 - g_2 = 0$ or $\deg g_1 - g_2 < m$, a contradiction, then $g_1 = g_2$. This gives (i) and (ii).

(iii) By division algorithm $x^n - 1 = g(x)h(x) + r(x)$, where $r(x) = 0$ or $\deg r(x) < \deg g(x)$
Suppose that \( \text{deg } g(x) = n - k \). By parts (ii) and (iii), if \( c(x) \in C \) with \( c(x) = 0 \) or \( \text{deg } c(x) < n \), then \( c(x) = g(x)f(x) \in \mathbb{F}_q[x] \). If \( c(x) = 0 \), then \( f(x) = 0 \). If \( c(x) \neq 0 \), \( \text{deg } c(x) < n \) implies that \( \text{deg } f(x) < k \), (because the degree of the product of two polynomials is the sum of degrees of polynomials.) Therefore 
\[ C = \{ g(x)f(x) | f(x) = 0 \text{ or } \text{deg } f(x) < k \} \]

Therefore \( C \) has dimension at most \( k \) and \( \{ g(x), xg(x), \ldots, x^{k-1}g(x) \} \) spans \( C \). Since these \( k \) polynomials are of different degrees they are independent in \( \mathbb{F}_q[x] \). Since they are of at most \( n - 1 \), they remain independent in \( \mathbb{R}_n \), yielding (iv) and (v). The codewords in this basis, written as \( n \)-tuples, give \( G \) in part (vi). Part (vii) follows from theorem 4.1.1.

\[ \square \]

**Remark 4.2.1.** (i) \( \mathbb{R}_n = \mathbb{F}_q[x]/(x^n - 1) \) is a principal ideal ring.

(ii) Part (vii) requires that \( \text{gcd}(n, q) = 1 \) because \( x^n - 1 \) has no repeated roots if and only if \( \text{gcd}(n, q) = 1 \).

**Corollary 4.2.2.** Let \( C \) be nonzero cyclic code in \( \mathbb{R}_n \). The following are equivalent:

(i) \( g(x) \) is monic polynomial of minimum degree in \( C \).

(ii) \( C = \langle g(x) \rangle \), \( g(x) \) is monic, and \( g(x)|x^n - 1 \).

**Proof.** \( (i) \implies (ii) \) was shown in the proof of theorem 4.2.1.

Assume (ii). Let \( g_1(x) \) be monic polynomial of minimum degree in \( C \). By part (i) and (ii) of theorem 4.2.1 \( g(x) = g_1(x)h(x) + r(x) \), \( \text{deg } r(x) < \text{deg } g_1(x) \) or \( r(x) = 0 \implies g_1(x)|g(x) \) in \( \mathbb{F}_q[x] \) and \( C = \langle g_1(x) \rangle \). As \( g_1(x) \in C = \langle g(x) \rangle \), \( g_1(x) = g(x)a(x) (\text{mod } x^n - 1) \implies g_1(x) = g(x)a(x) + (x^n - 1)b(x) \in \mathbb{F}_q[x] \). Since \( g(x)|x^n - 1 \implies g(x)|g(x)a(x) + (x^n - 1)b(x) \) or \( g(x)|g_1(x) \). As both \( g_1(x) \) and \( g(x) \) are monic and divide one another in \( \mathbb{F}_q[x] \implies g_1(x) = g(x) \).

\[ \square \]

**Definition 4.2.1.** The monic polynomial \( g(x)|x^n - 1 \), generating \( C \) is called the generator matrix polynomial of the cyclic code \( C \). So there exists one to one correspondence between the nonzero cyclic codes and the divisors of \( x^n - 1 \), not equal to \( x^n - 1 \).
Remark 4.2.2. The generator of the zero cyclic code \( \{0\} \) is \( x^n - 1 \).

Corollary 4.2.3. The number of cyclic codes in \( R_n \) equals \( 2^m \) is the number of \( q \)-cyclotomic cosets modulo \( n \). Moreover, the dimension of cyclic codes in \( R_n \) are all possible sums of sizes of the \( q \)-cyclotomic cosets modulo \( n \).

Example 4.2.1. We showed that, over \( F_2 \), \( x^9 - 1 = (x + 1)(1 + x + x^6)(1 + x^3 + x^6) \), and so there are eight binary cyclic codes \( C_i \) of length 9 with generator polynomials \( g_i(x) \) are given in the following table

<table>
<thead>
<tr>
<th>( i )</th>
<th>( dim )</th>
<th>( g_i(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>( 1 + x^9 )</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>( (1 + x + x^2)(1 + x^3 + x^6) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7 + x^8 )</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>( (1 + x)(1 + x^3 + x^6) = 1 + x + x^3 + x^4 + x^6 + x^7 )</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>( 1 + x^3 + x^6 )</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>( (1 + x)(1 + x + x^2) = 1 + x^3 )</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>( 1 + x + x^2 )</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>( 1 + x )</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>( 1 )</td>
</tr>
</tbody>
</table>

Corollary 4.2.4. Let \( C_1 \) and \( C_2 \) be cyclic codes over \( F_q \) with generator polynomials \( g_1(x) \) and \( g_2(x) \) respectively. Then \( C_1 \subseteq C_2 \) if and only if \( g_2(x) | g_1(x) \).

Proof. Let \( \deg g_1(x) = t_1 \), \( \deg g_2(x) = t_2 \). If \( g_2(x) | g_1(x) \Rightarrow \deg g_2 \leq \deg g_1 \Rightarrow -\deg g_1 \leq -\deg g_2 \Rightarrow n - \deg g_1 \leq n - \deg g_2 \Rightarrow \dim C_1 \leq \dim C_2 \).

\( \Leftarrow \) if \( C_1 \subseteq C_2 \Rightarrow n - \deg g_1 \leq n - \deg g_2 \Rightarrow \deg g_2 \leq \deg g_1 \Rightarrow g_2 | g_1 \). \qed

Example 4.2.2. Consider \( R_3 = F_2[x]/(x^3 - 1) \) and consider the cyclic code \( C = \langle 1 + x \rangle \), then \( \dim C = 3 - 1 = 2 \) and \( C \) contains the codewords, 0, \( 1 + x \), \( x(1 + x) = x + x^2 \), \( x^2(1 + x) = x^2 + x^3 = x^2 + 1 = (x + 1)(x + 1) \), because \( x^3 = 1 \). Thus \( C = \{0, 1 + x, 1 + x^2, x + x^2\} = \{000, 110, 101, 011\} \). Also we can verify that \( \langle 1 + x^2 \rangle = \{f(x)(1 + x^2) | f(x) \in R_3\} = C \) and so \( C \) is generated by the polynomial \( 1 + x^2 \) as well.
Example 4.2.3. $x^9 - 1$ factors over $\mathbb{F}_2$ into irreducible factors $x^9 - 1 = (x + 1)(x^2 + x + 1)(x^6 + x^3 + 1)$.

Consider $C = \langle x^6 + x^3 + 1 \rangle$, therefore $\dim C = 9 - 6 = 3$ and has generator matrix,

$$ G = \begin{pmatrix} x^6 + x^3 + 1 \\ x(x^6 + x^3 + 1) \\ x^2(x^6 + x^3 + 1) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} $$

Theorem 4.2.5. The dual code of a cyclic code is cyclic.

Proof. If $a \cdot b = 0$, then $\pi(a) \cdot \pi(b) = 0$, where $\pi$ is the cyclic shift. As $a \cdot b = a_0b_0 + a_1b_1 + \cdots + a_nb_n$, then $\pi(a) \cdot \pi(b) = a_nb_n + a_1b_1 + \cdots + a_0b_0 = 0$. Consider the cyclic code $C$ which is generated by the word $v$; so $C = \{ v, \pi(v), \pi^2(v), \cdots, \pi^{n-1}(v) \}$. If $u \in C^\perp \implies \pi^i(v) \cdot u = 0$, for $i = 0, \cdots, n - 1$. However this means that $\pi^{i+1}(v) \cdot u = 0 \implies u$ is orthogonal to $C$, because $\pi^n(v) = v$. Since $u \in C^\perp \implies u \in C^\perp \implies C^\perp$ is cyclic. \square

Definition 4.2.2. The generator matrix of the dual code of cyclic code $C$ is the parity check matrix of the original cyclic code.

