Coding Theory

version: February 6, 2024

Alberto Ravagnani

Eindhoven University of Technology

Contents

Notes and Acknowledgement **3**

List of Symbols **4**

0 Introduction **5**

1 Channels and Codes **6**
1.1 Communication Channels 6
1.2 Codes and Decoders 8
1.3 The q-ary Symmetric Channel 11
1.4 Other Exercises 15

2 Codes with the Hamming Metric **16**
2.1 Definitions and First Examples 16
2.2 The Gilbert-Varshamov Bound 17
2.3 Linear Codes and Their Defining Matrices 18
2.4 Syndrome Decoding 22
2.5 Weight Distribution and Its Significance 24
2.6 New Codes from Old 26
2.7 The Dual Code 29
2.8 Equivalence of Linear Codes 31
2.9 Information Sets 33
2.10 Other Exercises 34

3 Bounds **38**
3.1 The Singleton Bound and MDS Codes 38
3.2 The Hamming Bound and Perfect Codes 40
3.3 The Griesmer Bound 43
3.4 The Plotkin Bound 46
3.5 Other Exercises 48

4 Reed-Solomon and Goppa Codes **51**
4.1 Reed-Solomon Codes 51
4.2 The Berlekamp-Welch Algorithm 52
4.3 Goppa Codes 55
4.4 Other Exercises 56

5 Duality Theory **57**
5.1 Preliminary Results 57
5.2 The MacWilliams Identities 59
5.3 Computation of Some Weight Distributions 60
5.4 Duality and MDS Codes 61
5.5 Other Exercises 63

6 Reed-Muller Codes **64**
6.1 Definition and First Properties 64
6.2 Structure of Reed-Muller Codes 66
6.3 The Dual of a Reed-Muller Code 67
6.4 Other Exercises 68

7 Distributed Storage and Locality **69**
7.1 Storage Strategies 69
7.2 Locality 70
7.3 Bounds for Codes with Locality 72
7.4 The Tamo-Barg Construction 73
7.5 Other Exercises 76

8 Code-Based Cryptography **77**
8.1 The McEliece Cryptosystem 77
8.2 A Note on Attack Strategies 78
8.3 Information Set Decoding 78
8.4 Other Exercises 81

9 Network Coding **83**
9.1 Recombining Messages 83
9.2 Multicast Networks and the Edge-Cut Bound 85
9.3 Communication Schemes 87
9.4 The Max-Flow-Min-Cut Theorem 89

A Finite Fields **93**

B Complexity Essentials **97**

C Solutions to Some of the Exercises **99**

Notes and Acknowledgement

1. These are the lecture notes for the course "Coding Theory" (code: 2MMC30) taught at Eindhoven University of Technology. Please keep in mind the following:

   - These notes do not substitute the lectures. Coming to class regularly is important to understand the material, ask questions, and figure out what the most relevant concepts are. Please come to class if you are taking this course.
   - These notes grow and evolve as the course progresses: please make sure you always have the latest version.
   - Do not hesitate to contact me or to ask questions in class if some parts of these notes are not clear to you. Your feedback helps me to improve the material.

2. The notes are organized into chapters covering different topics. Each chapter is divided into sections. Some important exercises are spread out over the various sections. Each chapter has a final section with more exercises.

3. I am grateful to Ruud Pellikaan for the material and exercises he generously shared with me when I became responsible for this course. The book he recently co-authored [6] is an excellent reference for some sections of this course.

4. Nice coding theory references, by which these notes are partly inspired, are listed at the end of these notes.

List of Symbols

$\mathbb{N}$	The natural numbers (with zero)
$\mathbb{Z}$	The integers
$\mathbb{Q}$	The rationals
$\mathbb{R}$	The reals
$\mathbb{C}$	The complex numbers
$\mathbb{F}_q$	The finite field with q elements, with q a prime power
$[n]$	The set $\{1, ..., n\}$, for $n \in \mathbb{N}$
S^c	The complement of a set S with respect to an ambient set, usually $\{1, ..., n\}$
$\binom{a}{b}$	The binomial coefficient of a and b
$\mathrm{rowsp}(M)$	The rowspace of the matrix M
$\mathrm{colsp}(M)$	The columnspace of the matrix M
$M^\top$	The transpose of the matrix M
$\mathrm{RREF}(M)$	The reduced row echelon form of the matrix M
d^{H}	The Hamming distance
ω^{H}	The Hamming weight
σ^{H}	The Hamming support
$\mathbb{F}_q[X]_{<s}$	The space of univariate polynomials over $\mathbb{F}_q$ of degree strictly smaller than s
$\mathbb{F}_2[X_1, ..., X_m]^{\times}_{\leq s}$	The space of multivariate square-free polynomials over $\mathbb{F}_2$ of degree at most s

Chapter 0

Introduction

See the slides (separate file) for an introduction to the main ideas behind coding theory and its real-world applications.

Chapter 1

Channels and Codes

In 1948, Claude Shannon had the brilliant idea of formalizing the principles of communication using mathematics. His seminal paper [8] marks the birth of information theory. Shannon's intuition consists in the fact that a communication channel is fully described by a collection of probabilities, which specify how often a symbol y turns into a different symbol x during transmission.

1.1 Communication Channels

We start by formally defining discrete communication channels.

Definition 1.1. A (**communication**) **channel** is a triple $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$, where $\mathscr{X}$ and $\mathscr{Y}$ are finite non-empty sets (which we call **input** and **output alphabet** respectively) and $\mathbb{P} : \mathscr{Y} \times \mathscr{X} \to \mathbb{R}$ is a function that satisfies the following properties:

1. $0 \leq \mathbb{P}(y, x) \leq 1$ for all $y \in \mathscr{Y}$ and $x \in \mathscr{X}$,
2. $\sum_{y \in \mathscr{Y}} \mathbb{P}(y, x) = 1$ for all $x \in \mathscr{X}$.

The elements of $\mathscr{X}$ are called **input symbols** and those of $\mathscr{Y}$ **output symbols**. Input and output symbols are sometimes called **messages**.

One often writes $\mathbb{P}(y \mid x)$ instead of $\mathbb{P}(y, x)$, and reads it "$\mathbb{P}$ of y, given x". The number $\mathbb{P}(y \mid x)$ is to be interpreted as the probability that y is received, if x was transmitted.

Remark 1.2. A communication channel $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$ gives rise to a collection of probability spaces as follows. Fix an input symbol $x \in \mathscr{X}$, and let $\mathbb{P}_x : 2^{\mathscr{Y}} \to \mathbb{R}$ be the function defined by $\mathbb{P}_x(\mathscr{Y}') := \sum_{y \in \mathscr{Y}'} \mathbb{P}(y \mid x)$ for all $\mathscr{Y}' \subseteq \mathscr{Y}$. Then the triple $(\mathscr{Y}, 2^{\mathscr{Y}}, \mathbb{P}_x)$ is a probability space. For $\mathscr{Y}' \in 2^{\mathscr{Y}}$, the number $\mathbb{P}_x(\mathscr{Y}')$ is interpreted as the probability that some symbol from $\mathscr{Y}'$ is received when x is transmitted.

A fundamental example of channel is the following. It can be constructed over any alphabet $\mathscr{X}$ of cardinality $2 \leq |\mathscr{X}| < +\infty$, but in these notes we directly focus on finite fields for convenience (see Appendix A).

Definition 1.3 (q-ary symmetric channel)**.** Let $\mathbb{F}_q$ be the finite field with q elements and let $0 \leq \alpha \leq 1$ be a real number. The **q-ary symmetric channel** with **error probability** α is the triple $\mathrm{Sym}(q, \alpha) := (\mathbb{F}_q, \mathbb{F}_q, \mathbb{P})$, where the function $\mathbb{P} : \mathbb{F}_q \times \mathbb{F}_q \to \mathbb{R}$ is defined by

$$\mathbb{P}(y \mid x) = \begin{cases} 1-\alpha & \text{if } x = y, \\ \alpha/(q-1) & \text{if } x \neq y, \end{cases}$$

for all $y, x \in \mathbb{F}_q$. When $q = 2$ we call this the **binary symmetric channel**.

A second important example is the following.

Definition 1.4 (erasure channel)**.** Let $\mathscr{X}$ be a finite non-empty set and let $0 \leq \alpha \leq 1$ be a real number. Take $? \notin \mathscr{X}$ and let $\mathscr{Y} := \mathscr{X} \cup \{?\}$. Finally, let $\mathbb{P} : \mathscr{Y} \times \mathscr{X} \to \mathbb{R}$ be defined as

$$\mathbb{P}(y \mid x) = \begin{cases} 1-\alpha & \text{if } x = y, \\ \alpha & \text{if } y = ?, \\ 0 & \text{otherwise,} \end{cases}$$

for all $y \in \mathscr{Y}$ and $x \in \mathscr{X}$. Then $\mathrm{Er}(\mathscr{X}, \alpha) := (\mathscr{X}, \mathscr{Y}, \mathbb{P})$ is called an **erasure channel** with **erasure probability** α over the alphabet $\mathscr{X}$. The question mark ? is called **erasure symbol** in this context.

Exercise 1.5. Show that $\mathrm{Sym}(q, \alpha)$ and $\mathrm{Er}(\mathscr{X}, \alpha)$ are indeed communication channels, according to Definition 1.1.

In digital communications, a channel $(\mathscr{X}, \mathscr{Y}, \mathbb{P})$ is typically used multiple times, say n, one after another. In practice what we transmit are therefore "words" of given length, say n, made of "letters" from the input alphabet $\mathscr{X}$. Similarly, we receive "words" of length n made of "letters" from the output alphabet $\mathscr{Y}$. This is modeled by the product channel.

Definition 1.6. Let $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$ be a channel and let $n \geq 1$ be an integer. The **n-th product** of $(\mathscr{X}, \mathscr{Y}, \mathbb{P})$ is the channel $\mathscr{K}^n := (\mathscr{X}^n, \mathscr{Y}^n, \mathbb{P}^n)$, where $\mathbb{P}^n$ is defined by

$$\mathbb{P}^n(y \mid x) := \prod_{i=1}^{n} \mathbb{P}(y_i \mid x_i),$$

for all $y \in \mathscr{Y}^n$ and $x \in \mathscr{X}^n$.

The product construction models what information theorists call a *discrete memoryless channel*. "Memoryless" means that what happens in a channel use is indepent from what happens in the other channel uses. This behaviour is captured by the *product* of the $\mathbb{P}(y_i \mid x_i)$'s. In these notes we only treat memoryless channels.

Proposition 1.7. If $\mathscr{K}$ is a channel and $n \geq 1$, then $\mathscr{K}^n$ is a channel.

Proof. Let $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$. The function $\mathbb{P}^n$ clearly satisfies Property 1 of Definition 1.1. To see that Property 2 holds we compute

$$\begin{aligned}\sum_{y \in \mathscr{Y}^n} \mathbb{P}^n(y \mid x) &= \sum_{(y_1,\ldots,y_n) \in \mathscr{Y}^n} \mathbb{P}(y_1 \mid x_1) \cdots \mathbb{P}(y_n \mid x_n) \\ &= \sum_{y_1 \in \mathscr{Y}} \mathbb{P}(y_1 \mid x_1) \cdots \sum_{y_n \in \mathscr{Y}} \mathbb{P}(y_n \mid x_n) \;=\; 1 \cdots 1 = 1.\end{aligned}$$

We conclude that the triple $\mathscr{K}^n = (\mathscr{X}^n, \mathscr{Y}^n, \mathbb{P}^n)$ is a channel. □

We are particularly interested in the n-th product of the q-ary symmetric channel and of the erasure channel with input alphabet $\mathscr{X} = \mathbb{F}_q$; see Definitions 1.3 and 1.4. We denote these by $\mathrm{Sym}^n(q, \alpha)$ and $\mathrm{Er}^n(q, \alpha)$ respectively. The transmitted messages in those situations are vectors of length n with entries in $\mathbb{F}_q$, i.e., elements of the vector space $\mathbb{F}_q^n$.

1.2 Codes and Decoders

In this section we explain the main idea behind error correction in communications. In the sequel, $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$ denotes an arbitrary channel, unless otherwise stated.

In digital communications what happens is the following: source and destination agree on a set of admissible messages $\mathscr{C} \subseteq \mathscr{X}$ that can be transmitted over the channel. In this context, $\mathscr{C}$ is called an (**error-correcting**) **code**. The reason why it makes sense to restrict the choice of admissible messages from $\mathscr{X}$ to $\mathscr{C}$ will be explained later. A message $x \in \mathscr{C}$ is transmitted, and a message $y \in \mathscr{Y}$ is received. The function $\mathbb{P}$ is modeling a probabilistic noise affecting the channel, telling us the probability that y is received, when x is sent. The receiver tries to guess the transmitted message x from the received one, y. This guessing process is called **decoding**. In some applications, decoding needs to be modeled with quite sophisticated probabilistic objects, whereas the following definition is good enough for us.

Definition 1.8. Let $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$ be a channel and let $\mathscr{C} \subseteq \mathscr{X}$ be non-empty. A **decoder** for $\mathscr{C}$ is a function $D : \mathscr{Y} \to \mathscr{C} \cup \{\boldsymbol{f}\}$, where $\boldsymbol{f} \notin \mathscr{X}$ is called a **failure message**.

The decoder D is modeling the attempt of the receiver to guess the transmitted message $x \in \mathscr{C}$ from y, the received one. We say that "decoding is successful" when $D(y) = x$, i.e., when the transmitted message is guessed correctly. The failure message $\boldsymbol{f}$ is modeling the situation where the receiver is not confident in making a guess.

Example 1.9. Let $\mathrm{Sym}^4(2,\alpha) = (\mathbb{F}_2^4, \mathbb{F}_2^4, \mathbb{P}^4)$ be fourth product of the binary symmetric channel with probability $\alpha < 1/2$ (in practice, α is much smaller than $1/2$). Let

$$\mathscr{C} = \{(0,0,0,0), (1,1,1,1), (1,1,0,0)\} \subseteq \mathbb{F}_2^4.$$

Suppose that $y = (0,0,1,1)$ is received, and that we want to guess the transmitted message. A very natural way of doing this is looking for the messages $x \in \mathscr{C}$ that maximize the probability

$$\mathbb{P}^4(y \mid x),$$

as these are the most likely to have been sent. The latter fact relies on the (standard) assumption that all elements of $\mathscr{C}$ have exactly the same probability to be sent. Under this assumption, by Bayes theorem we have

$$\mathrm{Prob}(x \text{ sent} \mid y \text{ received}) = \mathbb{P}(y \mid x)\, \frac{\mathrm{Prob}(x \text{ sent})}{\mathrm{Prob}(y \text{ received})} = \mathbb{P}(y \mid x)\, \frac{1}{|\mathscr{C}| \cdot \mathrm{Prob}(y \text{ received})}.$$

Therefore, for a *given* received message y, the maxima of the sets

$$\{\mathbb{P}(y \mid x) \mid x \in \mathscr{C}\}, \quad \{\mathrm{Prob}(x \text{ sent} \mid y \text{ received}) \mid x \in \mathscr{C}\}$$

are attained by the same elements $x \in \mathscr{C}$. In other words, maximizing the probability $\mathrm{Prob}(x \text{ sent} \mid y \text{ received})$ is the same as maximizing $\mathbb{P}(y \mid x)$.

In our example, $\mathscr{C}$ has cardinality three and the computations are easy:

$$\begin{aligned} \mathbb{P}^4(y \mid (0,0,0,0)) &= (1-\alpha)^2\alpha^2, \\ \mathbb{P}^4(y \mid (1,1,1,1)) &= \alpha^2(1-\alpha)^2, \\ \mathbb{P}^4(y \mid (1,1,0,0)) &= \alpha^4. \end{aligned}$$

Since $\alpha < 1/2$ by assumption, the two elements $(0,0,0,0)$ and $(1,1,1,1)$ of $\mathscr{C}$ are the most probable, but they are *equally* probable. In such a situation we are undecided and prefer to return the failure message $\boldsymbol{f}$.

The probabilistic decoding strategy illustrated in the previous example applies to general channels.

Definition 1.10 (maximum likelihood decoding)**.** Let $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$ be a channel and let $\mathscr{C} \subseteq \mathscr{X}$ be non-empty. The **maximum likelihood decoder** $D : \mathscr{Y} \to \mathscr{C} \cup \{\boldsymbol{f}\}$ for $\mathscr{C}$ is defined by

$$D(y) := \begin{cases} x & \text{if } x \text{ is the unique element of } \mathscr{C} \text{ that maximizes } \mathbb{P}(y \mid x), \\ \boldsymbol{f} & \text{otherwise.} \end{cases}$$

Now suppose that we are transmitting over a channel $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$, and that we agreed on a non-empty set of admissible messages $\mathscr{C} \subseteq \mathscr{X}$ and on a decoder $\mathscr{D} : \mathscr{Y} \to \mathscr{C} \cup \{\boldsymbol{f}\}$ for $\mathscr{C}$. A message $x \in \mathscr{C}$ is transmitted, and $y \in \mathscr{Y}$ is received. The following

are the possible scenarios.