Theorem 4.2.6. Let $C$ be an $[n, k]$ cyclic code with generator polynomial $g(x)$. Let $h(x) = (x^n - 1)/g(x) = (x^n - 1)/g(x) = \sum_{i=0}^{k} h_i x^i$. Then the generator polynomial of $C^\perp$ is $g^\perp(x) = \frac{x^n h(x^{-1})}{h(0)}$. Furthermore, a generator matrix for $C^\perp$, and hence a parity check matrix for $C$ is

$$ H = \begin{pmatrix} h_k & h_{k-1} & h_{k-2} & \cdots & h_0 & 0 \\ 0 & h_k & h_{k-1} & \cdots & h_1 & h_0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & h_k & \cdots & h_0 \end{pmatrix} $$

Proof. The dual of a cyclic code is cyclic, we will show that $g^\perp(x)$ divides $x^n - 1$, then we will know that it is the generator polynomial for a cyclic code $\langle g^\perp(x) \rangle$ that has generator matrix $H$, and so $\langle g^\perp(x) \rangle = C^\perp$.

Let $h(x)g(x) = x^n - 1$, then $h(0)g(0) = -1 \implies h(0) = -1$ since $g(0) = 1$ as $g(x) = g_0 + g_1 x + \cdots + g_{n-k} x^{n-k}$, $g_{n-k} = 1$. \[82\]
\[\Rightarrow h(x^{-1})g(x^{-1}) = x^{-n} - 1,\]
\[x^n h(x^{-1}) g(x^{-1}) = 1 - x^n,\]
\[x^k h(x^{-1}) \cdot x^{n-k} g(x^{-1}) = 1 - x^n,\]
\[\frac{x^k h(x^{-1})}{h(0)} \cdot x^{n-k} g(x^{-1}) = x^n - 1,\]
\[g^\perp \cdot x^{n-k} g(x^{-1}) = x^n - 1,\]
\[\Rightarrow g^\perp(x)|x^n - 1\]
\[\therefore g^\perp(x) \text{ is a generator polynomial of } \mathcal{C}^\perp.\]

Since \(g(x)\) is the generator polynomial for \(\mathcal{C}\), and \(h(x) = \frac{x^n-1}{g(x)}\), then \(g(x)h(x) = 0\). If \(c(x) \in \mathcal{C}\), then \(c(x)h(x) = 0\).

Now \(\deg (c(x)h(x)) < n + k = n + n - \deg g(x) = 2n - \deg g(x)\), from this we deduce that the coefficients of \(x^k, x^{k+1}, \ldots, x^{n-1}\) in the product \(c(x)h(x)\) must be 0, that is
\[c_0h_k + c_1h_{k-1} + \cdots + c_kh_0 = 0,\]
\[c_1h_k + c_2h_{k-1} + \cdots + c_{k+1}h_0 = 0,\]
\[\vdots \quad \vdots \quad \vdots \]
\[c_{n-k-1}h_k + c_{n-k}h_{k-1} + \cdots + c_{n-1}h_0 = 0.\]

But this is equivalent to \(c_0c_1 \cdots c_{n-1})H^\perp = 0\), and so \(H\) generates a code \(\mathcal{C}'\) that is orthogonal to \(\mathcal{C}\), that is \(\mathcal{C}' \subset \mathcal{C}^\perp\). However, since \(h_k \neq 0\), it follows that \(\dim (\mathcal{C}') = \deg g(x)\) and so \(\mathcal{C}' = \mathcal{C}^\perp \Rightarrow H\) is the parity check matrix for \(\mathcal{C}\). \(\square\)

**Example 4.2.4.** The code \(\mathcal{C} = \langle x^6 + x^3 + 1 \rangle\) has generator polynomial \(g(x) = x^6 + x^3 + 1\) and has dimension \(k = 9 - 6 = 3\).

\[\therefore h(x) = \frac{x^n-1}{g(x)} = (1 + x)(1 + x + x^2) = 1 + x^3.\]

The generator polynomial of \(\mathcal{C}^\perp\) is \(g^\perp(x) = \frac{x^k h(x^{-1})}{h(0)} = \frac{x^3(1+x^{-3})}{1} = x^3 + 1,\)

83
The generator matrix for $C_{\text{bot}}$ is

$$
H = \begin{pmatrix}
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\vspace{4pt}
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0
\vspace{4pt}
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0
\vspace{4pt}
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0
\vspace{4pt}
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0
\vspace{4pt}
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1
\end{pmatrix}
$$

Remark 4.2.3. (i) The polynomial $h(x)$ is called the check polynomial for a cyclic code $C$; it is not the generator polynomial for $C^\perp$.

(ii) The polynomial $g^\perp(x) = \frac{x^k h(x^{-1})}{h(0)}$ is called the reverse polynomial of the check polynomial $h(x)$ and $g^\perp(x)$ is the generator polynomial for $C^\perp$.

If $h(x) = h_0 + h_1 x + \cdots + h_k x^k$, then the generator polynomial of $C^\perp$ is

$$
g^\perp(x) = \frac{x^k h(x^{-1})}{h(0)} = h^{-1}(0)(h_0 + h_1 x^{-1} + \cdots + h_k x^{-k})
$$

$$
= h^{-1}(0)(h_0 x^k + h_1 x^{k-1} + \cdots + h_k) = h^{-1}(0)(h_k + h_{k-1} + \cdots + h_0 x^k)
$$

Theorem 4.2.7. Let $C$ be a cyclic code over $\mathbb{F}_{q^t}$. Then $C|_{\mathbb{F}_q}$ is also cyclic.

Proof. Since $C$ is cyclic code over $\mathbb{F}_{q^t}$, there exists a generator polynomial $g(x)$ over $\mathbb{F}_{q^t}[x]$ that generates $C$ and $C = \langle g(x) \rangle$, $g(x)|x^n - 1$ such that $\gcd(q^t, n) = 1$ and $g(x)$ containing a primitive $n$th root of unity of the extension field $\mathbb{F}_{q^t}$ of $\mathbb{F}_{q^t}$, when $n|(q^t)^s - 1$. Since $\mathbb{F}_{q^t}$ contains a primitive $n$th root of unity, so there exists $g_1(x)$ over $\mathbb{F}_q$ and $g_1(x)|x^n - 1$, $\gcd(q, n) = 1$ so $C|_{\mathbb{F}_q} = \langle g_1(x) \rangle$, so $C|_{\mathbb{F}_q}$ is cyclic.

Encoding and decoding cyclic codes There is three ways to encode the cyclic codes. If $C$ is a cyclic code of length $n$ over $\mathbb{F}_q$ with generator polynomial $g(x)$ of degree $n - k$, so $C$ has dimension $k$.

The first way:(This method is non systematic) Let $G$ be the generator matrix obtained
from the shifts of \( g(x) \) in Theorem 4.2.1

\[
G = \begin{pmatrix}
g(x) \\
xg(x) \\
\vdots \\
x^{k-2}g(x)
\end{pmatrix} = \begin{pmatrix}
g_0 & \cdots & g_{n-k} & 0 \\
0 & g_0 & \cdots & g_{n-k} \\
\vdots & \vdots & \ddots & \vdots \\
0 & g_0 & \cdots & g_{n-k}
\end{pmatrix}.
\]

To encode the message \( m \in \mathbb{F}_q^k \) as the codeword \( c = mG \), let \( m(x) = a_0 + a_1 x + \cdots + a_{k-1} x^k \in \mathbb{F}_q[x] \). Then to encode \( m(x) \) as a codeword \( c(x) \) by forming the product

\[
c(x) = m(x)g(x).
\]

**Example 4.2.5.** Let \( C \) be a cyclic code of length 15 with generator polynomial \( g(x) = (1 + x + x^4)(1 + x + x^2 + x^3 + x^4) \). Encode the message \( m(x) = 1 + x^2 + x^5 \) using the first encoding procedure.

**Solution:** \( g(x) = (1 + x + x^4)(1 + x + x^2 + x^3 + x^4) = 1 + x^4 + x^6 + x^7 + x^8 \). Then \( c(x) = m(x)g(x) = (1 + x^2 + x^5)(1 + x^4 + x^6 + x^7 + x^8) = 1 + x^2 + x^4 + x^5 + x^7 + x^{10} + x^{11} + x^{12} + x^{13} \) as a vector in \( \mathbb{F}_2^{15} = (101001101001101) \).

The second way of encoding cyclic codes: (This method is systematic).

The polynomial \( m(x) \) associated to the message is of degree at most \( k - 1 \) or (the zero polynomial). The polynomial \( x^{n-k}m(x) \) has degree at most \( n - 1 \) and has its first \( n - k \) coefficients equals to 0. Thus the message is contained in the coefficients of \( x^{n-k}, x^{n-k+1}, \ldots, x^{n-1} \). By the division algorithm, \( x^{n-k}m(x) = g(x)a(x) + r(x) \), where \( \deg r(x) < n - k \) or \( r(x) = 0 \).