1. *Successful decoding*: $D(y) = x$. This is what we are aiming for.
2. *Decoding failure*: $D(y) = \boldsymbol{f}$, i.e., decoding returns a failure message. That's something we want to avoid as much as possible.
3. *Decoding error*: $D(y) \notin \{x, \boldsymbol{f}\}$, i.e., decoding returns the wrong message. That's a complete disaster and makes communication highly unreliable.

In scenarios 2 and 3 we say that "decoding was unsuccessful".

Remark 1.11. The whole point of coding theory is to minimize the chance to be in the second and third scenarios in the above list, without reducing too much the cardinality of the code $\mathscr{C}$. Given a channel $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$, this goal is achieved by carefully designing $\mathscr{C} \subseteq \mathscr{X}$ and the decoder D.

On the one hand, restricting the choice of messages to a proper subset $\mathscr{C} \subseteq \mathscr{X}$ can reduce the chance of unsuccessful decoding, as we will see shortly. On the other hand, the smaller $|\mathscr{C}|$ is, the poorer our language is. For example, when $|\mathscr{C}| = 2$ our language is binary, i.e., we can only transmit information of the form "yes" or "no". If $|\mathscr{C}| = 3$, we can say "yes", "no", or "maybe". That's slightly better but still very limiting. The code used in the *Mariner 9* mission in 1971 (a *Reed-Muller* code) had cardinality 64, and could only transmit low-quality pictures of Mars. Twenty-six years later, one of the codes used in the *Mars Pathfinder* mission (a *Reed-Solomon* code) had cardinality close to 10^{400}.

The set $\mathscr{C}$ plays a crucial role in error correction, as most problems in coding theory reduce to identifying a good selection of admissible messages. Constructing $\mathscr{C}$ is essentially what these notes are about. We therefore repeat the definition of code more formally.

Definition 1.12. An (**error-correcting**) **code** for a channel $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$ is a non-empty subset $\mathscr{C} \subseteq \mathscr{X}$. Its elements are called **codewords**.

Since unsuccessful decoding is what we want to avoid, a good idea is to compute the probability that this happens. Fix a channel $\mathscr{K} = (\mathscr{X}, \mathscr{Y}, \mathbb{P})$, a code $\mathscr{C} \subseteq \mathscr{X}$, and a decoder $D : \mathscr{Y} \to \mathscr{C} \cup \{\boldsymbol{f}\}$. For a given transmitted codeword $x \in \mathscr{C}$, what is the probability the decoding is unsuccessful? As $\mathbb{P}(y|x)$ measures the probability that y is received when x is transmitted, the decoder D returns $\widehat{x} \neq x$ or $\boldsymbol{f}$ with probability

$$\sum_{\substack{y \in \mathscr{Y} \\ D(y) \neq x}} \mathbb{P}(y \mid x).$$

Assuming that all codewords are equally likely to be transmitted (standard assumption in information theory), the **probability of unsuccessful decoding** is therefore given by the average

$$\mathrm{PUD}(\mathscr{K}, \mathscr{C}, D) := \frac{1}{|\mathscr{C}|} \sum_{x \in \mathscr{C}} \sum_{\substack{y \in \mathscr{Y} \\ D(y) \neq x}} \mathbb{P}(y \mid x). \tag{1.1}$$

Note that if $|\mathscr{C}| = 1$ then $\mathrm{PUD}(\mathscr{K}, \mathscr{C}, D) = 0$, i.e., we never make a mistake when decoding. However, a code of cardinality one is completely useless. Observe moreover that, by Property 2 of Definition 1.1, we have

$$\mathrm{PUD}(\mathscr{K}, \mathscr{C}, D) := 1 - \frac{1}{|\mathscr{C}|} \sum_{x \in \mathscr{C}} \sum_{\substack{y \in \mathscr{Y} \\ D(y) = x}} \mathbb{P}(y \mid x). \tag{1.2}$$

1.3 The q-ary Symmetric Channel

In this section we discuss how the objectives stated in Remark 1.11 can be achieved for the n-th product of the q-ary symmetric channel. We only consider codes endowed with a maximum likelihood decoder (Definition 1.10). This situation is highly relevant for applications.

Notation 1.13. In the sequel $n \geq 1$ is an integer and $(\mathbb{F}_q^n, \mathbb{F}_q^n, \mathbb{P}^n) = \mathrm{Sym}^n(q, \alpha)$ is the n-th product of the q-ary symmetric channel with error probability $0 \leq \alpha \leq 1$.

We want to analyze the PUD for different codes $\mathscr{C} \subseteq \mathbb{F}_q^n$ endowed with the maximum likelihood decoder, in order to get a sense of how changing the code affects the PUD. We start by giving a more explicit formula for the $\mathbb{P}(y \mid x)$'s. These can be conveniently expressed in terms of the Hamming distance between x and y.

Definition 1.14. The **Hamming distance** between vectors $x, y \in \mathbb{F}_q^n$ is the integer $d^{\mathrm{H}}(x, y) := |\{1 \leq i \leq n \mid x_i \neq y_i\}|$.

The Hamming distance indeed satisfies the properties of a distance function, in the following precise sense.

Proposition 1.15. The Hamming distance $d^{\mathrm{H}} : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{R}$ is a distance function on $\mathbb{F}_q^n$, i.e., it satisfies the following properties:

1. for all $x, y \in \mathbb{F}_q^n$ we have $d^{\mathrm{H}}(x, y) \geq 0$, and equaliy holds if and only if $x = y$;
2. for all $x, y \in \mathbb{F}_q^n$ we have $d^{\mathrm{H}}(x, y) = d^{\mathrm{H}}(y, x)$;
3. for all $x, y, z \in \mathbb{F}_q^n$ we have $d^{\mathrm{H}}(x, y) \leq d^{\mathrm{H}}(x, z) + d^{\mathrm{H}}(z, y)$.

Proof. Exercise (the solution can be found in Appendix C). □

Remark 1.16. The Hamming distance can be defined over any non-empty set $\mathscr{X}$ (which is not necessarily a finite field $\mathbb{F}_q$). For $x, y \in \mathscr{X}^n$, let $d^{\mathrm{H}}(x, y) := |\{1 \leq i \leq n \mid x_i \neq y_i\}|$. The function d^{H} satisfies all the properties of Proposition 1.15.

Next, we show the connection between the Hamming distance and the q-ary symmetric channel.

Proposition 1.17. Let $\mathrm{Sym}^n(q,\alpha) = (\mathbb{F}_q^n, \mathbb{F}_q^n, \mathbb{P}^n)$ be the n-th product of the q-ary symmetric channel with error probability $0 \leq \alpha \leq 1$. For all $x, y \in \mathbb{F}_q^n$ we have

$$\mathbb{P}^n(y \mid x) = \left(\frac{\alpha}{q-1}\right)^d (1-\alpha)^{n-d} = \left(\frac{\alpha}{(1-\alpha)(q-1)}\right)^d (1-\alpha)^n,$$

where $d = d^{\mathrm{H}}(x,y)$.

Proof. Let $S := \{1 \leq i \leq n \mid y_i \neq x_i\} \subseteq [n]$. Then $|S| = d$ and $|S^c| = n - d$. Therefore by definition of $\mathrm{Sym}^n(q,\alpha)$ we have

$$\mathbb{P}^n(y \mid x) = \prod_{i=1}^n \mathbb{P}(y_i \mid x_i) = \left(\prod_{i \in S} \frac{\alpha}{q-1}\right)\left(\prod_{i \in S^c} 1-\alpha\right) = \left(\frac{\alpha}{q-1}\right)^d (1-\alpha)^{n-d},$$

which is the desired expression. □

Another natural way to decode over the q-ary symmetric channel is to look for the message $x \in \mathscr{C} \subseteq \mathbb{F}_q^n$ that is closer to the received one, y, with respect to the Hamming distance.

Definition 1.18. Let $\mathrm{Sym}^n(q,\alpha) = (\mathbb{F}_q^n, \mathbb{F}_q^n, \mathbb{P}^n)$ be the n-th product of the q-ary symmetric channel with error probability $0 \leq \alpha \leq 1$. Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code. The **minimum-distance decoder** $D : \mathbb{F}_q^n \to \mathscr{C} \cup \{\boldsymbol{f}\}$ for $\mathscr{C}$ is defined by

$$D(y) := \begin{cases} x & \text{if } x \text{ is the unique element of } \mathscr{C} \text{ that minimizes } d^{\mathrm{H}}(y,x), \\ \boldsymbol{f} & \text{otherwise.} \end{cases}$$

Under very reasonable assumptions from an applied viewpoint, maximum likelihood decoding is the same as minimum-distance decoding.

Proposition 1.19. Consider the n-th product $\mathrm{Sym}^n(q,\alpha) = (\mathbb{F}_q^n, \mathbb{F}_q^n, \mathbb{P}^n)$ of the q-ary symmetric channel with error probability $0 \leq \alpha \leq 1$. Assume $\alpha < (q-1)/q$ and let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code. Let D and D' be the maximum likelihood and the minimum distance decoders for $\mathscr{C}$, respectively. For all $y \in \mathscr{Y}$ we have

$$D(y) = D'(y).$$

Proof. It is easy to see that $\alpha < (q-1)/q$ implies $\alpha < (1-\alpha)(q-1)$. Therefore the result follows from Proposition 1.17. □

We familiarize ourselves with the q-ary symmetric channel and its minimum distance decoder in the next example. We will repeatedly need the following elementary combinatorial result.

Lemma 1.20. Let $x \in \mathbb{F}_q^n$ and let $0 \leq i \leq n$ be an integer. The number of vectors $y \in \mathbb{F}_q^n$ with $d^{\mathrm{H}}(y, x) = i$ is

$$\binom{n}{i}(q-1)^i.$$

Proof. Exercise (the solution can be found in Appendix C). □

Example 1.21. Let $q = 3$, $n = 3$, and let $\mathscr{K} = \mathrm{Sym}^3(3, \alpha) = (\mathbb{F}_3^3, \mathbb{F}_3^3, \mathbb{P}^3)$ be the ternary symmetric channel with error probability α. Assume $\alpha < 2/3$ (in practice, α is much smaller). Consider the following two codes in $\mathbb{F}_3^3$:

$$\mathscr{C}_1 = \{0\}, \qquad \mathscr{C}_2 = \mathbb{F}_3^3.$$

Let D_1 and D_2 be the minimum distance decoders for $\mathscr{C}_1$ and $\mathscr{C}_2$ respectively. We want to compute $\mathrm{PUD}(\mathscr{K}, \mathscr{C}_1, D_1)$ and $\mathrm{PUD}(\mathscr{K}, \mathscr{C}_2, D_2)$.

The code $\mathscr{C}_1$ has only one codeword, namely $0 = (0, 0, 0)$. Therefore all vectors $y \in \mathbb{F}_3^3$ decode to 0 and it is to be expected that the PUD is zero. Let us check this using the definition in (1.2). We have

$$\begin{aligned} \mathrm{PUD}(\mathscr{K}, \mathscr{C}_1, D_1) &= 1 - \frac{1}{|\mathscr{C}_1|} \sum_{x \in \mathscr{C}_1} \sum_{\substack{y \in \mathbb{F}_3^3 \\ D_1(y) = x}} \mathbb{P}^3(y \mid x) \\ &= 1 - \sum_{y \in \mathbb{F}_3^3} \left(\frac{\alpha}{2}\right)^{d^{\mathrm{H}}(y,0)} (1-\alpha)^{3 - d^{\mathrm{H}}(y,0)}, \end{aligned}$$

where the last equality follows from Proposition 1.17. By Lemma 1.20 we then have

$$\begin{aligned} \mathrm{PUD}(\mathscr{K}, \mathscr{C}_1, D_1) &= 1 - \sum_{d=0}^{3} \sum_{\substack{y \in \mathbb{F}_3^3 \\ d^{\mathrm{H}}(y,0) = d}} \left(\frac{\alpha}{2}\right)^d (1-\alpha)^{n-d} \\ &= 1 - \sum_{d=0}^{3} 2^d \binom{3}{d} \left(\frac{\alpha}{2}\right)^d (1-\alpha)^{3-d} \\ &= 1 - \sum_{d=0}^{3} \binom{3}{d} \alpha^d (1-\alpha)^{3-d} = 0, \end{aligned}$$

where the latter equality follows from the Binomial Theorem. While the probability of unsuccessful decoding is zero, the code $\mathscr{C}_1$ has cardinality 1 and is therefore completely useless from a communication viewpoint.

The code $\mathscr{C}_2$ has instead 3^3 codewords, and for every $x \in \mathscr{C}_2$ and $y \in \mathbb{F}_3^3$ we have that y decodes to x if and only if $y = x$. Thus

$$\mathrm{PUD}(\mathscr{K}, \mathscr{C}_2, D_2) = 1 - \frac{1}{3^3} \sum_{x \in \mathscr{C}_2} \mathbb{P}^3(x \mid x) = 1 - \frac{1}{3^3} \sum_{x \in \mathbb{F}_3^3} (1-\alpha)^3 = 1 - (1-\alpha)^3.$$

Exercise **1.22.** Let $q = 3$, $n = 3$, and let $\mathscr{K} = \mathrm{Sym}^3(3, \alpha) = (\mathbb{F}_3^3, \mathbb{F}_3^3, \mathbb{P}^3)$ be the ternary symmetric channel with error probability α. Assume $\alpha < 2/3$. Let

$$\mathscr{C} := \{x \in \mathbb{F}_3^3 \mid x_1 + x_2 = 0,\ x_2 + x_3 = 0\}$$

and let D be the minimum distance decoder for $\mathscr{C}$. Compute $\mathrm{PUD}(\mathscr{K}, \mathscr{C}, D)$.

Closely related to the PUD is the following fundamental parameter of an error-correcting code $\mathscr{C} \subseteq \mathbb{F}_q^n$.

Definition 1.23. The **minimum** (**Hamming**) **distance** of a code $\mathscr{C} \subseteq \mathbb{F}_q^n$ with $|\mathscr{C}| \geq 2$ is the positive integer

$$d^{\mathrm{H}}(\mathscr{C}) := \min\{d^{\mathrm{H}}(x, y) \mid x, y \in \mathscr{C},\ x \neq y\}.$$

We let $n + 1$ be the minimum distance of any code $\mathscr{C} \subseteq \mathbb{F}_q^n$ with cardinality $|\mathscr{C}| = 1$.

A lower bound for the minimum distance of a code $\mathscr{C} \subseteq \mathbb{F}_q^n$ yields an upper bound for the corresponding PUD.

Proposition 1.24. Consider the n-th product $\mathscr{K} = \mathrm{Sym}^n(q, \alpha)$ of the q-ary symmetric channel with error probability $0 \leq \alpha \leq 1$. Assume $\alpha < (q-1)/q$. Let d be a positive integer and let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a non-zero code with $d^{\mathrm{H}}(\mathscr{C}) \geq d$. Let D be the maximum likelihood decoder for $\mathscr{C}$. We have

$$\mathrm{PUD}(\mathscr{K}, \mathscr{C}, D) \leq 1 - (1-\alpha)^n \sum_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i} \left(\frac{\alpha}{1-\alpha}\right)^i.$$

Proof. Write $\mathrm{Sym}^n(q, \alpha) = (\mathbb{F}_q^n, \mathbb{F}_q^n, \mathbb{P}^n)$. By Proposition 1.19, we shall assume that D is the minimum distance decoder. Let $t = \lfloor (d-1)/2 \rfloor$. Note that for $x \in \mathscr{C}$ and $y \in \mathbb{F}_q^n$ we have $D(y) = x$ whenever $d^{\mathrm{H}}(y, x) \leq t$ (explain why). Therefore by Proposition 1.17 we deduce

$$\begin{aligned}
\sum_{\substack{y \in \mathbb{F}_q^n \\ D(y) = x}} \mathbb{P}^n(y \mid x) &\geq \sum_{\substack{y \in \mathbb{F}_q^n \\ d^{\mathrm{H}}(y,x) \leq t}} \mathbb{P}^n(y \mid x) \\
&= \sum_{i=0}^{t} \sum_{\substack{y \in \mathbb{F}_q^n \\ d^{\mathrm{H}}(y,x) = i}} \mathbb{P}^n(y \mid x) \\
&= \sum_{i=0}^{t} \sum_{\substack{y \in \mathbb{F}_q^n \\ d^{\mathrm{H}}(y,x) = i}} \left(\frac{\alpha}{(1-\alpha)(q-1)}\right)^i (1-\alpha)^n.
\end{aligned}$$

Combining this with Lemma 1.20 we obtain

$$\sum_{\substack{y\in\mathbb{F}_q^n\\ D(y)=x}} \mathbb{P}^n(y \mid x) \geq (1-\alpha)^n \sum_{i=0}^{t} \binom{n}{i} \left(\frac{\alpha}{1-\alpha}\right)^i.$$

Finally, by Eq. (1.2) we conclude

$$\mathrm{PUD}(\mathscr{K},\mathscr{C},D) \leq 1-(1-\alpha)^n \sum_{i=0}^{t} \binom{n}{i} \left(\frac{\alpha}{1-\alpha}\right)^i. \qquad \square$$

Proposition 1.24 shows that if the minimum distance of a code $\mathscr{C}$ is large, then $\mathrm{PUD}(\mathscr{K},\mathscr{C},D)$ is small, where $\mathscr{K}$ is a product of the q-ary symmetric channel and D is the maximum likelihood decoder for $\mathscr{C}$. This motivates the following central problem in coding theory.

Problem 1.25. Construct error-correcting codes $\mathscr{C} \subseteq \mathbb{F}_q^n$ having large cardinality and large minimum distance *simultaneously*.

1.4 Other Exercises

Exercise **1.26.** Show that the Hamming distance is invariant under translations. In other words, show that for all $x, y, z \in \mathbb{F}_q^n$ we have $d^{\mathrm{H}}(x,y) = d^{\mathrm{H}}(x+z, y+z)$.

Exercise **1.27.** Let $\mathscr{K} = \mathrm{Er}(\mathscr{X},\alpha) = (\mathscr{X},\mathscr{Y},\mathbb{P})$ be the erasure channel as in Definition 1.4, and let $\mathscr{K}^n = (\mathscr{X}^n,\mathscr{Y}^n,\mathbb{P}^n)$ be its n-th product. Fix $x \in \mathscr{X}^n, y \in \mathscr{Y}^n$ and define the set $S := \{1 \leq i \leq n \mid y_i = ?\}$. Show that $\mathbb{P}^n(y \mid x) = 0$ if $y_i \neq x_i$ for at least a value of $i \in \{1,...,n\} \setminus S$, and that $\mathbb{P}^n(y \mid x) = \alpha^{|S|}\,(1-\alpha)^{n-|S|}$ otherwise.

Exercise **1.28.** Do the same calculations as in Example 1.9 with the fourth power $\mathrm{Er}^4(\mathbb{F}_2,\alpha)$ of the erasure channel, the same code $\mathscr{C} \subseteq \mathbb{F}_2^4$ and $y = (1,?,0,0)$.

Exercise **1.29** (solved in Appendix C)**.**

1. Construct a code $\mathscr{C} \subseteq \mathbb{F}_2^4$ with $|\mathscr{C}| = 2$ and $d^{\mathrm{H}}(\mathscr{C}) \geq 3$.
2. Show that there is no code $\mathscr{C} \subseteq \mathbb{F}_2^4$ with $d^{\mathrm{H}}(\mathscr{C}) \geq 3$ and $|\mathscr{C}| \geq 3$.
3. Show that there exists a code $\mathscr{C} \subseteq \mathbb{F}_3^4$ with $d^{\mathrm{H}}(\mathscr{C}) \geq 3$ and $|\mathscr{C}| \geq 3$.

Chapter 2

Codes with the Hamming Metric

2.1 Definitions and First Examples

We start by establishing the notation for the remainder of these notes.

Notation 2.1. In the sequel q is a prime power and n, k and d denote integers with $n \geq 1$, $0 \leq k \leq n$ and $1 \leq d \leq n+1$ (unless otherwise stated).

In these lecture notes we mainly focus on codes that live in the Hamming distance space $(\mathbb{F}_q^n, d^{\mathrm{H}})$, also called *block codes*.

Definition 2.2. A (**block**) **code** is a non-empty subset $\mathscr{C} \subseteq \mathbb{F}_q^n$. Its elements are the **codewords**. We call n the **length** of $\mathscr{C}$. We say that $\mathscr{C}$ is **linear** if it is an $\mathbb{F}_q$-subspace of $\mathbb{F}_q^n$. In this case we write $\mathscr{C} \leq \mathbb{F}_q^n$ to stress the linear structure. The **dimension** of $\mathscr{C} \leq \mathbb{F}_q^n$ is its dimension as a linear space over $\mathbb{F}_q$.

Remark 2.3. If $\mathscr{C} \leq \mathbb{F}_q^n$ is a linear code of dimension k, then we have $|\mathscr{C}| = q^k$. To see this, fix an $\mathbb{F}_q$-basis $\{v_1, ..., v_k\}$ of $\mathscr{C}$ and observe that all the elements of $\mathscr{C}$ can be uniquely written as a linear combination of the v_i's. The number of such linear combinations is q^k.

Recall from Definition 1.23 that the **minimum** (**Hamming**) **distance** of a code $\mathscr{C} \subseteq \mathbb{F}_q^n$ is the positive integer $d^{\mathrm{H}}(\mathscr{C}) = \min\{d^{\mathrm{H}}(x,y) \mid x, y \in \mathscr{C},\ x \neq y\}$, with the convention that $d^{\mathrm{H}}(\mathscr{C}) = n+1$ if $|\mathscr{C}| = 1$.

The following property of the minimum distance follows immediately from the definitions.

Remark 2.4. Let $\mathscr{D} \subseteq \mathscr{C} \subseteq \mathbb{F}_q^n$ be codes. Then $d^{\mathrm{H}}(\mathscr{D}) \geq d^{\mathrm{H}}(\mathscr{C})$.

Notation 2.5. Several references say that "$\mathscr{C}$ is an $(n, M, d)_q$ code" if $\mathscr{C} \subseteq \mathbb{F}_q^n$, $|\mathscr{C}| = M$, and $d^{\mathrm{H}}(\mathscr{C}) = d$. Similarly, they say that "$\mathscr{C}$ is an $[n, k, d]_q$ code" if is a linear $(n, q^k, d)_q$ code. The notation $(n, M, \geq d)_q$ or $[n, k, \geq d]$ is also sometimes used to denote codes whose minimum distance is lower-bounded by d.

Some simple codes can be constructed as follows.

Example 2.6 (trivial codes)**.** $\mathbb{F}_q^n$ and all sets $\mathscr{C} \subseteq \mathbb{F}_q^n$ of cardinality one are called **trivial codes**. Their minimum distances are 1 and $n+1$, respectively. The only linear trivial codes are $\{0\}$ and $\mathbb{F}_q^n$.

Example 2.7 (repetition codes)**.** The set $\mathscr{C} = \{(\alpha, ..., \alpha) \mid \alpha \in \mathbb{F}_q\} \leq \mathbb{F}_q^n$ is a linear code of dimension 1 and minimum distance n. It is called the n**-times repetition code**.

A third example is the following.

Example 2.8 (parity-check codes)**.** For $n \geq 2$ the set

$$\mathscr{C} = \left\{ (x_1, ..., x_n) \in \mathbb{F}_q^n \mid x_n = -\sum_{i=1}^{n-1} x_i \right\} \leq \mathbb{F}_q^n$$

is a linear code of dimension $n-1$. It is called the **parity-check code** of length n.

Exercise 2.9. Compute the minimum distance of the parity-check code of Example 2.8 for an arbitrary $n \geq 2$.

2.2 The Gilbert-Varshamov Bound

As already mentioned in Chapter 1, good codes should have large cardinality and large minimum distance at the same time. In this section we show the existence of codes of sufficiently large cardinality and minimum distance.

Definition 2.10. For $0 \leq r \leq n$ and $x \in \mathbb{F}_q^n$, the **Hamming ball** of radius r centered at x is the set $B_r^{\mathrm{H}}(x) = \{y \in \mathbb{F}_q^n \mid d^{\mathrm{H}}(y, x) \leq r\} \subseteq \mathbb{F}_q^n$.

The following result is an immediate consequence of Lemma 1.20.

Proposition 2.11. For all $0 \leq r \leq n$ and $x \in \mathbb{F}_q^n$ we have

$$|B_r^{\mathrm{H}}(x)| = \sum_{i=0}^{r} \binom{n}{i} (q-1)^i.$$

In particular, the latter cardinality does not depend on the choice of $x \in \mathbb{F}_q^n$.

We can now show the existence of codes of sufficiently large minimum distance and cardinality. The proof only uses the metric structure of $(\mathbb{F}_q^n, d^{\mathrm{H}})$.

Theorem 2.12 (Gilbert-Varshamov bound)**.** There exists a code $\mathscr{C} \subseteq \mathbb{F}_q^n$ with minimum distance $d^{\mathrm{H}}(\mathscr{C}) \geq d$ and cardinality

$$|\mathscr{C}| \geq \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i}. \tag{2.1}$$

Proof. Let β denote the quantity on the RHS of (2.1), and observe that β is the ratio between the cardinality of $\mathbb{F}_q^n$ and the size of the Hamming ball of radius $d-1$. Note moreover that the desired theorem is immediate if $d = n+1$, as in such case $\beta = 1$ and we can take e.g. $\mathscr{C} = \{0\}$. We henceforth assume $1 \leq d \leq n$.

Let $\mathscr{C}$ be a code having the maximum cardinality among all codes with minimum distance $\geq d$ (the set of codes with minimum distance $\geq d$ is non-empty, and therefore the mentioned maximum does exist). We will show that $|\mathscr{C}| \geq \beta$. For every $x \in \mathbb{F}_q^n$ there must exist a codeword $c \in C$ such that $d^{\mathrm{H}}(c, x) \leq d-1$, as otherwise $\mathscr{D} = \mathscr{C} \cup \{x\}$ would be a code with $d^{\mathrm{H}}(\mathscr{D}) \geq d$ and cardinality exceeding that of $\mathscr{C}$, a contradiction. It follows that $\mathbb{F}_q^n$ is contained within the union of the Hamming balls of radius $d-1$ centered at the codewords of $\mathscr{C}$. In symbols,

$$\mathbb{F}_q^n \subseteq \bigcup_{c \in \mathscr{C}} B_{d-1}(c).$$

We therefore have

$$q^n = |\mathbb{F}_q^n| \leq \left| \bigcup_{c \in \mathscr{C}} B_{d-1}(c) \right| \leq \sum_{c \in \mathscr{C}} |B_{d-1}(c)| = |\mathscr{C}| \cdot \sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i,$$

where the last equality follows from Proposition 2.11. Thus $|\mathscr{C}| \geq \beta$, as desired. □

2.3 Linear Codes and Their Defining Matrices

Most interesting codes are linear subspaces $\mathscr{C} \leq \mathbb{F}_q^n$. These can be conveniently represented via matrices, which make them easy to handle.

Definition 2.13. We say that a matrix G with entries in $\mathbb{F}_q$ is a **generator matrix** of/for a linear code $\mathscr{C} \leq \mathbb{F}_q^n$ if it has full rank and its rows generate $\mathscr{C}$ over $\mathbb{F}_q$. We also say that $\mathscr{C}$ is **generated** by G.

Observe that the empty matrix $G : \emptyset \times \{1, ..., n\} \to \mathbb{F}_q$ is a generator matrix for the zero code $\{0\} \leq \mathbb{F}_q^n$.

Remark 2.14. If a code $\mathscr{C} \leq \mathbb{F}_q^n$ has dimension $k \geq 1$ and G is a generator matrix of $\mathscr{C}$, then $\mathscr{C} = \{x \cdot G \mid x \in \mathbb{F}_q^k\}$.

Every code $\mathscr{C} \leq \mathbb{F}_q^n$ has a generator matrix, which is not unique in general. However, we can select a canonical generator matrix using the notion of reduced row-echelon form.

Recall that a matrix M over a field is in **reduced row-echelon form** if:

1. the zero rows of M are grouped at the bottom;
2. each non-zero row of M has more initial zeros than the previous rows;

3. the first non-zero entry of any non-zero row of M (called the **pivot entry** of the row) equals 1 and is the only non-zero entry in its column.

Every matrix M can be put in reduced row-echelon form by performing elementary row operations. Such a reduced row-echelon form is unique and denoted by $\mathrm{RREF}(M)$. Therefore every linear code $\mathscr{C} \leq \mathbb{F}_q^n$ has a unique generator matrix in reduced row-echelon form, which we call the **standard generator matrix** of $\mathscr{C}$.

Example 2.15. The n-times repetition code (Example 2.7) has standard generator matrix

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

Note that any matrix $G = \begin{pmatrix} \alpha & \alpha & \cdots & \alpha \end{pmatrix}$ with $\alpha \neq 0$ is a generator matrix of the n-times repetition code.

Exercise **2.16.** Write down the standard generator matrix of the codes of Examples 2.6 and 2.8.

Exercise **2.17.** Let G be a generator matrix of a code $\mathscr{C} \leq \mathbb{F}_q^n$ and let $A \in \mathbb{F}_q^{k \times k}$ be an invertible matrix. Show that $A \cdot G$ is a generator matrix of $\mathscr{C}$ as well.

When a code is linear, its minimum distance can be computed by taking the minimum of the *Hamming weights* of the non-zero codewords. These are defined as follows.

Definition 2.18. The (**Hamming**) **weight** of a vector $x \in \mathbb{F}_q^n$ is $\omega^{\mathrm{H}}(x) := |\{i \mid x_i \neq 0\}|$.

Example 2.19. The ternary vectors $(0,1,0,2)$, $(0,0,0,0)$ and $(1,2,2,1)$ in $\mathbb{F}_3^3$ have weight 2, 0 and 4 respectively.

Proposition 2.20. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a non-zero linear code. We have

$$d^{\mathrm{H}}(\mathscr{C}) = \min\{\omega^{\mathrm{H}}(x) \mid x \in \mathscr{C},\ x \neq 0\}.$$

Proof. Exercise (the proof can be found in Appendix C). □

Example 2.21. The matrix

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix} \in \mathbb{F}_3^{3\times 3}$$

is not the generator matrix of any code. Indeed, it is does not have full rank as the last row is the sum of the previous two. The matrix

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \in \mathbb{F}_3^{2\times 3}$$

instead generates a linear code $\mathscr{C} \leq \mathbb{F}_3^3$ of dimension 2 over $\mathbb{F}_3$. Therefore $|\mathscr{C}| = 3^2 = 9$ by Remark 2.3. The elements of $\mathscr{C}$ are

$$\mathscr{C} = \{(0,0,0), (1,1,0), (0,1,2), (2,2,0), (0,2,1), (1,2,2), (2,1,1), (1,0,1), (2,0,2)\}.$$

We also have $\{\omega^{\mathrm{H}}(v) \mid v \in \mathscr{C},\, v \neq 0\} = \{2, 3\}$, from which $d^{\mathrm{H}}(\mathscr{C}) = 2$.

Example 2.22. Let $x, y \in \mathbb{F}_2^n$ be vectors of even Hamming weight. Then it is easy to see (exercise) that $x + y$ also has even Hamming weight. Therefore the set $\mathscr{C}$ of vectors of even Hamming weight in $\mathbb{F}_2^n$ is a linear code (called the **even weight code** of length n). We have $\dim(\mathscr{C}) = n - 1$ and $d^{\mathrm{H}}(\mathscr{C}) = 2$.

Another (very useful) way of describing a linear code is via a *parity-check matrix*.

Definition 2.23. We say that a matrix H with entries in $\mathbb{F}_q$ is a **parity-check matrix** of/for a linear code $\mathscr{C} \leq \mathbb{F}_q^n$ if it has full rank and $\mathscr{C}$ is the left kernel of $H^\top$, i.e, if

$$\mathscr{C} = \left\{x \in \mathbb{F}_q^n \mid x \cdot H^\top = 0\right\}. \tag{2.2}$$

Note that a parity-check matrix H for a code $\mathscr{C}$ uniquely determines $\mathscr{C}$ because of the defining property in (2.2).

The following result is an immediate consequence of the rank-nullity theorem and of the definitions.

Proposition 2.24. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code of dimension k, and let G, H be a generator and a parity-check matrix of $\mathscr{C}$, respectively. Then G has size $k \times n$ and H has size $(n - k) \times n$. Moreover, $GH^\top = 0$.