Let \( c(x) = x^{n-k}m(x) - r(x) \), as \( c(x) \) is a multiple of \( g(x) \), \( c(x) \in C \). Also \( c(x) \) differs from \( x^{n-k}m(x) \) in the coefficients of 1, \( x, \cdots, x^{n-k-1} \) as \( \deg r(x) < n - k \). So \( c(x) \) contains the message \( m \) in the coefficients of the terms of degree at least \( n - k \).

**Example 4.2.6.** Encode the message \( m(x) = 1 + x^2 + x^5 \) using the second encoding procedure. (The systematic encoding).

**Solution:** Since from previous example \( g(x) = 1 + x^4 + x^6 + x^7 + x^8 \), \( n = 15 \), \( k = n - \deg g(x) = 15 - 8 = 7 \)

85
\[ x^{n-k} = x^{15-7} = x^8 \implies x^{n-k}m(x) = x^8(1 + x^2 + x^5) = x^8 + x^{10} + x^{13} . \]

By dividing \( x^8 + x^{10} + x^{13} \) by \( g(x) = 1 + x^4 + x^6 + x^7 + x^8 \) we get
\[ x^{13} + x^{10} + x^8 = (x^5 + x^4 + x + 1)g(x) + (x^6 + x + 1) \text{ where } a(x) = x613 + x^{10} + x^8 \text{ and } r(x) = x^6 + x + 1. \]
\[ c(x) = x^{13} + x^{10} + x^8 + x^6 + x + 1. \text{ The message is contained in the coefficients } x^{n-k}, x^{n-k+1}, \ldots, x^{n-1} \text{ systematics.} \]

**The third method of encoding:** this method is systematic. Let \( C = \langle g(x) \rangle \) be a cyclic code. Let \( g^i(x) \) be the generator polynomial of \( C^\perp \) and \( C \) is an \([n, k]-\)code. If \( c = (c_0 c_1 \cdots c_{n-1}) \in C \) once \( c_0 c_1 \cdots c_{k-1} \) are known, then the remaining components \( c_k, \ldots, c_{n-1} \) are determined from \( Hc^\perp = 0 \), where \( H \) is the parity check matrix. We can scale the rows of \( H \) so that its rows are shifts of the monic polynomial \( g^i(x) = h'_0 + h'_1 x + \cdots + h'_{k-1} x^{k-1} + x^k \). To encode \( C \), we chose \( k \) information bits \( c_0 c_1 \cdots c_{k-1} \), then \( c_i = -\sum_{j=0}^{k-1} h'_j c_{i-k+j} \), where the computation \( c_i \) is performed in the order \( i = k, k + 1, \ldots, n - 1 \).

**Example 4.2.7.** Encode the message \( m(x) = 1 + x^2 + x^5 \) using the third encoding procedure.

**Solution:** \( h(x) = \frac{x^{15-1}}{g(x)} \), where \( g(x) = 1 + x^4 + x^6 + x^7 + x^8 \). Then \( h(x) = x^7 + x^6 + x^4 + 1 \).
\( n = 15, \operatorname{deg} g(x) = 8 \implies k = 7. \)
\( g^i(x) = x^k h(x^{-1})/h(0) = x^7(x^{-7} + x^{-6} + x^{-4} + 1)/1 = 1 + x + x^3 + x^7. \)
\( h_0 = 1, h_1 = 1, h_2 = 0, h_3 = 1, h_4 = 0, h_5 = 0, h_6 = 0, h_7 = 1. \)

\[
H = \begin{pmatrix}
1 + x + x^3 + x^7 \\
xg^i(x) = x(1 + x + x^3 + x^7) \\
\vdots \\
x^7(1 + x + x^3 + x^7)
\end{pmatrix} = \begin{pmatrix}
1 + x + x^3 + x^7 \\
x + x^2 + x^4 + x^8 \\
x^2 + x^3 + x^5 + x^9 \\
x^3 + x^4 + x^6 + x^{10} \\
x^4 + x^5 + x^7 + x^{11} \\
x^5 + x^6 + x^8 + x^{12} \\
x^6 + x^7 + x^9 + x^{13} \\
x^7 + x^8 + x^{10} + x^{14}
\end{pmatrix}
\]

86
Let the information bits are \((c_0 c_1 \cdots c_6)\), since \(m(x) = 1 + x^2 + x^5 \mapsto (1010010)\). Then
\[
c_7 = -\sum_{j=0}^{6} h_j c_{7-j} = +(h_0 c_0 + h_1 c_2 + h_2 c_3 + \cdots + h_6 c_6) = (c_0 + c_1 + 0 + c_3 + 0 + 0 + 0) = (c_0 + c_1 + c_3)
\]
\[
c_8 = \sum_{j=0}^{6} h_j c_{8-j} = +(h_0 c_1 + h_1 c_2 + h_2 c_3 + \cdots + h_6 c_7) = (c_1 + c_2 + c_4)
\]
similarly we do for all other bits.

4.3 Idempotent and multipliers

An element \(e\) of a ring satisfying \(e^2 = e\) is called an idempotent. Each cyclic code in \(R_n = \mathcal{U}_q[x]/\langle x^n - 1 \rangle\) contains a unique idempotent which generates the ideal. This idempotent is called the generating idempotent of the cyclic code.

**Definition 4.3.1.** The unity in a ring is (an non zero) multiplicative identity in the ring.

**Example 4.3.1.** \(e(x) = x^3 + x^5 + x^6 \in R_7\) over \(\mathbb{F}_2[x]\) \(\implies e^2(x) = (x^3 + x^5 + x^6)^2 = x^3 + x^5 + x^6\).
\[\therefore\] \(e(x)\) is an idempotent in \(R_7\).

**Example 4.3.2.** The generating idempotent for the zero cyclic code \(\{0\}\) is 0, while that for the cyclic code \(R_n\) is 1.

**Theorem 4.3.1.** Let \(C\) be a cyclic code in \(R_n\). Then:

(i) there exists a unique idempotent \(e(x) \in C\) such that \(C = \langle e(x) \rangle\), and
(ii) if \( e(x) \) is a nonzero idempotent in \( C \), then \( C = \langle e(x) \rangle \) if and only if \( e(x) \) is a unity of \( C \).

**Proof.** If \( C \) is the zero code, then the idempotent is the zero polynomial and (i) is clear. To prove (ii). Assume \( C \) is nonzero cyclic code. Suppose that \( e(x) \) is a unity, then \( \langle e(x) \rangle \subseteq C \) as \( C \) is an ideal....(1).

If \( c(x) \in C \), then \( c(x)e(x) = c(x) \in C \implies c(x) \in \langle e(x) \rangle \implies C \subseteq \langle e(x) \rangle \) ...(2).

From (1) and (2), we have \( C = \langle e(x) \rangle \).

Conversely, suppose that \( e(x) \neq 0 \) idempotent such that \( C = \langle e(x) \rangle \). Then every element \( c(x) \in C \) can be written in the form \( c(x) = f(x)e(x) \). But \( c(x)e(x) = f(x)(e(x))^2 = f(x)e(x) = c(x) \) implying \( e(x) \) is a unity for \( C \).

To prove (i). As \( C \) is nonzero by part (ii), if \( e_1(x) \) and \( e_2(x) \) are generating idempotents, then both are unities and \( e(x) = e_2(x)e_1(x) = e_2(x) \implies e_1(x) = e_2(x) \). So the generating idempotent is unique.

To prove the existence of generating idempotent. If \( g(x) \) is the generator polynomial for \( C \), then \( g(x)|x^n - 1 \), by theorem 4.2.1, let \( h(x) = \frac{x^n - 1}{g(x)} \). Thus \( \gcd(g(x), h(x)) = 1 \) in \( \mathbb{F}_q[x] \) because \( x^n - 1 \) has distinct roots. Now by Euclidean algorithm there exist polynomials \( a(x), b(x) \in \mathbb{F}_q[x] \), so that \( a(x)g(x) + b(x)h(x) = 1 \). Let \( e(x) \equiv a(x)g(x)(\mod x^n - 1) \); that is \( e(x) \) is the coset representative of \( e(x) = a(x)g(x) + (x^n - 1) \in \mathcal{R}_n \). Then \( (e(x))^2 \equiv a(x)g(x)(1 - b(x)h(x)) \equiv a(x)g(x) - a(x)g(x)b(x)h(x) \equiv a(x)g(x) \equiv e(x)(\mod x^n - 1) \), (because \( g(x)h(x) = x^n - 1 \)). Also if \( c(x) \in C \), then \( c(x) = f(x)g(x) \) implying \( c(x)e(x) \equiv f(x)g(x)(1 - b(x)h(x)) \equiv f(x)g(x) - f(x)g(x)b(x)h(x) \equiv f(x)g(x) \equiv c(x)(\mod x^n - 1) \).