Example 2.25. If $\mathscr{C} = \{0\}$ is the zero code, then any invertible $n \times n$ matrix H is a parity-check matrix of $\mathscr{C}$. The empty matrix is the parity-check matrix of $\mathbb{F}_q^n$.

The parity-check matrix of a code can be easily computed from a generator matrix, provided that the latter has a special form.

Proposition 2.26. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $k \geq 1$. Suppose that $\mathscr{C}$ has a generator matrix of the form $G = (I_k \mid A)$, where I_k is the identity $k \times k$ matrix. Then

$$H = (-A^\top \mid I_{n-k})$$

is a parity-check matrix of $\mathscr{C}$

Proof. One easy checks that $G \cdot H^\top = 0$, from which we see that the code $\mathscr{C}$ is contained in the left kernel of $H^\top$. Since H has rank $n-k$, $H^\top$ has rank $n-k$ as well. In particular, the left kernel of $H^\top$ has dimension $n-(n-k)$ by the rank-nullity theorem. Therefore $\mathscr{C}$ must be exactly the left kernel of $H^\top$ and H is a parity-check matrix of $\mathscr{C}$. □

Example 2.27. If $\mathscr{C}$ is the 5-times repetition code, then combining Example 2.15 with Proposition 2.26 we see that a parity-check matrix of $\mathscr{C}$ is

$$H = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{F}_q^{4\times 5}.$$

Exercise **2.28.** Find a parity-check matrix of the codes of Examples 2.6 and 2.8.

Every linear code has at least one parity-check matrix, as the following linear algebra result result shows.

Proposition 2.29. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code. There exists a parity-check matrix of $\mathscr{C}$.

Proof. Let k be the dimension of $\mathscr{C}$. Then the quotient space $\mathbb{F}_q^n/\mathscr{C}$ has dimension $n-k$ and is therefore isomorphic to $\mathbb{F}_q^{n-k}$. Let $g : \mathbb{F}_q^n/\mathscr{C} \to \mathbb{F}_q^{n-k}$ be an isomorphism and let $f : \mathbb{F}_q^n \to \mathbb{F}_q^n/\mathscr{C}$ be the projection on the quotient. The composition $g \circ f$ is linear, surjective, and its kernel is $\mathscr{C}$. We can therefore take as H the matrix of $g \circ f$ with respect to the canonical bases of $\mathbb{F}_q^n$ and $\mathbb{F}_q^{n-k}$ (the images are put in the rows of H). □

The minimum distance of a linear code can be computed by looking at the columns of a parity-check matrix H of $\mathscr{C}$. We will give a slightly more general result using the notion of *Hamming support* of a vector.

Definition 2.30. The (**Hamming**) **support** of $x \in \mathbb{F}_q^n$ is $\sigma^{\mathrm{H}}(x) := \{1 \leq i \leq n \mid x_i \neq 0\}$.

Remark 2.31. It easily follows from the definitions that for all vectors $x \in \mathbb{F}_q^n$ we have $|\sigma^{\mathrm{H}}(x)| = \omega^{\mathrm{H}}(x)$. In words, the weight of a vector is the cardinality of its support.

We can now show that the linear dependencies among the columns of a parity-check matrix gives us information on the supports of the codewords.

Proposition 2.32. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $0 \leq k \leq n-1$, and let H be a parity-check matrix of $\mathscr{C}$. Then for all subsets $S \subseteq \{1, ..., n\}$ with $|S| \geq 1$ the following are equivalent:

1. the columns of H indexed by S are linearly dependent,
2. there exists $x \in \mathscr{C}$ with $x \neq 0$ and $\sigma^{\mathrm{H}}(x) \subseteq S$.

Proof. Let $h_1, ..., h_n \in \mathbb{F}_q^{n-k}$ be the columns of H. If the vectors $\{h_i \mid i \in S\}$ are linearly dependent, there exist field elements $(\alpha_i \mid i \in S)$, not all zero, such that $\sum_{i \in S} \alpha_i h_i = 0$. Let $x \in \mathbb{F}_q^n$ be the vector whose i-th component is α_i for all $i \in S$, and that is zero elsewhere. By definition, $\sigma^{\mathrm{H}}(x) \subseteq S$. We also have $\sum_{i=1}^n x_i h_i = 0$, i.e., $H \cdot x^\top = 0$ (or equivalently $x \cdot H^\top = 0$). Therefore $x \in \mathscr{C}$. The other direction is analogous and left as exercise. □

Using Proposition 2.32 we can therefore express the minimum distance of a linear code as follows.

Corollary 2.33. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $0 \leq k \leq n-1$, and let H be a parity-check matrix of $\mathscr{C}$. We have

$$\begin{aligned} d^{\mathrm{H}}(\mathscr{C}) &= 1 + \max\{1 \leq i \leq n \mid \text{every } i \text{ columns of } H \text{ are linearly independent}\} \\ &= \min\{1 \leq i \leq n \mid \text{there exist } i \text{ columns of } H \text{ that are linearly dependent}\}. \end{aligned}$$

Proof. If $k = 0$ then $d^{\mathrm{H}}(\mathscr{C}) = n+1$ by definition and H is an invertible $n \times n$ matrix by Proposition 2.24. Therefore

$$\max\{1 \leq i \leq n \mid \text{every } i \text{ columns of } H \text{ are linearly independent}\} = n,$$

and the result follows. If $1 \leq k \leq n-1$ then the corollary is an immediate consequence of Proposition 2.32. □

Example 2.34. Let $\mathscr{C} \leq \mathbb{F}_2^6$ be the linear code defined by the parity-check matrix

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

Every two columns of H are linearly independent, so $d^{\mathrm{H}}(\mathscr{C}) \geq 3$. Columns 2,3 and 4 are linearly dependent, and therefore there exists a non-zero codeword $x \in \mathscr{C}$ whose Hamming support is contained in $\{2,3,4\}$. In particular, $d^{\mathrm{H}}(\mathscr{C})$ must be 3. The fact that the columns 1, 2 and 3 are independent tells us that there is no codeword $x \in \mathscr{C}$ whose Hamming support is contained in $\{1,2,3\}$.

2.4 Syndrome Decoding

In this section we describe a decoding algorithm that works for *any* linear code. The main idea behind the decoding procedure is the following.

Suppose that we are working with a linear code $\mathscr{C} \leq \mathbb{F}_q^n$. Let $x \in \mathscr{C}$ be transmitted and let $y \in \mathbb{F}_q^n$ be received. Pick any $\widehat{e} \in \mathbb{F}_q^n$ of minimal Hamming weight with the property that $y - \widehat{e} \in \mathscr{C}$. In other words, $\widehat{e}$ is a vector that attains

$$\min\{\omega^{\mathrm{H}}(e) \mid e \in \mathbb{F}_q^n,\ y - e \in \mathscr{C}\}.$$

Then $\widehat{x} := y - \widehat{e} \in \mathscr{C}$ turns out to be a good estimate for the transmitted codeword x with respect to the minimum distance decoder (Definition 1.18). More precisely, we have

$$d^{\mathrm{H}}(y - \widehat{e}, y) \leq d^{\mathrm{H}}(x', y) \quad \text{for all } x' \in \mathscr{C}.$$

Indeed, suppose towards a contradiction that $d^{\mathrm{H}}(y - \widehat{e}, y) > d^{\mathrm{H}}(x', y)$ for some $x' \in \mathscr{C}$. Setting $e' := y - x'$, this implies $\omega^{\mathrm{H}}(\widehat{e}) > \omega^{\mathrm{H}}(e')$. Since $y - e' = x' \in \mathscr{C}$, this contradicts the definition of $\widehat{e}$.

The above discussion gives us an idea for a decoding algorithm for an arbitrary linear code $\mathscr{C}$: When $y \in \mathbb{F}_q^n$ is received, we search for a vector $e \in \mathbb{F}_q^n$ of minimal Hamming weight in the set $y + \mathscr{C} = \{y + x \mid x \in \mathscr{C}\}$, and we decode y to $y - e$.

In order to make this algorithm efficient, we need a good strategy to compute a vector e of minimal weight from y. Using a parity-check matrix H of $\mathscr{C}$ can help. We start by observing that the set $y + \mathscr{C}$ is the equivalence class of y with respect to the equivalence relation on $\mathbb{F}_q^n$ defined by $y \sim y'$ if and only if $y - y' \in \mathscr{C}$.

Lemma 2.35. For all $y, y' \in \mathbb{F}_q^n$ we have $y \sim y'$ if and only if $y \cdot H^\top = y' \cdot H^\top$.

Proof. We have $y \cdot H^\top = y' \cdot H^\top$ if and only if $(y - y') \cdot H^\top = 0$. Since $\mathscr{C}$ is the left kernel of $H^\top$, this is true if and only if $y - y' \in \mathscr{C}$. □

The previous lemma shows that the equivalence classes of $\sim$ are in bijection with the elements of the set $\{y \cdot H^\top \mid y \in \mathbb{F}_q^n\}$, which are called **syndromes**. We can therefore prepare a table that lists all syndromes, and that for each syndrome s lists a vector of minimal weight e having that syndrome, also called a **coset leader**. Once we have done that we can apply the following algorithm.

Algorithm 2.36 (Syndrome Decoding)**.** A code $\mathscr{C}$ is used, and H is a parity-check matrix of $\mathscr{C}$. A vector $y \in \mathbb{F}_q^n$ is received.

1. Compute the syndrome $s = y \cdot H^\top$.
2. Find s in the table and the corresponding vector $e \in \mathbb{F}_q^n$.
3. Decode y to $y - e$.

Note that the syndromes of $\mathscr{C}$ are in bijection with the elements of the quotient space $\mathbb{F}_q^n/\mathscr{C}$. In particular, there are q^{n-k} syndromes (i.e., rows in our table).

Example 2.37. Let $\mathscr{C} \leq \mathbb{F}_2^6$ be the linear code defined by the parity-check matrix

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

Write

$$H^\top = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

and observe that, by definition, each row of $H^\top$ is a syndrome. The vector $(0, 0, 0)$ is another syndrome. As the rows of $H^\top$ are distinct we already found 7 syndromes out

of $2^{6-(6-3)} = 2^3 = 8$. Note moreover that $(1,0,1,0,0,0) \cdot H^\top = (1,0,1)$, which is then the eigth syndrome. For each syndrome s we now need to find a vector e having that syndrome and minimal Hamming weight, and fill in the table. One possibility is the following:

Syndrome	Coset leader
$(0,0,0)$	(0,0,0,0,0,0)
$(1,0,0)$	(1,0,0,0,0,0)
$(0,1,0)$	(0,1,0,0,0,0)
$(0,0,1)$	(0,0,1,0,0,0)
$(0,1,1)$	(0,0,0,1,0,0)
$(1,1,1)$	(0,0,0,0,1,0)
$(1,1,0)$	(0,0,0,0,0,1)
$(1,0,1)$	(1,0,1,0,0,0)

Now suppose that $x = (0,1,1,1,0,0) \in \mathscr{C}$ is transmitted, that one error occurs and $y = (0,1,1,1,0,1)$ is received. We apply syndrome decoding and compute

$$y \cdot H^\top = (1,1,0).$$

The corresponding coset leader in the table is $e = (0,0,0,0,0,1)$. We therefore decode to $y - e = (0,1,1,1,0,0)$, which is the transmitted codeword, x.

2.5 Weight Distribution and Its Significance

When working with a code $\mathscr{C} \subseteq \mathbb{F}_q^n$ for the q-ary symmetric channel (Definition 1.3), a particularly undesirable situation is when a codeword $x \in \mathscr{C}$ is transmitted and a different codewords $y \in \mathscr{C}$, $y \neq x$, is received. In such a scenario we are unable to realize that an error has occurred in the transmission and we certainly make a decoding error.

Let $(\mathbb{F}_q^n, \mathbb{F}_q^n, \mathbb{P}^n) = \mathrm{Sym}^n(q,\alpha)$ be the n-th product of the q-ary symmetric channel with error probability $0 \leq \alpha \leq 1$. Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code with $|\mathscr{C}| \geq 2$. The **probability of an undetected error** associated with n, q, α and $\mathscr{C}$ is

$$\mathrm{PUE}(n,q,\alpha,\mathscr{C}) := \frac{1}{|\mathscr{C}|} \sum_{x \in \mathscr{C}} \sum_{\substack{y \in \mathscr{C} \\ y \neq x}} \mathbb{P}^n(y \mid x).$$

When the code $\mathscr{C}$ is linear, the above probability can be bounded in terms of a fundamental parameter of the underlying code.

Definition 2.38. Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code. The **weight distribution** of $\mathscr{C}$ is the integer vector $(W_0^{\mathrm{H}}(\mathscr{C}), ..., W_n^{\mathrm{H}}(\mathscr{C})) \in \mathbb{N}^{n+1}$, where for all $0 \leq i \leq n$ the number

$$W_i^{\mathrm{H}}(\mathscr{C}) := |\{x \in \mathscr{C} \mid \omega^{\mathrm{H}}(x) = i\}|$$

counts the number of codewords in $\mathscr{C}$ with Hamming weight i.

Note that for any code $\mathscr{C} \subseteq \mathbb{F}_q^n$ we have $\sum_{i=0}^{n} W_i^{\mathrm{H}}(\mathscr{C}) = |\mathscr{C}|$.

Example 2.39. The n-times repetition code $\mathscr{C}$ (Example 2.7) has weight distribution $W_0^{\mathrm{H}}(\mathscr{C}) = 1$, $W_n^{\mathrm{H}}(\mathscr{C}) = q-1$ and $W_i^{\mathrm{H}}(\mathscr{C}) = 0$ for all $1 \leq i \leq n-1$. Observe that $1 + (q-1) = q$, which is the cardinality of the $\mathscr{C}$.

Exercise 2.40. Write down the weight distribution of the trivial codes $\{0\}$ and $\mathbb{F}_q^n$.

Weight distribution and probability of an undetected error relate as follows.

Theorem 2.41. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code for the n-th product $\mathrm{Sym}^n(q, \alpha)$ of the q-ary symmetric channel with error probability $0 \leq \alpha \leq 1$. We have

$$\mathrm{PUE}(n, q, \alpha, \mathscr{C}) = \sum_{i=1}^{n} \left(\frac{\alpha}{q-1}\right)^i (1-\alpha)^{n-i}\, W_i^{\mathrm{H}}(\mathscr{C}).$$

Proof. By Proposition 1.17 we have

$$\begin{aligned}
\mathrm{PUE}(n, q, \alpha, \mathscr{C}) &= \frac{1}{|\mathscr{C}|} \sum_{x \in \mathscr{C}} \sum_{\substack{y \in \mathscr{C} \\ y \neq x}} \mathbb{P}^n(y \mid x) \\
&= \frac{1}{|\mathscr{C}|} \sum_{x \in \mathscr{C}} \sum_{i \geq 1} \sum_{\substack{y \in \mathscr{C} \\ d^{\mathrm{H}}(y,x) = i}} \left(\frac{\alpha}{q-1}\right)^i (1-\alpha)^{n-i} \\
&= \frac{1}{|\mathscr{C}|} \sum_{i \geq 1} \left(\frac{\alpha}{q-1}\right)^i (1-\alpha)^{n-i} \sum_{x \in \mathscr{C}} |\{y \in \mathscr{C} \mid d^{\mathrm{H}}(x, y) = i\}|.
\end{aligned}$$

Now observe that, as $\mathscr{C}$ is linear, for all fixed $x \in \mathscr{C}$ and $i \geq 1$ the map

$$\{y \in \mathscr{C} \mid d^{\mathrm{H}}(x, y) = i\} \to \{y \in \mathscr{C} \mid \omega^{\mathrm{H}}(y) = i\}$$

defined by $y \mapsto x - y$ is a bijection. Therefore we conclude that

$$\mathrm{PUE}(n, q, \alpha, \mathscr{C}) = \sum_{i=1}^{n} \left(\frac{\alpha}{q-1}\right)^i (1-\alpha)^{n-i}\, W_i^{\mathrm{H}}(\mathscr{C}),$$

as desired. □

A convenient way of writing the weight distribution of a code is by encoding it in a bivariate homogeneous polynomial.

Definition 2.42. Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code. The **weight enumerator** of $\mathscr{C}$ is the polynomial

$$W^{\mathrm{H}}(\mathscr{C}) = \sum_{i \geq 0} W_i^{\mathrm{H}}(\mathscr{C}) X^i Y^{n-i} \in \mathbb{R}[X, Y].$$

Note that $W^{\mathrm{H}}(\mathscr{C})$ is homogeneous of degree n. We conclude by re-stating Theorem 2.41 in a polynomial form.