So \( e(x) \) is a unity in \( C \) and (i) follows from (ii). \( \square \)

**Remark 4.3.1.** to find the generating idempotent \( e(x) \) for a cyclic code \( C \) from the generator polynomial \( g(x) \) is to solve \( 1 = a(x)g(x) + b(x)h(x) \) for \( a(x) \) using the Euclidean algorithm, where \( h(x) = \frac{x^n - 1}{g(x)} \). Then reducing \( a(x)g(x) \) modulo \( x^n - 1 \) produces \( e(x) \).

**Theorem 4.3.2.** Let \( C \) be a cyclic code over \( \mathbb{F}_q \) with generating idempotent \( e(x) \). Then the generator polynomial of \( C \) is \( g(x) = \gcd(e(x), x^n - 1) \) computed in \( \mathbb{F}_q[x] \).
Proof. Let \( d(x) = \gcd(e(x), x^n - 1) \) in \( \mathbb{F}_q[x] \), and let \( g(x) \) be the generator polynomial for \( C \). As \( d(x)|e(x) \implies e(x) = d(x)k(x) \) implying that every element of \( C = \langle e(x) \rangle \) is also a multiple of \( d(x) \), thus \( C \subseteq \langle d(x) \rangle \). (1). By theorem 4.2.1 in \( \mathbb{F}_q[x] \), \( g(x)|x^n - 1 \) and \( g(x)|e(x) \) as \( e(x) \in C \implies g(x)|\gcd(e(x), x^n - 1) \implies g(x)|d(x) \implies d(x) = g(x)l(x) \implies d(x) \in C \). Thus \( \langle d(x) \rangle \subseteq C \). Thus \( \langle d(x) \rangle \subseteq C \). ... (2).

From (1) and (2) \( C = \langle d(x) \rangle \). Since \( d(x) \) is a monic divisor of \( x^n - 1 \) generating \( C \) \( \implies d(x) = g(x) \) by corollary 4.2.2

Example 4.3.3. The following table gives all the cyclic codes \( C_i \) of length 7 over \( \mathbb{F}_2 \) together with their generator polynomials \( g_i(x) \) and their generating idempotents \( e_i(x) \).

<table>
<thead>
<tr>
<th>( i )</th>
<th>( \text{dim} )</th>
<th>( g_i(x) )</th>
<th>( e_i(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>( 1 + x^7 )</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>( 1 + x + x^2 + x^3 + x^4 + x^5 + x^6 )</td>
<td>( 1 + x + x^2 + x^3 + x^4 + x^5 + x^6 )</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>( 1 + x^2 + x^3 + x^4 )</td>
<td>( 1 + x^3 + x^5 + x^6 )</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>( 1 + x + x^2 + x^4 )</td>
<td>( 1 + x + x^2 + x^4 )</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>( 1 + x + x^3 )</td>
<td>( x + x^2 + x^4 )</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>( 1 + x^2 + x^3 )</td>
<td>( x^3 + x^5 + x^6 )</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>( 1 + x )</td>
<td>( x + x^2 + x^3 + x^4 + x^5 + x^6 )</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

To find \( e_1(x) \) if \( g(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6 \) we first find \( h(x) = \frac{x^7 + 1}{g(x)} = x + 1 \).

Then we form the equation \( 1 = a(x)b(x) + b(x)h(x) \) by using Euclidean algorithm, we divide \( g(x) \) by \( h(x) \) we get

\[
g(x) = (x + 1)(x^5 + x^3 + x) + 1, \text{ then } 1 = 1 \cdot g(x) + (x + 1)(x^5 + x^3 + x), \text{ where } b(x) = (x^5 + x^3 + x) \text{ and } a(x) = 1
\]

\[
\therefore e(x) = 1 \cdot g(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6.
\]

Theorem 4.3.3. Let \( C \) be an \([n, k]\) cyclic code with generating idempotent \( e(x) = \sum_{i=0}^{n-1} e_i x^i. \)
Then the $k \times n$ matrix

\[
\begin{pmatrix}
    e_0 & e_1 & e_2 & \cdots & e_{n-2} & e_{n-1} \\
    e_{n-1} & e_0 & e_1 & \cdots & e_{n-3} & e_{n-2} \\
    \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
    e_{n-k+1} & e_{n-k+2} & \cdots & \cdots & e_{n-k-1} & e_{n-k}
\end{pmatrix}
\]

is a generator matrix for $C$.

Proof. The matrix is equivalent to $\{e(x), xe(x), \cdots, x^{k-1}e(x)\}$. The basis of $C$. Therefore it suffices to show that $a(x) \in \mathbb{F}_q[x]$ has degree less than $k$ such that $a(x)e(x) = 0$ then $a(x) = 0$. Let $g(x)$ be the generator polynomial for $C$. If $a(x)e(x) = 0$, then $0 = a(x)e(x)g(x) = a(x)g(x)$ as $e(x)$ is the unity of $C$, contradiction to theorem 4.2.1(v). \(\square\)

Definition 4.3.2. If $C_1$ and $C_2$ are codes of length $n$ over $\mathbb{F}_q$, then $C_1 + C_2 = \{c_1 + c_2 | c_1 \in C_1$ and $c_2 \in C_2\}$ is the sum of $C_1$ and $C_2$. Both the sum of and intersection of two cyclic codes are cyclic.

Theorem 4.3.4. Let $C_i$ be a cyclic code of length $n$ over $\mathbb{F}_q$ with generator polynomial $g_i(x)$ and generating $e_i(x)$ for $i = 0$ and 2. Then

(i) $C_1 \cap C_2$ has generator polynomial $lcm(g_1(x), g_2(x))$ and generating idempotent $e_1(x)e_2(x)$, and

(ii) $C_1 + C_2$ has generator polynomial $gcd(g_1(x), g_2(x))$ and generating idempotent $e_1(x) + e_2(x) - e_1(x)e_2(x)$.

Proof. (i) Let $l(x) = lcm(g_1(x), g_2(x)) \implies g_1(x)|l(x)$, $g_2(x)|l(x) \implies e(x) = g_1(x)k_1(x)$, $l(x) = g_2(x)k_2(x) \implies e(x) \in C_1$ and $l(x) \in C_2$, since $C_1 \cap C_2$ is cyclic, then $\langle l(x) \rangle \subseteq C_1 \cap C_2$. (1)

Let $l'(x) \in C_1 \cap C_2 \implies l'(x) \in C_1, l'(x) \in C_2 \implies l'(x) = g_1(x)f_1(x)$, $l'(x) = g_2(x)f_2(x) \implies g_1|l'$ and $g_2|l'(x) \implies l'$ is a common multiple $\implies l(x)|l'(x) \implies \langle l'(x) \rangle \subseteq \langle l(x) \rangle \implies C_1 \cap C_2 \subseteq \langle l(x) \rangle$. (2)

From (1) and (2) we have $C_1 \cap C_2 = lcm(g_1(x), g_2(x))$, since $g_1(x)|x^n - 1$ and $g_2(x)|x^n - 1 \implies l(x)|x^n - 1 \implies l(x)$ is a monic polynomial and generates $C_1 \cap C_2$. 

90
If \( c(x) \in C_1 \cap C_2 \implies c(x) \in C_1, \ c(x) \in C_2 \) and \( c(x)(e_1(x)e_2(x) = c(x)e_2(x) = c(x) \implies e_1e_2 \in C_1 \cap C_2 \) and \( e_1e_2 \) is the unity generator idempotent of \( C_1 \cap C_2 \).

(ii) Let \( g(x) = \gcd(g_1(x), g_2(x)) \) It follows from the Euclidean algorithm that \( g(x) = g_1(x)a(x) + g_2(x)b(x) \) for some \( a(x) \) and \( b(x) \) in \( \mathbb{F}_q[x] \). So \( g_1(x) \in C_1 + C_2 \). Since \( C_1 + C_2 \) is cyclic, \( \langle g(x) \rangle \subseteq C_1 + C_2 \). On other hand \( g(x)|g_1(x) \), which shows that \( C_1 \subseteq \langle g(x) \rangle \), similarly \( C_2 \subseteq \langle g(x) \rangle \) implying that \( C_1 + C_2 \subseteq \langle g(x) \rangle \). So \( C_1 + C_2 = \langle g(x) \rangle \). Since \( g(x)|x^n - 1 \) as \( g(x)|g_1(x) \) and \( g(x) \) is monic, \( g(x) \) is the generator polynomial for \( C_1 + C_2 \). If \( c(x) = c_1(x) + c_2(x) \) where \( c_i(x) \in C_i \) for \( i = 1 \) and \( 2 \), then \( c(x)(e_1(x) + e_2(x) - e_1(x)e_2(x)) = c_1(x) + c_1(x)e_2(x) - c_1(x)e_2(x) + c_2(x)e_1(x) + c_2(x)e_2(x) - c_2(x)e_1(x) = c(x) \).