Corollary 2.43. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code for the n-th product $\mathrm{Sym}^n(q,\alpha)$ of the q-ary symmetric channel with error probability $0 \leq \alpha \leq 1$. We have

$$\mathrm{PUE}(n,q,\alpha,\mathscr{C}) = W^{\mathrm{H}}(\mathscr{C})\left(\frac{\alpha}{q-1}, 1-\alpha\right) - (1-\alpha)^n.$$

Example 2.44. Continuing Example 2.39, the n-times repetition code $\mathscr{C}$ has weight enumerator $Y^n + (q-1)X^n$. Therefore by Theorem 2.41 or Corollary 2.43 we have

$$\mathrm{PUE}(n,q,\alpha,\mathscr{C}) = \frac{\alpha^n}{(q-1)^{n-1}}.$$

Exercise **2.45.** Compute the weight distribution and the weight enumerator of the code generated over $\mathbb{F}_3$ by the matrix

$$G = \begin{pmatrix} 2 & 0 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Exercise **2.46.** Let $\mathscr{C} \leq \mathbb{F}_2^n$ be a binary linear code of dimension $k \geq 1$ having a generator matrix whose rows all have even Hamming weight. Prove that $W_i^{\mathrm{H}}(\mathscr{C}) = 0$ for all odd integers $1 \leq i \leq n$.

2.6 New Codes from Old

In this subsection we discuss ways of obtaining new codes starting from an old one. These operations will be needed later to study structural properties of error-correcting objects. In the sequel, we denote by S^c the complement of a subset $S \subseteq \{1,...,n\}$.

Notation 2.47. Let $S \subseteq \{1,...,n\}$ be a non-empty set and let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a linear code. We let $\mathscr{C}(S) := \{x \in \mathscr{C} \mid \sigma^{\mathrm{H}}(x) \subseteq S\}$ be the set of codewords of $\mathscr{C}$ whose Hamming support is contained in S.

Note that $\mathscr{C}(S)$ is a subcode of $\mathscr{C} \subseteq \mathbb{F}_q^n$ for any non-empty set S. In particular, by Remark 2.4 we have $d^{\mathrm{H}}(\mathscr{C}(S)) \geq d^{\mathrm{H}}(\mathscr{C})$.

After constructing $\mathscr{C}(S)$, the coordinates in S^c are often deleted as they are zero. This is formally done via a projection map, which we will often use in the sequel.

Notation 2.48. For a non-empty set $S \subseteq \{1,...,n\}$ we let $\pi_S : \mathbb{F}_q^n \to \mathbb{F}_q^{|S|}$ be the projection on the coordinates indexed by S.

We can obtain a new code from an old one by deleting/selecting some coordinates.

Definition 2.49. Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code and let $S \subseteq \{1, ..., n\}$ be a non-empty set. The S**-puncturing** of $\mathscr{C}$ is $\pi_S(\mathscr{C})$ and its S**-shortening** is $\pi_S(\mathscr{C}(S))$.

Note that the S-puncturing and the S-shortening of $\mathscr{C}$ are codes of length $|S|$. Moreover, if both $\mathscr{C}$ and $\mathscr{C}(S)$ have cardinality at least two then

$$d^{\mathrm{H}}(\mathscr{C}(S)) = d^{\mathrm{H}}(\pi_S(\mathscr{C}(S))) \geq d^{\mathrm{H}}(\mathscr{C}).$$

This easily follows from the definitions and Remark 2.4.

Remark 2.50. If $\mathscr{C} \leq \mathbb{F}_q^n$ is a linear code then its S-puncturing and S-shortening are linear codes as well.

Example 2.51. Let $\mathscr{C} \leq \mathbb{F}_3^4$ be the code generated by

$$G = \begin{pmatrix} 1 & 1 & 2 & 0 \\ 1 & 2 & 2 & 1 \end{pmatrix}.$$

Let $S = \{1, 3, 4\}$. We have $\mathscr{C}(S) = \{(0, 0, 0, 0), (2, 0, 1, 1), (1, 0, 2, 2)\}$, which is a linear code of dimension 1 in $\mathbb{F}_3^4$. Therefore the S-shortening of $\mathscr{C}$ is

$$\pi_S(\mathscr{C}(S)) = \{(0, 0, 0), (2, 1, 1), (1, 2, 2)\} \leq \mathbb{F}_3^3.$$

The S-puncturing of $\mathscr{C}$ is instead the code $\pi_S(\mathscr{C})$ generated by

$$G' = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

and has therefore dimension 2 over $\mathbb{F}_3$. Note that $\mathscr{C}$ has minimum distance 2, $\pi_S(\mathscr{C})$ has minimum distance 1, while $\pi_S(\mathscr{C}(S))$ has minimum distance 3.

Remark 2.52. The S-puncturing of a code $\mathscr{C}$ may have minimum distance strictly larger than that of $\mathscr{C}$. Take for example the code $\mathscr{C} \leq \mathbb{F}_2^4$ generated by

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

and let $S = \{1, 2, 3\}$. We have $d^{\mathrm{H}}(\mathscr{C}) = 1$ while $d^{\mathrm{H}}(\pi_S(\mathscr{C})) = 3$.

Exercise **2.53.** Let $\mathscr{C} \leq \mathbb{F}_q^n$ be the n-times repetition code. Compute the dimension of $\mathscr{C}(S)$ for all non-empty sets $S \subseteq \{1, ..., n\}$. Show that $\pi_S(\mathscr{C})$ is the $|S|$-times repetition code for all non-empty sets $S \subseteq \{1, ..., n\}$.

The following simple construction increases the length of a code by one.

Definition 2.54. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code. The **extension** of $\mathscr{C}$ is the linear code

$$\mathscr{C}^{\mathrm{ext}} := \{(x_1, ..., x_n, -x_1 - \cdots - x_n) \mid (x_1, ..., x_n) \in \mathscr{C}\} \leq \mathbb{F}_q^{n+1}.$$

Exercise **2.55** (solved in Appendix C). 1. Following the notation of Definition 2.54, show that $\mathscr{C}^{\text{ext}}$ is a code of length $n+1$, with the same dimension as $\mathscr{C}$, and with $d^{\text{H}}(\mathscr{C}^{\text{ext}}) \geq d^{\text{H}}(\mathscr{C})$.

2. Show that if $\mathscr{C} \leq \mathbb{F}_2^n$ is a non-zero binary code of odd minimum distance d, then $\mathscr{C}^{\text{ext}}$ has minimum distance $d+1$.

Exercise **2.56.** Show that $\mathscr{C} \leq \mathbb{F}_q^n$ is a linear code of dimension $1 \leq k \leq n-1$ with generator and parity-check matrices $G = (G_{ij})$ and H, respectively, then $\mathscr{C}^{\text{ext}}$ has generator and parity-check matrices

$$G' = \begin{pmatrix} G_{11} & G_{12} & \cdots & G_{1n} & -G_{11} - \cdots - G_{1n} \\ G_{21} & G_{22} & \cdots & G_{2n} & -G_{21} - \cdots - G_{2n} \\ \vdots & & & & \\ G_{k1} & G_{k2} & \cdots & G_{kn} & -G_{k1} - \cdots - G_{kn} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 & \cdots & 1 \\ & & & 0 \\ & H & & 0 \\ & & & \vdots \\ & & & 0 \end{pmatrix},$$

respectively.

In the remainder of the section we describe two constructions that produce a new code combining two old codes.

Definition 2.57. Let $n_1, n_2 \geq 1$ be integers and let $\mathscr{C}_1 \subseteq \mathbb{F}_q^{n_1}$, $\mathscr{C}_2 \subseteq \mathbb{F}_q^{n_2}$ be codes. The **product** of $\mathscr{C}_1$ and $\mathscr{C}_2$ is the code

$$\mathscr{C}_1 \times \mathscr{C}_2 := \{(x, y) \mid x \in \mathscr{C}_1,\, y \in \mathscr{C}_2\} \subseteq \mathbb{F}_q^{n_1+n_2}.$$

Following the notation of Definition 2.57, if $|\mathscr{C}_1|, |\mathscr{C}_2| \geq 2$ then $\mathscr{C}_1 \times \mathscr{C}_2$ has length n_1+n_2, cardinality $|\mathscr{C}_1| \cdot |\mathscr{C}_2|$, and minimum distance $d^{\text{H}}(\mathscr{C}_1 \times \mathscr{C}_2) = \min\{d^{\text{H}}(\mathscr{C}_1), d^{\text{H}}(\mathscr{C}_2)\}$.

The last code construction of this section is simple but very interesting. We will use it to study the family of Reed-Muller codes in Chapter 6.

Definition 2.58. Let $\mathscr{C}_1, \mathscr{C}_2 \leq \mathbb{F}_q^n$ be linear codes. The **Plotkin sum** of $\mathscr{C}_1$ and $\mathscr{C}_2$ is the linear code

$$\mathscr{C}_1 \oplus_{\text{P}} \mathscr{C}_2 := \{(x, x+y) \mid x \in \mathscr{C}_1,\, y \in \mathscr{C}_2\} \leq \mathbb{F}_q^{2n}.$$

Exercise **2.59.** Following the notation of Definition 2.58, show that the dimension of the Plotkin sum $\mathscr{C}_1 \oplus_{\text{P}} \mathscr{C}_2$ is the sum of the dimensions of $\mathscr{C}_1$ and $\mathscr{C}_2$. Write down a generator matrix of $\mathscr{C}_1 \oplus_{\text{P}} \mathscr{C}_2$ in terms of generator matrices G_1 and G_2 for $\mathscr{C}_1$ and $\mathscr{C}_2$, respectively.

Example 2.60. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be the n-times repetition code (Example 2.7). Then the code $\mathscr{C} \oplus_{\text{P}} \mathscr{C} \leq \mathbb{F}_q^{2n}$ has

$$\begin{pmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \end{pmatrix} \in \mathbb{F}_q^{2\times 2n}$$

as generator matrix. In particular, it has dimension 2 and minimum distance n.

The Plotkin sum has the following important general property.

Proposition 2.61. Let $\mathscr{C}_1, \mathscr{C}_2$ be as in Definition 2.58. Assume that $\mathscr{C}_1$ and $\mathscr{C}_2$ are both non-zero and denote by d_1 and d_2 their minimum distances, respectively. Then $\mathscr{C}_1 \oplus_{\mathrm{P}} \mathscr{C}_2$ has minimum distance $\min\{2d_1, d_2\}$.

Proof. Exercise (*Hint*: if $y \in \mathscr{C}_2$ and $y_i \neq 0$, then either $x_i \neq 0$ or $x_i \neq y_i$). The solution can be found in Appendix C. □

2.7 The Dual Code

The vector space $\mathbb{F}_q^n$ is endowed with a symmetric, bilinear, non-degenerate form $\mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{F}_q$ defined by

$$(x, y) \mapsto \langle x, y \rangle := \sum_{i=1}^{n} x_i y_i.$$

If $\langle x, y \rangle = 0$ we say that x and y are **orthogonal**. Recall that being non-degenerate means that if a vector $y \in \mathbb{F}_q^n$ satisfies $\langle x, y \rangle = 0$ for all $x \in \mathbb{F}_q^n$, then we must have $y = 0$. In other words, the zero vector is the only one to be orthogonal to all vectors.

We call $\langle x, y \rangle$ the **scalar product** (or the **inner product**) of x and y. The following properties are easy to verify. In the sequel we will apply these properties without explicitly referring to them.

Proposition 2.62. Let $x, x', y, y' \in \mathbb{F}_q^n$ and let $\alpha \in \mathbb{F}_q$. The following hold:

1. $\langle x + x', y + y' \rangle = \langle x, y \rangle + \langle x', y \rangle + \langle x, y' \rangle + \langle x', y' \rangle$,
2. $\alpha \langle x, y \rangle = \langle \alpha x, y \rangle = \langle x, \alpha y \rangle$.

Definition 2.63. The **dual** of a linear code $\mathscr{C} \leq \mathbb{F}_q^n$ is

$$\mathscr{C}^{\perp} := \{y \in \mathbb{F}_q^n \mid \langle x, y \rangle = 0 \text{ for all } x \in \mathscr{C}\}.$$

We summarize the properties of the dual code in the next result.

Theorem 2.64. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code. The following hold.

1. $\mathscr{C}^{\perp}$ is a linear code.
2. If G is a generator matrix of $\mathscr{C}$, then G is a parity-check matrix of $\mathscr{C}^{\perp}$.
3. $\dim(\mathscr{C}^{\perp}) = n - \dim(\mathscr{C})$.
4. If H is a parity-check matrix of $\mathscr{C}$, then H is a generator matrix of $\mathscr{C}^{\perp}$.
5. $(\mathscr{C}^{\perp})^{\perp} = \mathscr{C}$.
6. For any linear code $\mathscr{D} \leq \mathbb{F}_q^n$ we have $(\mathscr{C} \cap \mathscr{D})^{\perp} = \mathscr{C}^{\perp} + \mathscr{D}^{\perp}$ and $(\mathscr{C} + \mathscr{D})^{\perp} = \mathscr{C}^{\perp} \cap \mathscr{D}^{\perp}$.

Proof. 1. The fact that $\mathscr{C}^\perp$ is linear easily follows from the definitions.

2. By the linearity of $\mathscr{C}$, a vector x belongs to $\mathscr{C}^\perp$ if and only if it belongs to the left kernel of $G^\top$. Therefore $\mathscr{C}^\perp$ is precisely the left kernel of $G^\top$. This shows that G is a parity-check matrix of $\mathscr{C}^\perp$.

3. Use part 2 and Proposition 2.24.

4. By definition, the left kernel of $H^\top$, say $\mathscr{L}$, is contained in $\mathscr{C}^\perp$. By Proposition 2.24 and part 3, the spaces $\mathscr{L}$ and $\mathscr{C}^\perp$ have the same dimension. Therefore they are equal.

5. The inclusion $\mathscr{C} \subseteq (\mathscr{C}^\perp)^\perp$ is easy. The other follows from part 3.

6. This is left as an exercise (*Hint*: for each identity, prove one inclusion directly and obtain the other by a dimension argument). □

Example 2.65. Consider the finite field $\mathbb{F}_4 = \mathbb{F}_2[\alpha]$, where $\alpha^2 = \alpha + 1$. Let $\mathscr{C} \leq \mathbb{F}_4^5$ be the code generated by the matrix

$$G = \begin{pmatrix} 1 & \alpha^2 & 0 & 1 & \alpha \\ \alpha & \alpha & \alpha & 1 & 1 \end{pmatrix}.$$

Then $\mathscr{C}$ has dimension 2 and minimum distance 3. The reduced row-echelon form of G is

$$G' = \begin{pmatrix} 1 & 0 & \alpha & \alpha & 0 \\ 0 & 1 & \alpha^2 & 1 & \alpha^2 \end{pmatrix},$$

which is another generator matrix of $\mathscr{C}$. By Proposition 2.26,

$$H = \begin{pmatrix} \alpha & \alpha^2 & 1 & 0 & 0 \\ \alpha & 1 & 0 & 1 & 0 \\ 0 & \alpha^2 & 0 & 0 & 1 \end{pmatrix}$$

is a parity-check matrix of $\mathscr{C}$, and therefore a generator matrix of the dual code $\mathscr{C}^\perp$. Its reduced row-echelon form is

$$H' = \begin{pmatrix} 1 & 0 & 0 & \alpha^2 & 1 \\ 0 & 1 & 0 & 0 & \alpha \\ 0 & 0 & 1 & 1 & \alpha^2 \end{pmatrix}.$$

The dual code $\mathscr{C}^\perp$ has dimension 3 and minimum distance 2. We can apply again Proposition 2.26 to obtain a parity-check matrix of $\mathscr{C}^\perp$:

$$G'' = \begin{pmatrix} \alpha^2 & 0 & 1 & 1 & 0 \\ 1 & \alpha & \alpha^2 & 0 & 1 \end{pmatrix}.$$

Finally, observe that G'' has reduced row-echelon form G', as one expects.

Exercise 2.66. Let $S \subseteq \{1, ..., n\}$ be a set and recall Notation 2.47. Show that $\mathbb{F}_q^n(S)^\perp = \mathbb{F}_q^n(S^c)$, where $\mathbb{F}_q^n$ is viewed as a code and S is the complement of S in $\{1, ..., n\}$.

2.8 Equivalence of Linear Codes

A ubiquitous question in mathematics is: "When are two objects *essentially* the same?" In coding theory we ask ourselves: "When are codes $\mathscr{C}$ and $\mathscr{D}$ essentially the same code?". In this section we only study this question for linear codes.