Thus (ii) follows, since \( e_1(x) + e_2(x) - e_1(x)e_2(x) \in C_1 + C_2 \) and \( (e_1 + e_2 - e_1e_2)^2 = e_1 + e_2 - e_1e_2 \).

Example 4.3.4. If \( e_1(x) \) and \( e_2(x) \) are idempotents, so \( e_1(x)e_2(x), e(x) + e(x) - e_1(x)e_2(x) \) and \( 1 - e_1(x) \).

Solution: \( (e_1e_2)^2 = e_1^2e_2^2 = e_1e_2 \) and \( (1 - e_1)^2 = 1 - 2e_1 + e_1^2 = 1 - 2e_1 + e_1 = 1 - e_1 \).

**primitive idempotents:** Let \( x^n - 1 = f_1(x) \cdots f_s(x) \), where \( f_i(x) \) is irreducible over \( \mathbb{F}_q \) for \( 1 \leq i \leq s \). Then \( f_i(x) \)'s are distinct as \( x^n - 1 \) has distinct roots. Let \( \hat{f}_i(x) = \frac{x^n - 1}{f_i(x)} \), the ideals \( \langle \hat{f}_i(x) \rangle \) of \( \mathcal{R}_n \) are the minimal ideals of \( \mathcal{R}_n \).

Definition 4.3.3. An ideal \( I \) in a ring \( \mathcal{R} \) is a minimal ideal provided there is no proper ideal between \( \{0\} \) and \( I \). The generating idempotent of \( \langle \hat{f}_i(x) \rangle \) is \( \hat{e}_i(x) \) and is called the primitive idempotents of \( \mathcal{R}_n \).

Theorem 4.3.5. The following hold in \( \mathcal{R}_n \).

(i) The ideals \( \langle \hat{f}_i(x) \rangle \) for \( 1 \leq i \leq s \) are all the minimal ideals of \( \mathcal{R}_n \).

(ii) \( \mathcal{R}_n \) is the vector space direct sum of \( \langle \hat{f}_i(x) \rangle \) for \( 1 \leq i \leq s \).

(iii) If \( i \neq j \), then \( \hat{e}_i(x)\hat{e}_j(x) = 0 \) in \( \mathcal{R}_n \).

(iv) \( \sum_{i=1}^{s} \hat{e}_i = 1 \) in \( \mathcal{R}_n \).

91
(v) The only idempotents in \( \langle \hat{f}_i(x) \rangle \) are 0 and \( \hat{e}_i(x) \).

(vi) If \( e(x) \) is a non zero idempotent in \( R_n \), then there is a subset \( T \) of \( \{1, 2, \ldots, s\} \) such that \( e(x) = \sum_{i \in T} \hat{e}_i(x) \) and \( \langle e(x) \rangle = \sum_{i \in T} \langle \hat{f}_i(x) \rangle \).

Proof. (i) Suppose that \( \langle \hat{f}_i(x) \rangle \) is not a minimal ideal of \( R_n \). So there is a generator polynomial \( g(x) \) of a nonzero ideal properly contained in \( \langle \hat{f}_i(x) \rangle \) such that \( \hat{f}_i(x) | g(x) \) and \( g(x) \neq \hat{f}_i(x) \). As \( \hat{f}_i(x) \) is irreducible and \( g(x) | x^n - 1 \), this is impossible, so \( \langle \hat{f}_i(x) \rangle \) is minimal ideal.

(ii) As \( \{ \hat{f}_i(x) | 1 \leq i \leq s \} \) has no common irreducible factor of \( x^n - 1 \) and each polynomial in the set divides \( x^n - 1 \), \( \gcd(\hat{f}_i(x), \ldots, \hat{f}_s(x)) = 1 \). Applying the Euclidean algorithm inductively, \( 1 = \sum_{i=1}^{s} a_i(x) \hat{f}_i(x) \ldots (1) \), for some \( a_i(x) \in F_q[x] \). So 1 is the sum of the ideals \( \langle \hat{f}_i(x) \rangle \), which is itself an ideal of \( R_n \). In any ring, the only ideal containing the identity of the ring is the ring itself. This proves that \( R_n \) is the vector space sum of the ideals \( \langle \hat{f}_i(x) \rangle \). To prove it is a direct sum, we must show that \( \langle \hat{f}_i(x) \rangle \cap \sum_{j \neq i} \langle \hat{f}_j(x) \rangle = \{0\} \), for \( 1 \leq i \leq s \).

As \( f_i(x) | \hat{f}_j(x) \) for \( j \neq i \), \( f_j(x) \nmid \hat{f}_j(x) \), and the irreducible factors of \( x^n - 1 \) are distinct, we conclude that \( f_i(x) = \gcd\{f_j(x) | 1 \leq j \leq s, j \neq i \} \). Applying induction to the results of the previous theorem (ii) shows that \( \langle f_i(x) \rangle = \sum_{j \neq i} \langle \hat{f}_j(x) \rangle \). So \( \langle \hat{f}_i(x) \rangle \cap \sum_{j \neq i} \langle \hat{f}_j(x) \rangle = \langle \hat{f}_i(x) \rangle \cap \langle f_i(x) \rangle = \langle \text{lcm}(\hat{f}_i(x), f_i(x)) \rangle = \langle x^n - 1 \rangle = \{0\} \) (by theorem 4.3.7). Then part (ii) is completed.

To complete the proof of part (i). Let \( M = \langle m(x) \rangle \) be any minimal ideal of \( R_n \). As \( 0 \neq m(x) = m(x) \cdot 1 = \sum_{i=1}^{s} m(x)a_i(x)\hat{f}_i(x) \), by (1) there is an \( i \) such that \( m(x)a_i(x)\hat{f}_i(x) \neq 0 \). Hence \( M \cap \langle \hat{f}_i(x) \rangle \neq \{0\} \) as \( m(x)a_i(x)\hat{f}_i(x) \in M \cap \langle \hat{f}_i(x) \rangle \), and therefore \( M = \langle \hat{f}_i(x) \rangle \) by minimality of \( M \) and \( \langle \hat{f}_i(x) \rangle \). Thus completes the proof of (i).

(iii) If \( i \neq j \), \( \hat{e}_i(x)\hat{e}_j(x) \in \langle \hat{f}_i(x) \rangle \cap \langle \hat{f}_j(x) \rangle = \{0\} \) by (ii) we have \( \hat{e}_i\hat{e}_j = 0 \) in \( R_n \).
(iv) By using (iii) and applying induction to theorem 4.3.7 (ii), \( \sum_{i=1}^{s} \hat{e}_i(x) \) is the generating idempotent of \( \sum_{i=1}^{s} \langle \hat{f}_i(x) \rangle = R_n \), by part (ii). The generating idempotent of \( R_n \) is 1.

(v) If \( e(x) \) is a nonzero idempotent in \( \langle \hat{f}_i(x) \rangle \), then \( \langle e(x) \rangle \) is an ideal contained in \( \langle \hat{f}_i(x) \rangle \).

By minimality as \( e(x) \) is nonzero, \( \langle \hat{f}_i(x) \rangle = \langle e(x) \rangle \), implying that \( e(x) = \hat{e}_i(x) \) as both are the unique unity of \( \langle \hat{f}_i(x) \rangle \).

(vi) Note that \( e(x)\hat{e}_i(x) \) is an idempotent in \( \langle \hat{f}_i(x) \rangle \). Thus either \( e(x)\hat{e}_i(x) = 0 \) or \( \hat{e}_i(x) \) by (v). Let \( T = \{ i | e(x)\hat{e}_i(x) \neq 0 \} \). Then by (iv), \( e(x) = e(x) \cdot 1 = e(x) \sum_{i=1}^{s} \hat{e}_i(x) = \sum_{i=1}^{s} e(x)\hat{e}_i(x) = \sum_{i \in T} \hat{e}_i(x) \). Furthermore \( \langle e(x) \rangle = \langle \sum_{i \in T} \hat{e}_i(x) \rangle = \sum_{i \in T} \langle \hat{e}_i(x) \rangle \) by theorem 4.3.7 (ii) and induction.

\[\square\]

**Theorem 4.3.6.** Let \( M \) be a minimal ideal of \( R_n \). Then \( M \) is an extension field of \( \mathbb{F}_q \).