Definition 2.67. A (**Hamming**) **linear isometry** is an $\mathbb{F}_q$-linear map $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ that preserves the Hamming weight, i.e., such that $\omega^{\mathrm{H}}(x) = \omega^{\mathrm{H}}(f(x))$ for all $x \in \mathbb{F}_q^n$.

Example 2.68. The binary codes $\mathscr{C}, \mathscr{D} \leq \mathbb{F}_q^n$ generated by

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad \text{and} \quad G' = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix},$$

respectively, are different but equivalent. Indeed, take the map $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ defined by $f : (x_1, x_2, x_3, x_4, x_5) \mapsto (x_4, x_3, x_2, x_1, x_5)$. It is easy to check that f is a linear isometry and that $f(\mathscr{C}) = \mathscr{D}$.

The next result describes some properties of linear isometries. The proof is left as an exercise.

Proposition 2.69.
1. Every linear isometry f is an $\mathbb{F}_q$-isomorphism $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$.
2. If $f, g : \mathbb{F}_q^n \to \mathbb{F}_q^n$ are linear isometries, then so is $f \circ g$.
3. If f is a linear isometry, then f^{-1} is a linear isometry as well.

The previous result implies that linear isometries form a group with respect to composition of functions. The group identity is the identity map.

Definition 2.70. We say that linear codes $\mathscr{C}, \mathscr{D} \leq \mathbb{F}_q^n$ are **equivalent** if there exists a linear isometry f such that $f(\mathscr{C}) = \mathscr{D}$. In such a case we write $\mathscr{C} \equiv \mathscr{D}$.

Note that, by Proposition 2.69, code equivalence is indeed an equivalence relation. Moreover, equivalent codes have the same parameters, as one expects.

Corollary 2.71. Let $\mathscr{C}, \mathscr{D} \leq \mathbb{F}_q^n$ be equivalent codes. Then $\mathscr{C}$ and $\mathscr{D}$ have the same dimension and the same minimum distance.

In Example 2.68 we encountered a linear isometry that acts on vectors by permuting their entries. We now introduce two general classes of linear isometries.

Notation 2.72.
1. For a permutation τ of the set $\{1, ..., n\}$, we let $f_\tau : \mathbb{F}_q^n \to \mathbb{F}_q^n$ be the map defined by $f_\tau(x_1, ..., x_n) := (x_{\tau(1)}, ..., x_{\tau(n)})$ for all $x \in \mathbb{F}_q^n$.
2. For a vector $\lambda \in \mathbb{F}_q^n$ with $\lambda_i \neq 0$ for all $1 \leq i \leq n$, we let $f_\lambda : \mathbb{F}_q^n \to \mathbb{F}_q^n$ be the map defined by $f_\lambda(x_1, ..., x_n) := (\lambda_1 x_1, ..., \lambda_n x_n)$ for all $x \in \mathbb{F}_q^n$.

It is easy to check that the maps introduced in Notation 2.72 are linear isometries. The following result shows that they generate the entire group of linear isometries $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$.

Theorem 2.73. Let $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ be a map. The following are equivalent:

1. f is a linear isometry,
2. $f = f_\lambda \circ f_\tau$ for some $\lambda \in \mathbb{F}_q^n$ with $\lambda_i \neq 0$ for all $1 \leq i \leq n$ and some permutation τ of the set $\{1, ..., n\}$.

Proof. We only need to show that 1 implies 2, as the other direction is immediate. Suppose that f is a linear isometry. Let $\{e_1, ..., e_n\}$ be the canonical basis of $\mathbb{F}_q^n$. Since f preserves the Hamming weight, for all $1 \leq i \leq n$ there exists $j_i \in \{1, ..., n\}$ and $\lambda_{j_i} \in \mathbb{F}_q \setminus \{0\}$ with $f(e_i) = \lambda_{j_i} e_{j_i}$.

We claim that the map $i \mapsto j_i$ is a permutation, say τ, of $\{1, ..., n\}$. Indeed, if $i \neq i'$ and $j_i = j_{i'}$, then $f(e_i + e_{i'})$ is a multiple of e_{j_i}. This contradicts the fact that f preserves the Hamming weight, because $e_i + e_{i'}$ has weight 2, while $f(e_i + e_{i'})$ has weight 1 or 0. Therefore the map $i \mapsto j_i$ is injective, and thus surjective (i.e., a permutation).

Now take any vector vector $x \in \mathbb{F}_q^n$. By definition, $x = \sum_{i=1}^n x_i e_i$. By the linearity of f we have

$$f(x) = \sum_{i=1}^{n} x_i \lambda_{j_i} e_{j_i}. \tag{2.3}$$

Next, let $\lambda \in \mathbb{F}_q^n$ be the vector whose j_i-th component is λ_{j_i}. Notice that this uniquely defines λ precisely because τ is a permutation. We have

$$(f_\lambda \circ f_\tau)(x) = (f_\lambda \circ f_\tau)\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i (f_\lambda \circ f_\tau)(e_i) = \sum_{i=1}^{n} x_i \lambda_{j_i} e_{j_i}. \tag{2.4}$$

Therefore comparing (2.3) and (2.4) we conclude that $f(x) = (f_\lambda \circ f_\tau)(x)$. Since $x \in \mathbb{F}_q^n$ was arbitrary, we have that $f = f_\lambda \circ f_\tau$ as functions. □

Exercise 2.74 (solved in Appendix C)**.** Show that the group of linear isometries $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ has order $(q-1)^n n!$.

A natural question is whether code equivalence is "compatible" with duality. The answer to this question is affirmative.

Proposition 2.75. Let $\mathscr{C}, \mathscr{D} \leq \mathbb{F}_q^n$ be linear codes. Suppose that $\mathscr{D} = (f_\lambda \circ f_\tau)(\mathscr{C})$ for some $\lambda \in \mathbb{F}_q^n$ with $\lambda_i \neq 0$ for all $1 \leq i \leq n$ and some permutation τ of the set $\{1, ..., n\}$. Then $\mathscr{D}^\perp = (f_{1/\lambda} \circ f_\tau)(\mathscr{C}^\perp)$, where $1/\lambda = (1/\lambda_1, ..., 1/\lambda_n)$.

Proof. Exercise. □

Corollary 2.76. Let $\mathscr{C}, \mathscr{D} \leq \mathbb{F}_q^n$ be linear codes. Then $\mathscr{C} \equiv \mathscr{D}$ if and only if $\mathscr{C}^\perp \equiv \mathscr{D}^\perp$.

2.9 Information Sets

A useful concept in the analysis of linear codes is the following.

Definition 2.77. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code of dimension $k \geq 1$. A non-empty set $S \subseteq \{1, ..., n\}$ is called an **information set** for $\mathscr{C}$ if $\pi_S(\mathscr{C})$ has dimension k.

Example 2.78. Let $\mathscr{C} \leq \mathbb{F}_3^5$ be the code generated by the matrix

$$G = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 \\ 2 & 2 & 1 & 1 & 2 \end{pmatrix}.$$

Then $\{1, 2\}$, $\{1, 3\}$ and $\{1, 2, 3\}$ are information sets, while $\{2\}$, $\{2, 3\}$ are not.

The terminology "information set" is motivated by the following property: if S is an information set for $\mathscr{C}$ and $x \in \mathscr{C}$, then the components $(x_i \mid i \in S)$ contain all the information to uniquely determine x. More formally, the following properties of information sets hold. They can be established using elementary linear algebra and are therefore left as an exercise (have a look at Example 2.80 before attempting to prove them).

Proposition 2.79. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $k \geq 1$.

1. $\mathscr{C}$ has an information set of cardinality k.
2. Every information set S for $\mathscr{C}$ has cardinality at least k.
3. Every information set S for $\mathscr{C}$ contains an information set for $\mathscr{C}$ of cardinality exactly k (we call such an information set **minimal**).
4. Let G be any generator matrix of $\mathscr{C}$. Then $S \subseteq \{1, ..., n\}$ is an information set for $\mathscr{C}$ if and only if the columns of G indexed by S form a matrix of rank k.
5. $\mathscr{C}$ is equivalent to a code having $\{1, ..., k\}$ as an information set (such a code is called **systematic**).
6. Suppose that S is an information set for $\mathscr{C}$, and let $x \in \mathscr{C}$. If $y \in \mathscr{C}$ is a codeword with $y_i = x_i$ for all $i \in S$, then $x = y$.
7. Suppose that S is a minimal information set for $\mathscr{C}$, and let $v \in \mathbb{F}_q^k$. Then there exists a unique $x \in \mathscr{C}$ with $\pi_S(x) = v$.

Proof. Exercise (*Hint*: look first at Example 2.80). □

Example 2.80. Consider the code $\mathscr{C} \leq \mathbb{F}_3^5$ generated by the matrix

$$G = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 \\ 2 & 1 & 1 & 2 & 1 \end{pmatrix}.$$

Then $\mathscr{C}$ has dimension $k = 2$. There are $\binom{5}{2} = 10$ submatrices of G of size 2×2, indexed by the subsets $S \subseteq \{1, ..., 5\}$ of cardinality 2. The submatrix corresponding to the set $\{2, 4\}$ has rank 2, because

$$\det \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix} = 1 \neq 0.$$

Therefore we can perform elementary row operations on G and obtain a new matrix, say G', that has an identity 2×2 matrix in block $\{2, 4\}$. More precisely, we can divide both rows of G by 2, then subtract the first row to the second row, and finally divide the second row by 2. This yields

$$G' = \begin{pmatrix} 2 & 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

The rows of G' are a basis of $\mathscr{C}$. Therefore

$$\mathscr{C} = \{(2\alpha, \alpha, 2\alpha + \beta, \beta, 2\alpha + \beta) \mid \alpha, \beta \in \mathbb{F}_3\}. \tag{2.5}$$

In particular, the restriction of $\pi_{\{2,4\}}$ to $\mathscr{C}$ is a bijection, showing that part 7 of Proposition 2.79 holds.

The code $\mathscr{C}$ is not systematic, as $\{1, 2\}$ is not an information set for $\mathscr{C}$. Let τ be any permutation of $\{1, ..., 5\}$ with $\tau(1) = 2$ and $\tau(2) = 4$. Directly from 2.5 we obtain

$$f_\tau(\mathscr{C}) = \{(\alpha, \beta, ..., ..., ...) \mid \alpha, \beta \in \mathbb{F}_3\},$$

where the missing entries depend on the specific choice of τ. Therefore $\{1, 2\} = \tau^{-1}(\{2, 4\})$ is an information set for $f_\tau(\mathscr{C})$.

2.10 Other Exercises

Exercise **2.81.** Let $f : \mathbb{F}_2^3 \to \mathbb{F}_2^6$ be the map defined by

$$f(x, y, z) := (x, y, z, x + y, y + z, x + z)$$

for all $(x, y, z) \in \mathbb{F}_3^3$. Let $\mathscr{C} \subseteq \mathbb{F}_3^6$ be the image of f.

1. Show that $\mathscr{C}$ is a linear code and compute its dimension.
2. Write down a generator and a parity-check matrix of $\mathscr{C}$.
3. Compute the minimum distance of $\mathscr{C}$ by applying Corollary 2.33.

Exercise **2.82.** Suppose $q = n$ and let $\mathscr{C} = \{x \in \mathbb{F}_q^n \mid \{x_1, ..., x_n\} = \{1, ..., n\}\}$.

1. Show that $\mathscr{C}$ is a non-linear code.
2. Compute the cardinality and the minimum distance of $\mathscr{C}$.

3. Show that for every $x \in \mathscr{C}$ and for every error vector $e \in \mathbb{F}_q^n$ of weight 1 there are exactly two codewords of $\mathscr{C}$ at distance 1 from $x + e$.

Exercise **2.83.** Let $\mathscr{C} \leq \mathbb{F}_2^5$ be the code generated by

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

1. Write G in standard form and compute a parity-check matrix of $\mathscr{C}$.
2. Compute the length, the dimension and the minimum distance of $\mathscr{C}$.
3. Compute a coset leader for each equivalence class of $\mathbb{F}_2^5$ modulo the code.
4. Use syndrome decoding to correct the following messages containing at most one error:
$$y_1 = (1,1,1,1,1), \qquad y_2 = (0,1,1,1,0), \qquad y_3 = (1,1,0,1,1).$$

Exercise **2.84** (solved in Appendix C)**.** Let $\mathscr{C} \leq \mathbb{F}_2^n$ be a binary code of dimension $1 \leq k \leq n-1$ and $d^{\mathrm{H}}(\mathscr{C}) \geq 3$. Let H be a parity-check matrix of $\mathscr{C}$.

1. Show that the columns of H are distinct.
2. Assume that $x \in \mathscr{C}$ is sent and $y \in \mathbb{F}_2^5$ is received, and that exactly one error occured in the transmission (i.e., there is a unique $1 \leq i \leq n$ with $x_i \neq y_i$). Show that i can be retrieved by looking at H and at the product $H \cdot y^\top$.

Exercise **2.85.** Let $\mathscr{C} \leq \mathbb{F}_2^6$ be the binary code defined by the generator matrix

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

1. Compute the parameters of $\mathscr{C}$ (length, dimension and minimum distance).
2. Compute the parameters of the extended code $\mathscr{C}^{\mathrm{ext}}$ and write down a parity-check matrix of it.

Exercise **2.86.** Show the following linear version of the Gilbert-Varshamov bound. Let $1 \leq d \leq n$ and let k be the smallest integer with

$$q^k > \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i}.$$

Then $k \geq 1$ and there exists a linear code $\mathscr{C} \leq \mathbb{F}_q^n$ of dimension k and minimum distance at least d.

Exercise **2.87.** Consider the code $\mathscr{D} := (\mathscr{C}^{\mathrm{ext}})^\perp$, where $\mathscr{C}$ is the code of Exercise 2.85. Show that the shortening $\pi_S(\mathscr{D}(S))$ has minimum distance at least 3 for all $S \subseteq \{1, ..., n\}$ with $|S| = 5$.

Exercise **2.88.** Let q be not a power of 2 and let $\mathscr{C}, \mathscr{D} \leq \mathbb{F}_q^n$ be non-zero linear codes of minimum distances d_1 and d_2, respectively. Define the code

$$\mathscr{E} := \{(x, x+y, x-y) \mid x \in \mathscr{C},\, y \in \mathscr{D}\} \leq \mathbb{F}_q^{3n}.$$

Show that $\mathscr{E}$ has dimension $\dim(\mathscr{C})+\dim(\mathscr{D})$ and minimum distance exactly $\min\{3d_1, 2d_2\}$.

Then show that the assumption "q is not a power of 2" is indeed necessary, exhibiting codes $\mathscr{C}, \mathscr{D} \leq \mathbb{F}_2^n$ for which $\mathscr{E}$ has minimum distance strictly smaller than $\min\{3d_1, 2d_2\}$.

Exercise **2.89** (solved in Appendix C)**.** Is there a linear code having weight enumerator $1Y^9 + 14X^3Y^6 + 16X^4Z^5 + 5X^7Y^2 + 10X^8Y + 9X^9$? Justify your answer.

Exercise **2.90.** For some n and some q of your choice (not all parameters will work):

1. construct a linear code $\mathscr{C} \leq \mathbb{F}_q^n$ with $\mathscr{C} \cap \mathscr{C}^\perp = \{0\}$;
2. construct a linear code $\mathscr{C} \leq \mathbb{F}_q^n$ with $\mathscr{C} \cap \mathscr{C}^\perp \neq \{0\}$;
3. construct a linear code $\mathscr{C} \leq \mathbb{F}_q^n$ with $\mathscr{C} = \mathscr{C}^\perp$.

Exercise **2.91** (solved in Appendix C)**.** Let $\mathscr{C} \leq \mathbb{F}_2^n$ be any linear binary code. Show the following statement: Either all the codewords of $\mathscr{C}$ have even weight, or exactly half of the codewords of $\mathscr{C}$ have even weight.

Exercise **2.92.** A code $\mathscr{C} \leq \mathbb{F}_q^n$ is called self-dual if $\mathscr{C}^\perp = \mathscr{C}$.

1. Show that $\mathscr{C} \leq \mathbb{F}_q^n$ is self-dual then n is even and the dimension of $\mathscr{C}$ is $n/2$.
2. Show that if $\mathscr{C} \leq \mathbb{F}_2^n$ is a binary self-dual code, then the all-1 vector $(1, 1, ..., 1)$ belongs to $\mathscr{C}$.

Exercise **2.93.** Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code of dimension $1 \leq k < n$.