**Proof.** We will show that every nonzero element in \( M \) has a multiplicative invese in \( M \).

Let \( a(x) \in M \) with \( a(x) \neq 0 \). Then \( \langle a(x) \rangle \) is a nonzero ideal of \( R_n \) contained in \( M \), and hence \( \langle a(x) \rangle = M \). So \( e(x) \) is the unity of \( M \), there is an element \( b(x) \) in \( R_n \) with \( a(x)b(x) = e(x) \). Now \( c(x) = b(x)e(x) \in M \) as \( e(x) \in M \). Hence \( a(x)c(x) = e^2(x) = e(x) \).

\[\square\]

### 4.4 Zeros of a cyclic code

If \( t = \text{ord}_n(q) \), then \( \mathbb{F}_{q^t} \) is a splitting field of \( x^n - 1 \), so \( \mathbb{F}_{q^t} \) contains a primitive nth root of unity \( \alpha \), and \( x^n - 1 = \Pi_{i=0}^{n-1} (x - \alpha^i) \) is the factorization of \( x^n - 1 \) into linear factors over \( \mathbb{F}_{q^t} \). Furthermore \( x^n - 1 = \Pi_s M_{\alpha^s} \) is the factorization of \( x^n - 1 \) into irreducible factors over \( \mathbb{F}_q \), where \( s \) runs through a set of representatives of \( q- \) cyclotomic cosets modulo \( n \).

Let \( C \) be a cyclic code in \( R_n \) with generator polynomial \( g(x) \). Then \( g(x) = \Pi_s M_{\alpha^s} = \Pi_s \Pi_{i \in C_s} (x - \alpha^i) \), where \( s \) runs through some subset of representatives of \( q- \) cyclotomic
cosets $C_s$ modulo $n$.

Let $T = \cup_s C_s$ be the union of these $q-$ cyclotomic cosets. The roots of unity $\mathcal{Z} = \{\alpha^i | i \in T\}$ are called the zeros of the cyclic code $\mathcal{C}$ and $\{\alpha^i | i \notin T\}$ are nonzeros of $\mathcal{C}$. The set $T$ is called defining set of $\mathcal{C}$. It follows that $c(x)$ belongs to $\mathcal{C}$ if and only if $c(\alpha^i) = 0$ for each $i \in T$ (because $c(x) = g(x)f(x)$ where $f(x) \in \mathbb{F}_q(x)$ and $\deg f(x) < k$).

**Note** The set $T$ where $T = \cup_s C_s$ and hence either the set of zeros or the set of nonzeros, completely determines the generator polynomial $g(x)$, and since $|T| = \deg g(x)$, then $\dim \mathcal{C} = n - |T|$.

**Example 4.4.1.** In the following table we give, generator polynomial $g_i(x)$ and generating idempotents $e_i(x)$ for all the cyclic codes $\mathcal{C}_i$ of length 7 over $\mathbb{F}_2$ and we add in the table the defining sets of each code relative to the primitive root $\alpha$ given in example 4.3.4.

<table>
<thead>
<tr>
<th>$i$</th>
<th>dim</th>
<th>$g_i(x)$</th>
<th>$e_i(x)$</th>
<th>Defining set</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>$1 + x^7$</td>
<td>0</td>
<td>${0, 1, 2, 3, 4, 5, 6}$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$1 + x + x^2 + x^3 + x^4 + x^5 + x^6$</td>
<td>$1 + x + x^2 + x^3 + x^4 + x^5 + x^6$</td>
<td>${1, 2, 3, 4, 5, 6}$</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>$1 + x^2 + x^3 + x^4$</td>
<td>$1 + x^3 + x^5 + x^6$</td>
<td>${0, 1, 2, 4}$</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>$1 + x + x^2 + x^4$</td>
<td>$1 + x + x^2 + x^4$</td>
<td>${0, 3, 5, 6}$</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>$1 + x + x^3$</td>
<td>$x + x^2 + x^4$</td>
<td>${1, 2, 4}$</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>$1 + x^2 + x^3$</td>
<td>$x^3 + x^5 + x^6$</td>
<td>${3, 5, 6}$</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>$1 + x$</td>
<td>$x + x^2 + x^3 + x^4 + x^5 + x^6$</td>
<td>${0}$</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>$1$</td>
<td>1</td>
<td>$\emptyset$</td>
</tr>
</tbody>
</table>

$\alpha^0, \alpha = \alpha, \alpha^2 = \alpha^2, \alpha^3 = \alpha + 1, \alpha^4 = \alpha^2 + \alpha, \alpha^5 = \alpha^2 + \alpha + 1, \alpha^6 = \alpha^2 + 1$

**Exercise:** What would be the defining sets of each of the codes in the above example if the primitive root $\beta = \alpha^3$ were used to determine the defining set rather than $\alpha$?

**Solution:** If $\beta = \alpha^3$ is a primitive 7 th root of unity then the defining sets for the a b ove
codes is given in the following table

<table>
<thead>
<tr>
<th>i</th>
<th>dim</th>
<th>$g_i(x)$</th>
<th>Defining set</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>$1 + x^7$</td>
<td>{0, 1, 2, 3, 4, 5, 6}</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$1 + x + x^2 + x^3 + x^4 + x^5 + x^6$</td>
<td>{1, 2, 3, 4, 5, 6}</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>$1 + x^2 + x^3 + x^4$</td>
<td>{0, 3, 5, 6}</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>$1 + x + x^2 + x^4$</td>
<td>{0, 1, 2, 4}</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>$1 + x^2 + x^3$</td>
<td>{1, 2, 4}</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>$1 + x + x^3$</td>
<td>{3, 5, 6}</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>$1 + x$</td>
<td>{0}</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>$1$</td>
<td>{}</td>
</tr>
</tbody>
</table>

To explain the code $C_5$. Since the generator polynomial $g(x) = 1 + x + x^3$, then $\beta = \alpha^3$ is a primitive 7-th root of unity, then the roots of the cyclic code are $(\alpha^3)^3 = \alpha^2$, $(\alpha^3)^5 = \alpha$, $(\alpha^3)^6 = \alpha^4$, because $g(\alpha^2) = 1 + \alpha^2 + \alpha^6 = 1 + \alpha^2 + 1 + \alpha^2 = 0$,
$g(\alpha) = 1 + \alpha + \alpha^3 = 1 + \alpha + \alpha + 1 = 0$,
$g(\alpha^4) = 1 + \alpha^4 + \alpha^{12} = 1 + \alpha^2 + \alpha + \alpha^3 = 1 + \alpha^2 + \alpha + \alpha^2 + \alpha + 1 = 0$,
Or $g(x) = (x - \alpha^2)(x - \alpha)(x - \alpha^4) = 1 + x + x^3$.

**Theorem 4.4.1.** Let $\alpha$ be a primitive nth root of unity in some extension field of $\mathbb{F}_q$. Let $C$ be a cyclic code of length $n$ over $\mathcal{U}_q$ with defining set $T$ and generator polynomial $g(x)$.

The following hold.

(i) $T$ is a union of $q$-cyclotomic cosets modulo $n$.

((ii) $g(x) = \Pi_{i \in T} (x - \alpha^i)$.

((iii) $c(x) \in \mathcal{R}_n$ is in $C$ if and only if $c(\alpha^i) = 0$ for all $i \in T$.

(iv) The dimension of $C$ is $n - |T|$.

For a cyclic code $C$ in $\mathcal{R}_n$, there are in general many polynomials $r(x) \in \mathcal{R}_n$ such that $C = \langle r(x) \rangle$. However by theorem 4.2.1 and its corollary, there is exactly one such
polynomial, namely the monic polynomial in $C$ of minimal degree, which also divides $x^n - 1$ and which we call the generator polynomial of $C$. In the next theorem we characterize all polynomials $r(x)$ which generates $C$.

**Theorem 4.4.2.** Let $C$ be a cyclic code of length $n$ over $\mathbb{F}_q$ with generator polynomial $g(x)$. Let $v(x)$ be a polynomial in $\mathbb{R}_n$.

(i) $C = \langle v(x) \rangle$ if and only if $\gcd(v(x), x^n - 1) = g(x)$.

(ii) $v(x)$ generates $C$ if and only if the $n$th roots of unity which are zeros of $v(x)$ are precisely the zeros of $C$.

**Proof.** (i) Let $g(x) = \gcd(v(x), x^n - 1) \implies g(x) \mid v(x)$, so multiples of $v(x)$ are multiples of $g(x)$ in $\mathbb{R}_n$ and so $\langle v(x) \rangle \subseteq C = \langle g(x) \rangle$...(1).

By the Euclidean algorithm there exist polynomials $a(x)$ and $b(x)$ in $\mathbb{F}_q[x]$ such that $g(x) = a(x)v(x) + b(x)(x^n - 1)$. Hence $g(x) = a(x)v(x)$ in $\mathbb{R}_n$ and so multiples of $g(x)$ are multiples of $v(x)$ in $\mathbb{R}_n$ implying, $\langle v(x) \rangle \supseteq C = \langle g(x) \rangle$...(2).

So from (1) and (2) we have $C = \langle v(x) \rangle$.

Conversely, assume that $C = \langle v(x) \rangle$. Let $d(x) = \gcd(v(x), x^n - 1)$. As $g(x) \mid v(x)$ and $g(x) \mid x^n - 1 \implies g(x) \mid d(x)$. As $g(x) \in C = \langle v(x) \rangle$, so there exist $a(x)$ such that $g(x) = a(x)v(x) \in \mathbb{R}_n$. So there exists $b(x)$ such that $g(x) = a(x)v(x) + b(x)(x^n - 1)$ in $\mathbb{F}_q[x] \implies d(x) \mid g(x)$. Hence as both $d(x)$ and $g(x)$ are monic and divide each other $d(x) = g(x)$ and (i) holds.

(ii) As the only roots of both $g(x)$ and $x^n - 1$ are $n$th roots of unity, $g(x) = \gcd(v(x), x^n - 1)$ if and only if the $n$th roots of unity which are zeros of $v(x)$ are precisely the zeros of $g(x)$, the latter are the zeros of $C$. \hfill $\square$

### 4.5 Meggitt decoding of cyclic codes

There are several variations of Meggitt decoding; we will present two of them.

Let $C$ be an $[n, k, d]$ cyclic code over $\mathbb{F}_q$ with generator polynomial $g(x)$ of degree $n - k$;
$\mathcal{C}$ will correct $t = \left\lfloor \frac{d-1}{2} \right\rfloor$ errors. Let $c(x) \in \mathcal{C}$ is transmitted and $y(x) = c(x) + e(x)$ is received, where $e(x) = e_0 + e_1 x + \cdots + e_{n-1} x^{n-1}$ is the error vector with $wt(e(x)) \leq t$. The Miggitt decoder stores syndromes of error patterns with coordinate $n$—in error.

The first version of Miggitt decoding algorithm described as follows: By shifting $y(x)$ at most $n$ times, the decoder finds the error vector $e(x)$ from the list and corrects the errors.

In the second version, by shifting $y(x)$ until an error appears in coordinate $n-1$, the decoder finds the error in that coordinate, correct only that error, and then corrects errors in coordinates $n−2, n−3, \cdots, 1, 0$ in that order by further shifting.

For any vector $v(x) \in \mathbb{F}_q[x]$ let $R_{g(x)}(v(x))$ be the unique remainder when $v(x)$ is divided by $g(x)$ according to the division algorithm, that is $R_{g(x)}(v(x)) = r(x)$ where $v(x) = g(x)f(x) + r(x)$ where $r(x) = 0$ or $\deg r(x) < n-k$. The function $R_{g(x)}(v(x))$ satisfies the following properties.

**Theorem 4.5.1.** With the preceding notation the following hold:

(i) $R_{g(x)}(av(x) + bv'(x)) = aR_{g(x)}(v(x)) + bR_{g(x)}(v'(x))$ for all $v(x), v'(x) \in \mathbb{F}_q[x]$ and all $a, b \in \mathbb{F}_q$.

(ii) $R_{g(x)}(v(x) + a(x)(x^n - 1)) = R_{g(x)}(v(x))$

(iii) $R_{g(x)}(v(x)) = 0$ if and only if $v(x) \mod (x^n - 1) \in \mathcal{C}$.

(iv) If $c(x) \in \mathcal{C}$, then $R_{g(x)}(c(x) + e(x)) = R_{g(x)}(e(x))$

(v) If $R_{g(x)}(e(x)) = R_{g(x)}(e'(x))$, where $e(x)$ and $e'(x)$ each have weight at most $t$, then $e(x) = e'(x)$

(vi) $R_{g(x)}(v(x)) = v(x)$ if $\deg v(x) < n-k$.

**Proof.** (i) Let $r(x) = R_{g(x)}(v(x))$, $r'(x) = R_{g(x)}(v'(x))$ where $v(x) = g(x)f(x) + r(x)$, $v'(x) = g(x)f'(x) + r'(x)$ with $r(x) = 0$ or $\deg r(x) < n-k$, and $r'(x) = 0$ or $\deg r'(x) < n-k$.

$r(x) = v(x) - g(x)f(x)$, $r'(x) = v'(x) - g(x)f'(x)$.
ar(x) = av(x) - ag(x)f(x), \quad br'(x) = bv'(x) - bg(x)f'(x)

ar(x) + br'(x) = av(x) - ag(x)f(x) + bv'(x) - bg(x)f'(x)

av(x) + bv'(x) = ag(x)f(x) + bg(x)f'(x) + ar(x) + br'(x) = g(x)(af(x) + bf'(x) + ar(x) + br'(x)), \text{where } \operatorname{deg} ar(x) + br'(x) < n - k \text{ or } ar(x) + br'(x) = 0 \implies \mathcal{R}_{g(x)}(av(x) + bv'(x)) = ar(x) + br'(x) = a\mathcal{R}_{g(x)}(v(x)) + b\mathcal{R}_{g(x)}(v'(x)) \text{ for all } v(x), r'(x) \in \mathbb{F}_q[x] \text{ and } a, b \in \mathbb{F}_q.

(ii) \text{Let } v(x) = g(x)f(x) + r(x) \text{ where } r(x) = 0 \text{ or } \deg r(x) < n - k. \quad \mathcal{R}_{g(x)}(v(x)) = r(x) = v(x) - g(x)f(x) \text{ since } g(x) \text{ is a generator polynomial, then } g(x)|x^n - 1 \implies x^n - 1 = k(x)g(x) \implies a(x)(x^n - 1) = a(x)k(x)g(x) \implies v(x) + a(x)(x^n - 1) = g(x)f(x) + r(x) + a(x)k(x)g(x) = g(x)[f(x) + a(x)k(x)] + r(x).

r(x) = \mathcal{R}_{g(x)}(v(x)) = \mathcal{R}_{g(x)}(v(x) + a(x)(x^n - 1)).

(iii) \text{If } r(x) = \mathcal{R}_{g(x)}(v(x)) = 0 \implies v(x) = g(x)f(x) \implies v(x) + a(x)(x^n - 1) = g(x)f(x) + a(x)k(x)g(x) = g(x)(f(x) + a(x)k(x)) \in \mathcal{C} \implies v(x) \mod (x^n - 1) \in \mathcal{C}.

Conversely, if \( v(x) \mod (x^n - 1) \in \mathcal{C} \implies v(x) + a(x)(x^n - 1) \in \mathcal{C} \implies \mathcal{R}_{g(x)}(v(x)) = 0.

(iv) \text{If } c(x) \in \mathcal{C}, \text{ then } \mathcal{R}_{g(x)}(c(x)) = 0 \text{ and } \implies \mathcal{R}_{g(x)}(c(x)) + \mathcal{R}_{g(x)}(e(x)) = 0 + \mathcal{R}_{g(x)}(e(x))

(v) \text{If } \mathcal{R}_{g(x)}(e(x)) = \mathcal{R}_{g(x)}(e'(x)) \text{ then there exists } c(x) \in \mathcal{C} \text{ such that } \mathcal{R}_{g(x)}(c(x) + e(x)) = \mathcal{R}_{g(x)}(c(x) + e'(x)) = \mathcal{R}_{g(x)}(c(x) + e'(x)) - \mathcal{R}_{g(x)}(e'(x)) = \mathcal{R}_{g(x)}(e(x) - e'(x)) = 0 \implies e(x) - e'(x) \mod (x^n - 1) \in \mathcal{C} \text{ where } \deg (e(x) - e'(x) < n - k \text{ but } g(x) \text{ is a unique monic polynomial of minimal degree } n - k \implies e(x) - e'(x) = 0 \implies e(x) - e'(x).

(vi) \mathcal{R}_{g(x)}(v(x)) = v(x) \text{ if } \deg v(x) < n - k.