1. Let τ be a permutation of $\{1, ..., n\}$. Show that a non-empty set $S \subseteq \{1, ..., n\}$ is an information set for $\mathscr{C}$ if and only if $\tau^{-1}(S)$ is an information set for $f_\tau(\mathscr{C})$.
2. Show that S is a minimal information set for $\mathscr{C}$ if and only if the complement S^c is a minimal information set for $\mathscr{C}^\perp$.

Exercise **2.94.** Recall that a **permutation matrix** over a field $\mathbb{F}_q$ is a square matrix P with the following properties:

- in each row of P there is exactly one non-zero entry, and that entry is a 1;
- in each column of P there is exactly one non-zero entry, and that entry is a 1.

1. Show that every permutation matrix P is invertible and that P^{-1} is a permutation matrix as well.
2. Show that for all permutation matrices $P \in \mathbb{F}_q^{n \times n}$ and all vectors $x \in \mathbb{F}_q^n$ we have $\omega^{\mathrm{H}}(x) = \omega^{\mathrm{H}}(x \cdot P)$.
3. Let $\mathscr{C}, \mathscr{C}' \leq \mathbb{F}_q^n$ be non-zero linear codes of the same dimension, say k. Show that the following are equivalent:

- $\mathscr{C}' = f_\tau(\mathscr{C})$ for some permutation τ of $\{1, ..., n\}$;
- there exist generator matrices G and G' of $\mathscr{C}$ and $\mathscr{C}'$ (respectively) and a permutation matrix P of size $n \times n$ with $G' = GP$.

Chapter 3

Bounds

In this chapter we illustrate various techniques to obtain upper and lower bounds for a given code parameter in terms of the others. The best known bounds are named after Singleton and Hamming, and yield the classes of *maximum distance separable* and *perfect* codes, respectively.

3.1 The Singleton Bound and MDS Codes

In this section we present one of the most important bounds for the cardinality of an error-correcting code, namely, the *Singleton bound*.

Theorem 3.1 (Singleton bound)**.** Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code. We have $|\mathscr{C}| \leq q^{n-d+1}$, where $d = d^{\mathrm{H}}(\mathscr{C})$. In particular, if $\mathscr{C}$ is linear then $\dim(\mathscr{C}) \leq n - d^{\mathrm{H}}(\mathscr{C}) + 1$.

Proof. The result follows from the definitions if $|\mathscr{C}| = 1$. Now suppose $|\mathscr{C}| \geq 2$, and let $S := \{1, 2, ..., n-d+1\}$. We claim that the restriction of π_S to $\mathscr{C}$ is injective. To see this, suppose towards a contradiction that there exist $x, y \in \mathscr{C}$ with $x \neq y$ but $\pi_S(x) = \pi_S(y)$. Then $x_i = y_i$ for $i \in \{1, ..., n-d+1\}$ and thus $d^{\mathrm{H}}(x, y) \leq n - (n-d+1) = d-1$, a contradiction. □

Definition 3.2. A code $\mathscr{C} \subseteq \mathbb{F}_q^n$ whose parameters attain the bound of Theorem 3.1 is called **MDS** (**maximum distance separable**).

Note that the trivial code (Example 2.6) are MDS.

Example 3.3. Let q be not a power of 2. The code $\mathscr{C} \leq \mathbb{F}_q^4$ generated by

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

is MDS. Indeed, it has dimension 2 and length 4. To see that it has minimum distance 3, observe that a non-zero codeword of $\mathscr{C}$ has the form

$$x_{a,b} = (a, b, a+b, a-b) \quad \text{for some } a, b \in \mathbb{F}_q \text{ with } (a,b) \neq (0,0).$$

If $a = 0$ and $b \neq 0$ then $x_{a,b}$ has weight 3. The same is true if $b = 0$ and $a \neq 0$. If $a \neq 0$ and $b \neq 0$ then $a+b$ and $a-b$ cannot be both zero, as q is not a power of 2 by assumption. Therefore $x_{a,b}$ has weight at least 3.

Exercise **3.4.** Check that the repetition code of Example 2.7 is MDS. Show that $\mathbb{F}_q^n$ is the only MDS code of minimum distance equal to 1 and length n.

Exercise **3.5.** Check that the code of Example 2.21 is MDS.

MDS codes are a central topic in coding theory. They have also been investigated in connection with several topics in combinatorics (combinatorial designs, hyperplane arrangements, posets, ...).

The next result gives a criterion to check if a code is MDS from a generator matrix.

Proposition 3.6. Let $k \geq 1$ and let G a $k \times n$ matrix over $\mathbb{F}_q$. Then G is the generator matrix of an MDS code if and only if all $k \times k$ minors of G are non-zero.

Proof. Suppose that G has a zero $k \times k$ minor, say the one corresponding to the columns indexed by a set S. Then there exists a non-zero linear combination of the rows of G, say $x \in \mathbb{F}_q^n$, whose Hamming support is contained in the complement of S. In particular, G generates a code $\mathscr{C}$ containing x. As $|S| = k$, we have $1 \leq \omega^{\mathrm{H}}(x) \leq n - k$ and $\mathscr{C}$ cannot be an MDS code.

Now suppose that all $k \times k$ minors of G are non-zero. Let $\mathscr{C}$ be the code generated by G, and suppose towards a contradiction that there exists a non-zero codeword $x \in \mathscr{C}$ with $\omega^{\mathrm{H}}(x) \leq n - k$. Let $S \subseteq \{1, ..., n\}$ be a set of size k and such that $x_i = 0$ for all $i \in S$. It is easy to see that the $k \times k$ minor of G whose columns are indexed by S is zero, a contradiction. □

A very natural question at this point is whether MDS codes exist or not. Finding all the triples (q, n, d) for which there exists an MDS code $\mathscr{C} \leq \mathbb{F}_q^n$ of dimension k is still an open problem in coding theory. However, it is very well-known that MDS codes exist over finite fields of sufficiently large cardinality with respect to n. The following result provides an explicit construction.

Proposition 3.7. Suppose $1 \leq k \leq n \leq q$. Let $\alpha_1, ..., \alpha_n$ be distinct elements of $\mathbb{F}_q$ and let

$$G = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & & & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_n^{k-1} \end{pmatrix} \in \mathbb{F}_q^{k \times n}.$$

Then G generates an MDS code of dimension k.

Proof. Fix any set $S \subseteq \{1, ..., n\}$ of cardinality k. The submatrix of G made of the columns indexed by S, say G_S, is a so-called *Vandermonde matrix*. It is known that the determinant of such a matrix is

$$\det(G_S) = \prod_{\substack{i,j \in S \\ i<j}} (\alpha_i - \alpha_j),$$

which is a non-zero element of $\mathbb{F}_q$ as the α_i's are distinct by assumption. Since S was arbitrary with $|S| = k$, the matrix G generates an MDS code by Proposition 3.6. □

Example 3.8. Let $\mathbb{F}_9 = \mathbb{F}_3[\beta]$, where $\beta^2 + 2\beta + 2 = 0$. Take $n = 5$ and consider the distinct elements $0, \beta, \beta^2, \beta^3, \beta^4$ of $\mathbb{F}_9$. The matrix

$$G = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & \beta & \beta^2 & \beta^3 & \beta^4 \\ 0 & \beta^2 & \beta^4 & \beta^6 & \beta^8 \end{pmatrix}$$

generates an MDS code of dimension 3 and minimum distance 3 as well.

Exercise **3.9.** Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $k \geq 1$. Show that the following are equivalent:

1. $\mathscr{C}$ is MDS,
2. every subset $S \subseteq \{1, ..., n\}$ with $|S| = k$ is an information set for $\mathscr{C}$ (see Definition 2.77).

Exercise **3.10.** Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code. Show that $d^{\mathrm{H}}(\mathscr{C}) + d^{\mathrm{H}}(\mathscr{C}^{\perp}) \leq n + 2$.

3.2 The Hamming Bound and Perfect Codes

Another famous bound for the size of an error-correcting code is the *Hamming bound*. Its proof uses an argument known as *sphere-packing*.

Theorem 3.11 (Hamming bound)**.** Let $d \geq 1$ and let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code with $|\mathscr{C}| \geq 2$ and $d^{\mathrm{H}}(\mathscr{C}) \geq d$. Define $t = \lfloor (d-1)/2 \rfloor$. We have

$$|\mathscr{C}| \leq \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i} (q-1)^i}.$$

Proof. The Hamming balls of radius t centered at the codewords of $\mathscr{C}$ are pairwise disjoint. Indeed, if there existed $x, x' \in \mathscr{C}$ and $y \in \mathbb{F}_q^n$ with $y \in B_t^{\mathrm{H}}(x) \cap B_t^{\mathrm{H}}(x')$, then by Proposition 1.15 we would have $d^{\mathrm{H}}(x, x') \leq d^{\mathrm{H}}(x, y) + d^{\mathrm{H}}(y, x') \leq 2t < d \leq d^{\mathrm{H}}(\mathscr{C})$, a

contradiction. We therefore conclude

$$\sum_{x\in\mathscr{C}} |B_t^{\mathrm{H}}(x)| = \left|\bigcup_{x\in\mathscr{C}} B_t^{\mathrm{H}}(x)\right| \leq |\mathbb{F}_q^n| = q^n.$$

By Proposition 2.11 we have

$$\sum_{x\in\mathscr{C}} |B_t^{\mathrm{H}}(x)| = |\mathscr{C}| \sum_{i=0}^{t} \binom{n}{i}(q-1)^i,$$

and the result follows. □

Definition 3.12. A code $\mathscr{C} \subseteq \mathbb{F}_q^n$ with $|\mathscr{C}| \geq 2$ and whose parameters meet the Hamming bound of Theorem 3.11 is called **perfect**.

The proof of the Hamming bound also gives us the following characterization of perfect codes.

Proposition 3.13. Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code with $|\mathscr{C}| \geq 2$, and let $t = \lfloor (d^{\mathrm{H}}(\mathscr{C}) - 1)/2 \rfloor$. The following are equivalent:

1. $\mathscr{C}$ is perfect,
2. for all $y \in \mathbb{F}_q^n$ there exists a unique $x \in \mathscr{C}$ with $y \in B_t^{\mathrm{H}}(x)$.

In other words, in a perfect code of minimum distance d the Hamming balls centered at the codewords and with radius $t = \lfloor (d-1)/2 \rfloor$ are pairwise disjoint and cover the entire space $\mathbb{F}_q^n$.

Example 3.14. Suppose that n is odd and let $\mathscr{C} \leq \mathbb{F}_2^n$ be the n-times binary repetition code. The minimum distance of $\mathscr{C}$ is $d = n = 2t + 1$ and therefore

$$\sum_{i=0}^{t} \binom{n}{i}(q-1)^i = \sum_{i=0}^{t} \binom{n}{i}. \tag{3.1}$$

The quantity in (3.1), say ε, is the number of subsets of $\{1, ..., n\}$ of cardinality at most t. These are in bijection with the subsets of $\{1, ..., n\}$ of cardinality at least $n - t$. Since $n - t = t + 1$, we have

$$\varepsilon + \varepsilon = 2^n,$$

from which $\varepsilon = 2^{n-1}$. Therefore $\mathscr{C}$ is a perfect code, since

$$|\mathscr{C}| = 2 = \frac{2^n}{2^{n-1}} = \frac{2^n}{\sum_{i=0}^{t} \binom{n}{i}}.$$

Another important example is the following.

Example 3.15. The binary code $\mathscr{G}_{23} \leq \mathbb{F}_2^{23}$ with generator matrix

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

is perfect. It has dimension 12 and minimum distance 7. It is a famous object in coding theory, known as the **binary Golay code**. The extension of the Golay code (Definition 2.54) has length 24, dimension 12, and minimum distance 8. It has been used by the spacecrafts *Voyager 1* and *Voyager 2* to transmit pictures of Jupiter and Saturn in 1979, 1980, and 1981.

It is interesting to observe that there are no perfect codes $\mathscr{C} \subseteq \mathbb{F}_q^n$ with $|\mathscr{C}| \geq 2$ and even minimum distance.

Proposition 3.16. Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code with $|\mathscr{C}| \geq 2$ and with even minimum distance. Then $\mathscr{C}$ is not perfect.

Proof. Let $d = 2s$ be the minimum distance of the code $\mathscr{C}$. Note first that $s \geq 2$ and that $t := \lfloor (d-1)/2 \rfloor = s - 1$. Take $x \in \mathscr{C}$ and pick any vector $y \in \mathbb{F}_q^n$ that is at distance exactly s from x.

We claim that there is no codeword $z \in \mathscr{C}$ with $y \in B_t^{\mathrm{H}}(z)$. By definition, $z \notin B_t^{\mathrm{H}}(x)$. Now suppose by contradiction that there exists $x' \in \mathscr{C}$ with $x' \neq x$ and $y \in B^{\mathrm{H}}(x')$. By the triangular inequality we would have

$$d^{\mathrm{H}}(x', x) \leq d^{\mathrm{H}}(x', y) + d^{\mathrm{H}}(y, x) \leq s - 1 + s = d - 1,$$

contradicting the fact that $\mathscr{C}$ has minimum distance d. So

$$y \notin \bigcup_{x \in \mathscr{C}} B_t^{\mathrm{H}}(x),$$

as claimed.

Finally, the claim implies

$$\left| \bigcup_{x \in \mathscr{C}} B_t^{\mathrm{H}}(x) \right| < q^n,$$

i.e. (as in the proof of Theorem 3.11),

$$|\mathscr{C}| \sum_{i=0}^{t} \binom{n}{i} (q-1)^i < q^n.$$

Therefore $\mathscr{C}$ is not perfect. □

There is a family of codes that are all perfect and have minimum distance 3. They are called *Hamming codes* are defined as follows.

Definition 3.17. Suppse $r \geq 3$ and let H be a matrix of size $r \times (q^r - 1)/(q-1)$ whose columns are all the non-zero vectors of $\mathbb{F}_q^r$ up to non-zero scalar multiples. The code having H as parity-check matrix is called a q-ary **Hamming code of redundancy** r.

Example 3.18. The binary code $\mathscr{C}$ with

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

as parity-check matrix is a Hamming code of redundancy 3. Since H has size 3×7, the code $\mathscr{C}$ has dimension $7 - 3 = 4$.

Proposition 3.19. Let $r \geq 3$ and let $\mathscr{C} \leq \mathbb{F}_q^n$ be a q-ary Hamming code of redundancy r, with $n = (q^r - 1)/(q-1)$. Then $\mathscr{C}$ has minimum distance 3 and dimension $n - r$. In particular, $\mathscr{C}$ is perfect.

Proof. The dimension of $\mathscr{C}$ is clearly $n - r$. Every two columns of H are linearly independent. Moreover, it is easy to see that there exist three columns of H that are linearly dependent. Therefore $\mathscr{C}$ has minimum distance exactly 3 by Corollary 2.33.

Since $\mathscr{C}$ has dimension $n - r$ and $n = (q^r - 1)/(q-1)$, the Hamming bound (Theorem 3.11) reads

$$q^{n-r} \leq \frac{q^n}{1 + n(q-1)} = \frac{q^n}{q^r} = q^{n-r}$$

and is therefore met with equality. □

Exercise **3.20.** Compute the weight distribution of the dual of the Hamming code $\mathscr{C}$ constructed in Example 3.18.

3.3 The Griesmer Bound

The Griesmer bound is a lower bound for the length of a linear code as a function of its dimension and minimum distance. The statement is as follows.

Theorem 3.21 (Griesmer bound)**.** Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code of dimension $k \geq 1$. Let d denote the minimum distance of $\mathscr{C}$. Then

$$n \geq \sum_{i=0}^{k-1} \lceil d/q^i \rceil.$$

The proof of the Griesmer bound relies on a specific code construction and its properties, which we exhamine first (recall Notation 2.48).

Definition 3.22. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code and let $x \in \mathscr{C}$ be a codeword of Hamming weight $w < n$. The **residual code** $\mathrm{Res}(\mathscr{C}, x)$ of $\mathscr{C}$ with respect to x is $\pi_S(\mathscr{C}) \leq \mathbb{F}_q^{n-w}$, where $S = \{1, ..., n\} \setminus \sigma^{\mathrm{H}}(x)$.

Note that $\mathrm{Res}(\mathscr{C}, x)$ does not depend on x itself but only on its support. Different choices of x might give the same residual code.

We will need the following preliminary result.

Proposition 3.23. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code and let $x \in \mathscr{C}$. Suppose that $\mathscr{C}$ has dimension $k \geq 2$ and that x has Hamming weight $1 \leq w < \min\{n, qd/(q-1)\}$, where d is the minimum distance of $\mathscr{C}$. Then the code $\mathrm{Res}(\mathscr{C}, x)$ has dimension $k-1$ and minimum distance at least $d - w + \lceil w/q \rceil$.