If \( \deg v(x) < n - k \implies v(x) = 0 \cdot g(x) + v(x) \implies \mathcal{R}_{g(x)}(v(x)) = v(x).

\[\square\]

**Theorem 4.5.2.** Let \( g(x) \) be a monic divisor of \( x^n - 1 \) of degree \( n - k \). If \( \mathcal{R}_{g(x)}(v(x)) = S(x) \), then \( \mathcal{R}_{g(x)}(xv(x)) \mod (x^n - 1) = \mathcal{R}_{g(x)}(xS(x)) = xS(x) - g(x)S_{n-k-1} \), where \( S_{n-k-1} \) is the coefficient of \( x^{n-k-1} \) in \( S(x) \).
Proof. By definition \( v(x) = g(x)f(x) + S(x) \) where \( S(x) = \sum_{i=0}^{n-k-1} S_i x^i \). So \( xv(x) = xg(x)f(x) + xs(x) = xg(x) \cdot f(x) + g(x)f_1(x) + S'(x) \), where \( S'(x) = R_{g(x)}(xS(x)) \). Also \( xv(x) \mod (x^n - 1) = xv(x) - (x^n - 1)v_{n-1} \). (Because \( xv(x) = v_0 x + \cdots + v_{n-1} x^n \implies xv(x) (\mod x^n - 1) = xv(x) - (x^n - 1) \). Thus \( xv(x) \mod (x^n - 1) = xg(x) + f(x)f_1(x) + S'(x) - (x^n - 1)v_{n-1} = (xf(x) + f_1(x) - h(x)v_{n-1})g(x) + s'(x) \), where \( g(x)h(x) = x^n \cdot 1 \). Therefore \( R_{g(x)}(xv(x)) \mod (x^n - 1) = s'(x) = R_{g(x)}(xS(x)) \), because \( \deg S'(x) < n - k = \deg g(x) \). As \( g(x) \) is monic of degree \( n - k \) and \( xS(x) \sum_{i=0}^{n-k-1} S_i x^{i+1} \), the remainder when \( xS(x) \) is divided by \( g(x) \) is \( xS(x) - g(x)S_{n-k-1} \). (because \( xS(x) = S_0 x + S_1 x^2 + \cdots + S_{n-k-1} x^{n-k} \) = \( g(x)S_{n-k-1} + R_{g(x)}(xS(x)) \implies R_{g(x)}(xS(x)) = xS(x) - g(x)S_{n-k-1} \). \( \square \)

Definition 4.5.1. The weight of a polynomial is the number of nonzero coefficients.

Definition 4.5.2. The syndrome \( S(v(x)) \) of \( v(x) \) is defined by \( S(v(x)) = R_{g(x)}(x^{n-k}v(x)) \).

Remark 4.5.1. If \( v(x) \in R_n \), then \( S(v(x)) = 0 \) if and only if \( v(x) \in C \).

The first version of Meggitt decoding algorithm

Step I: We find all the syndrome polynomials \( S(e(x)) \) of error patterns \( e(x) = \sum_{i=0}^{n-1} e_i x^i \) such that \( wt(e(x)) \leq t \) and \( e_{n-1} \neq 0 \).

Example 4.5.1. Let \( C \) be the [15, 7, 5] binary cyclic code with defining set \( T = \{1, 2, 3, 4, 6, 8, 9, 12\} \). Let \( \alpha \) be a 15th root of unity in \( \mathbb{F}_{16} \). Then \( g(x) = 1 + x^4 + x^6 + x^7 + x^8 \) is the generator polynomial of \( C \) and the syndrome of \( e(x) \) is \( S(e(x)) = R_{g(x)}(x^8e(x)) \). Step I produces the
For the following syndrome polynomials:

\[
\begin{array}{c|c}
  e(x) & S(e(x)) \\
  x^{14} & x^7 \\
x^{13} + x^{14} & x^6 + x^7 \\
x^{12} + x^{14} & x^5 + x^7 \\
x^{11} + x^{14} & x^4 + x^7 \\
x^{10} + x^{14} & x^3 + x^7 \\
x^{9} + x^{14} & x^2 + x^7 \\
x^{8} + x^{14} & x + x^7 \\
x^{7} + x^{14} & 1 + x^7 \\
x^{6} + x^{14} & x^3 + x^5 + x^6 \\
x^{5} + x^{14} & x^2 + x^4 + x^5 + x^6 + x^7 \\
x^{4} + x^{14} & x + x^3 + x^4 + x^5 + x^7 \\
x^{3} + x^{14} & 1 + x^2 + x^3 + x^4 + x^7 \\
x^{2} + x^{14} & x + x^2 + x^5 + x^6 \\
x + x^{14} & 1 + x + x^4 + x^5 + x^6 + x^7 \\
1 + x^{14} & 1 + x^4 + x^6 \\
\end{array}
\]

for example to compute \( S(x^{12} + x^{14}) = R_{g(x)}(x^8(x^{12} + x^{14})) = R_{g(x)}(x^{20} + x^{22}) = R_{g(x)}(x^5 + x^7) = x^5 + x^7 \) because \( \text{deg } x^7 + x^7 < 8 = \text{deg } g(x) \) also \( S(1 + x^{14}) = R_{g(x)}(x^8(1 + x^{14})) = R_{g(x)}(x^8 + x^7) = R_{g(x)}(x^8) + R_{g(x)}(x^7) = (1 + x^4 + x^6 + x^7) + x^7 = 1 + x^4 + x^6. \)

For \( R_{g(x)}(x^9) = R_{g(x)}(xx^8) \) by theorem 4.6.2 we have \( R_{g(x)}(x^9) = R_{g(x)}(x(1 + x^4 + x^6 + x^7)) = R_{g(x)}(x + x^5 + x^7 + x^8) = R_{g(x)}(x + x^5 + x^7 + 1 + x^4 + x^6 + x^7) = R_{g(x)}(1 + x + x^4 + x^5 + x^6 + x^7) = 1 + x + x^4 + x^5 + x^6 + x^7. \)

For \( S(x + x^{14}) = R_{g(x)}(x^8(x + x^{14}) = R_{g(x)}(x^9 + x^7) = R_{g(x)}(x^8) + R_{g(x)}(x^7) = 1 + x + x^4 + x^5 + x^6 + x^7. \)

**Step II:** Suppose that \( y(x) \) is the received vector, compute the syndrome \( S(y(x)) = R_{g(x)}(x^{n-k}y(x)) \), since \( y(x) = x(x) + e(x) \), with \( e(x) \in C \implies S(y(x)) = S(e(x)) + S(e(x)) = 0 + s(e(x)). \)
Example 4.5.2. Let \( y(x) = 1 + x^4 + x^7 + x^9 + x^{10} + x^{12} \) is received. Then \( S(y(x)) = x + x^2 + x^6 + x^7 \).

Step III: If \( S(y(x)) \) is in the list computed in step I, then you know the error polynomial \( e(x) \) and this can be subtracted from \( y(x) \) to obtain the codeword \( c(x) \). If \( S(y(x)) \) is not in the list go on step (iv).

Step iv: Compute the syndrome polynomial of \( xy(x), x^2y(x), \cdots \) in succession until the syndrome polynomial is in the list from step I. If \( S(xy(x)) \) is in the list and is associated with the error polynomial \( e'(x) \), then the received vector is decoded as \( y(x) - x^{n-k}e'(x) \).

The computation in step (iv) is most easily carried out using theorem 4.6.12 as

\[
R_g(x) = S(y(x)) = \sum_{i=0}^{n-k-1} S_ix^i,
\]

\[
S(xy(x)) = R_g(x)(x^n - k) = R_g(x)(xS(y(x))) = xS(y(x)) - S_{n-k-1}g(x) \ldots \ (1)
\]

We proceed in the same fashion to get the syndrome of \( x^2y(x) \) from that of \( xy(x) \).

Example 4.5.3. From example 4.5.2

\( S(y(x)) = x + x^2 + x^6 + x^7 \), that \( S(xy(x)) = x(x^2 + x^6 + x^7) - 1 \cdot g(x) = x^2 + x^3 + x^7 + x^8 - (1 + x^4 + x^6 + x^7 + x^8) = 1 + x^2 + x^3 + x^4 + x^6 \) which is not in the list, from equation (1), we have,

\[
S(x^2y(x)) = x \cdot S(xy(x)) - 0 \cdot g(x)
\]

\[
S(x^2y(x)) = x(1 + x^2 + x^3 + x^4 + x^6) - 0 = x + x^3 + x^4 + x^5 + x^7, \text{ which is in the list, which corresponds to the error } x^4 + x^{14} \implies y(x) \text{ is decoded as}
\]

\[
y(x) - (x^{15-2})(x^4 + x^{14}) = y(x) - (x^{17} + x^{27}) = y(x) - (x^2 + x^{15}) = 1 + x^4 + x^7 + x^9 + x^10 + x^{12} - (x^2 + x^{12}) = 1 + x^2 + x^4 + x^7 + x^9 + x^{10}.
\]

Note: this codeword is \((1 + x^2)g(x)\).