Remark 3.24. Following the notation of Proposition 3.23, it is not true in general that $qd/(q-1) \leq n$. For example, the code generated over $\mathbb{F}_3$ by

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

has dimension $k = 2$ and minimum distance $d = 5$. Therefore $3d = 15 > 14 = 2n$.

Proof. By Corollary 2.71 we can assume without loss of generality that x has the form $x = (1, ...1, 0, ..., 0)$. Let $S \subseteq \{1, ..., n\}$ denote the complement of the Hamming support of w, i.e., $S = \{w+1, ..., n\}$. Note that $S \neq \emptyset$ as $w < n$ by assumption. Moreover, $\mathrm{Res}(\mathscr{C}, x) = \pi_S(\mathscr{C})$ by definition of residual code. We need to prove that $\pi_S(\mathscr{C})$ has dimension $k-1$ and minimum distance at least $d - w + \lceil w/q \rceil$.

The restriction of the projection π_S to $\mathscr{C}$ is not injective, as $x \neq 0$ but $\pi_S(x) = 0$. Therefore $\pi_S(\mathscr{C})$ has dimension at most $k-1$. Suppose towards a contradiction that its dimension is strictly smaller than $k-1$. Then there must exist a non-zero $y \in \mathscr{C}$ that is not a multiple of x and such that $\pi_S(y) = 0$. Using a counting argument one sees that there is a field element $\alpha \in \mathbb{F}_q$ with

$$|\{1 \leq i \leq w \mid y_i = \alpha\}| \geq w/q.$$

Therefore the codeword $y - \alpha x$ is non-zero and has Hamming weight at most $w - w/q = w(q-1)/q$, from which $d \leq w(q-1)/q$. This contradicts our assumption on w and shows that $\pi_S(\mathscr{C})$ has dimension *exactly* $k-1$.

We now turn to the minimum distance of $\pi_S(\mathscr{C})$. Since $k \geq 2$ by assumption, $\pi_S(\mathscr{C}) \leq \mathbb{F}_q^{n-w}$ is not the zero code. Fix any non-zero vector $(z_{w+1}, ..., z_n) \in \pi_S(\mathscr{C})$ and let $z = (z_1, ..., z_w, z_{w+1}, ..., z_n) \in \mathscr{C}$. Note that z cannot be a multiple of x. Arguing as before, there exists $\alpha \in \mathbb{F}_q$ with

$$|\{1 \leq i \leq w \mid z_i = \alpha\}| \geq w/q.$$

Thus $z - \alpha x \in \mathscr{C}$ is a non-zero vector and has weight at most $w - w/q + \omega^{\mathrm{H}}(z_{w+1}, ..., z_n)$. In particular, we must have $d \leq w - w/q + \omega^{\mathrm{H}}(z_{w+1}, ..., z_n)$. This implies $\omega^{\mathrm{H}}(z_{w+1}, ..., z_n) \geq d - w + \lceil w/q \rceil$ and concludes the proof. □

We can finally establish the Griesmer bound combining Proposition 3.23 with an inductive argument.

Proof of Theorem 3.21. We proceed by induction on $k \geq 1$. The result is immediate if $k = 1$, as the bound reads $n \geq d$. Now suppose $k \geq 2$ (which automatically implies $d < n$) and apply Proposition 3.23 to a minimum weight codeword $x \in \mathscr{C}$ (note that $d < qd/(q-1)$ as well). We obtain the code $\mathrm{Res}(\mathscr{C}, x)$ of dimension $k-1$ and minimum distance $d' \geq \lceil d/q \rceil$. Applying the induction hypothesis to $\mathrm{Res}(\mathscr{C}, x)$ we get

$$n - d \geq \sum_{i=0}^{k-2} \lceil d'/q^i \rceil \geq \sum_{i=0}^{k-2} \lceil d/q^{i+1} \rceil,$$

from which

$$n \geq d + \sum_{i=0}^{k-2} \lceil d/q^{i+1} \rceil = d + \sum_{i=1}^{k-1} \lceil d/q^i \rceil = \sum_{i=0}^{k-1} \lceil d/q^i \rceil.$$

This concludes the proof. □

Example 3.25. The **ternary Golay code** $\mathscr{G}_{11} \leq \mathbb{F}_3^{11}$ is the code over $\mathbb{F}_3$ with parity-check matrix

$$H = \begin{pmatrix} 1 & 1 & 1 & 2 & 2 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 2 & 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 2 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 2 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 2 & 2 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{F}_3^{5 \times 11}.$$

It has length 11, dimension $11 - 5 = 6$ and minimum distance 5. The Griesmer bound for $\mathscr{C}$ reads

$$11 \geq \sum_{i=0}^{5} \lceil 5/3^i \rceil = 11.$$

Therefore $\mathscr{G}_{11}$ meets it with equality. The **extended ternary Golay code** is $\mathscr{G}_{11}^{\mathrm{ext}} \leq \mathbb{F}_3^{12}$; see Definition 2.54. It can be checked that it has length 12, dimension 6 and minimum distance 6. The Griesmer bound for $\mathscr{G}_{11}^{\mathrm{ext}}$ reads

$$12 \geq \sum_{i=0}^{5} \lceil 6/3^i \rceil = 12.$$

Therefore $\mathscr{G}_{11}^{\mathrm{ext}}$ also meets the Griesmer bound.

Exercise **3.26.** Use the statement of the Griesmer bound to establish the Singleton bound for linear codes $\mathscr{C} \leq \mathbb{F}_q^n$ of dimension $k \geq 1$.

3.4 The Plotkin Bound

In this section we prove a bound on the cardinality of a possibly non-linear code. It is called *Plotkin bound* and its proof combines a double-counting argument with the following analysis result.

Lemma 3.27 (Cauchy-Schwarz inequality)**.** Let $\ell \geq 1$ be an integer and let $v, w \in \mathbb{R}^\ell$. We have

$$\left(\sum_{j=1}^{\ell} v_j^2\right) \cdot \left(\sum_{j=1}^{\ell} w_j^2\right) \geq \left(\sum_{j=1}^{\ell} v_j w_j\right)^2 .$$

Theorem 3.28 (Plotkin bound)**.** Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code with $|\mathscr{C}| \geq 2$. Denote by d the minimum distance of $\mathscr{C}$ and suppose that $qd > (q-1)n$. We have

$$|\mathscr{C}| \leq \left\lfloor \frac{qd}{qd-(q-1)n} \right\rfloor .$$

Proof. We will evaluate the sum

$$\Sigma := \sum_{x \in \mathscr{C}} \sum_{y \in \mathscr{C}} d^{\mathrm{H}}(x, y)$$

in two different ways. For ease of notation we denote by M the cardinality of $\mathscr{C}$. Since $\mathscr{C}$ has minimum distance d, we have

$$\Sigma \geq dM(M-1). \tag{3.2}$$

On the other hand, observe that for vectors $x, y \in \mathbb{F}_q^n$ one has $d^{\mathrm{H}}(x, y) = \sum_{i=1}^n \delta(x_i, y_i)$, where for all $\alpha, \beta \in \mathbb{F}_q$ we let $\delta(\alpha, \beta) := 1$ if $\alpha \neq \beta$, and $\delta(\alpha, \beta) := 0$ otherwise. Therefore

$$\Sigma = \sum_{i=1}^{n} \sum_{x \in \mathscr{C}} \sum_{y \in \mathscr{C}} \delta(x_i, y_i) = \sum_{i=1}^{n} \sum_{\alpha \in \mathbb{F}_q} \sum_{\substack{x \in \mathscr{C} \\ x_i = \alpha}} \sum_{y \in \mathscr{C}} \delta(\alpha, y_i).$$

For $\alpha \in \mathbb{F}_q$ and $i \in \{1, ..., n\}$, let $\varepsilon_{\alpha,i} := |\{x \in \mathscr{C} \mid x_i = \alpha\}|$. With this notation we can re-write Σ as

$$\Sigma \;=\; \sum_{i=1}^{n} \sum_{\alpha \in \mathbb{F}_q} |\{(x, y) \in \mathscr{C} \mid x_i = \alpha,\, y_i \neq \alpha\}|$$

$$= \sum_{i=1}^{n} \sum_{\alpha \in \mathbb{F}_q} \varepsilon_{\alpha,i}(M - \varepsilon_{\alpha_i})$$
$$= nM^2 - \sum_{i=1}^{n} \sum_{\alpha \in \mathbb{F}_q} \varepsilon_{\alpha,i}^2.$$

For all $i \in \{1, ..., n\}$ we now apply the Cauchy-Schwarz inequality to the vectors $v = (1, ..., 1) \in \mathbb{R}^q$ and $w = (\varepsilon_{\alpha,i} \mid \alpha \in \mathbb{F}_q) \in \mathbb{R}^q$, obtaining

$$q \sum_{\alpha \in \mathbb{F}_q} \varepsilon_{\alpha,i}^2 \geq \left(\sum_{\alpha \in \mathbb{F}_q} \varepsilon_{\alpha,i} \right)^2 = M^2.$$

Therefore

$$\Sigma \leq nM^2 - nM^2/q = nM^2(q-1)/q. \tag{3.3}$$

Combining (3.2) with (3.3) we obtain $dM(M-1) \leq nM^2(q-1)/q$. Dividing by M and re-arranging the terms one gets

$$M \leq \frac{qd}{qd - (q-1)n},$$

which implies the desired bound. □

We conclude the section by introducing a family of codes that meet the Plotkin bound.

Definition 3.29. Let $r \geq 3$ be an integer. A **simplex code** of dimension r is the dual of a Hamming code of redundancy r; see Definition 3.17.

Simplex codes have an extremely regular weight distribution.

Proposition 3.30. Let $\mathscr{S}_r$ be a simplex code of dimension r. Then all the non-zero codewords of $\mathscr{S}_r$ have Hamming weight exactly q^{r-1}.

Proof. By Theorem 2.64, $\mathscr{S}_r$ has a generator matrix whose columns are all the vectors of $\mathbb{F}_q^r$, up to non-zero scalar multiples. Fix any non-zero codeword $x \in \mathscr{S}_r$. Then there exists a non-zero vector $y \in \mathbb{F}_q^r$ with $x = y \cdot H$. Now observe that for every non-zero $y \in \mathbb{F}_q^r$ there are exactly $(q^{r-1} - 1)/(q-1)$ non-zero vectors $h \in \mathbb{F}_q^r$ with $\langle y, h \rangle = 0$, up to non-zero scalar multiples. Therefore there are precisely $(q^{r-1} - 1)/(q-1)$ columns h of H for which $\langle y, h \rangle = 0$. This shows that x has Hamming weight

$$\frac{q^r - 1}{q-1} - \frac{q^{r-1} - 1}{q-1} = q^{r-1},$$

as desired. □

Simplex codes meet the Plotking bound, as the following example shows.

Example 3.31. Let $\mathscr{S}_r$ be a simplex code of dimension r. By Proposition 3.30, $\mathscr{S}_r$ has minimum distance q^{r-1}. The Plotkin bound reads

$$|\mathscr{S}_r| \leq \left\lfloor \frac{q \cdot q^{r-1}}{q \cdot q^{r-1} - (q-1) \cdot \frac{q^r-1}{q-1}} \right\rfloor = q^r.$$

Since $\mathscr{S}_r$ has dimension r, the Plotkin bound is met with equality.

3.5 Other Exercises

Exercise 3.32.

- Find all the minimal information sets of the code $\mathscr{C} \leq \mathbb{F}_5^6$ generated by the matrix

$$G = \begin{pmatrix} 2 & 1 & 3 & 0 & 1 & 4 \\ 0 & 3 & 3 & 1 & 1 & 4 \\ 1 & 1 & 2 & 0 & 4 & 1 \end{pmatrix}.$$

- Is the code $\mathscr{C}$ MDS? Justify your answer.
- Compute a parity-check matrix H of $\mathscr{C}$.
- Find all the minimal information set of $\mathscr{C}^\perp$.
- Compute the minimum distance of $\mathscr{C}^\perp$.

Exercise 3.33.

1. Are there values of $\alpha, \beta, \gamma, \delta \in \mathbb{F}_2$ for which

$$G = \begin{pmatrix} 1 & 0 & \alpha & \beta \\ 0 & 1 & \gamma & \delta \end{pmatrix} \in \mathbb{F}_2^{2\times 4}$$

generates an MDS code? Justify your answer.

2. Are there values of $\alpha, \beta, \gamma, \delta \in \mathbb{F}_3$ for which

$$G = \begin{pmatrix} 1 & 0 & \alpha & \beta \\ 0 & 1 & \gamma & \delta \end{pmatrix} \in \mathbb{F}_3^{2\times 4}$$

generates an MDS code? Justify your answer.

Exercise 3.34. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be an MDS code with minimum distance $d \geq 3$ and dimension k.

1. Let $S = \{1, 2, ..., n-d+3\}$ and $\mathscr{C}' = \pi_S(\mathscr{C})$. Show that $|\mathscr{C}| = |\mathscr{C}'|$ and that $\mathscr{C}'$ has minimum distance at least 3.
2. Use the Hamming bound on $\mathscr{C}'$ to show that $k \leq q - 1$.

Exercise 3.35. Let $\mathscr{C}$ be the code of Example 3.18. Find all the minimal information sets of $\mathscr{C}$ and $\mathscr{C}^\perp$.

Exercise **3.36** (solved in Appendix C)**.** Use the Griesmer bound to show that any binary linear code $\mathscr{C} \leq \mathbb{F}_2^n$ with dimension $k \geq 2$ and minimum distance $d = 3$ satisfies $k \leq n - 3$. What would the Singleton bound say about k?

Exercise **3.37.** Consider the binary Hamming code $\mathscr{C}$ having

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

as a parity check matrix. Compute the parameters of the residual code of $\mathscr{C}$ with respect to the codeword $(1, 1, 0, 1, 0, 0, 0)$.

Exercise **3.38.** Is there a code $\mathscr{C} \subseteq \mathbb{F}_2^{18}$ with minimum distance 10 and cardinality $|\mathscr{C}| = 11$? Justify your answer.

Exercise **3.39.** For each of the following 3-tuples (q, n, d) compute what the Singleton, Hamming, Griesmer and Plotkin (when applicable) bounds say about the dimension (equivalently, about the cardinality) of a linear code $\mathscr{C} \leq \mathbb{F}_q^n$ with minimum distance d. Regarding the Griesmer bound, proceed as in Exercise 3.36.

q	n	d
2	13	5
3	13	5
2	13	3
3	13	3
2	19	4
3	19	4
2	17	5
3	17	5
2	7	3
3	7	3

Exercise **3.40.** For each of the following 3-tuples (q, n, d) compute what the Singleton, Hamming and Plotkin (when applicable) bounds say about the dimension (equivalently, about the cardinality) of a linear code $\mathscr{C} \leq \mathbb{F}_q^n$ with minimum distance d.

q	n	d
2	8	6
3	8	6
4	8	6
5	8	6
7	8	6
8	8	6

Exercise **3.41.** Show that for all prime powers q and for all integers $n, t \geq 1$ with $n \leq q$ and $1 \leq t < n/2$ we have

$$\sum_{i=0}^{t} \binom{n}{i} (q-1)^i q^{n-2t} \leq q^n.$$

Exercise **3.42.** 1. Check that the existence of a binary linear code $\mathscr{C} \leq \mathbb{F}_2^{13}$ of dimension 6 and minimum distance 5 is in principle allowed by the Singleton, the Hamming and the Griesmer bounds (the Plotkin bound cannot be applied).

2. Show that such a code does not exist as follows:
   (a) prove first that if such a code $\mathscr{C}$ existed, then a code $\mathscr{D} \leq \mathbb{F}_2^8$ of dimension 5 and minimum distance at least 3 would exist;
   (b) then prove that the code $\mathscr{D}$ cannot exist.

Exercise **3.43.** Show that all simplex codes meet the Griesmer bound.

Chapter 4

Reed-Solomon and Goppa Codes

In this chapter we study Reed-Solomon and Goppa codes. Reed-Solomon codes are one of the most important families of error-correcting objects. They are constructed by evaluating univariate polynomials over distinct elements of the underlying finite field. Goppa codes are currently good candidates for constructing cryptosystems that can resist quantum attacks (see Chapter 8).

4.1 Reed-Solomon Codes

We start by defining Reed-Solomon codes. In fact, the codes defined in this section are sometimes called *generalized Reed-Solomon codes*.

In the sequel, for $k \geq 0$ we let $\mathbb{F}_q[X]_{<k}$ denote the $\mathbb{F}_q$-linear space of univariate polynomials of degree strictly smaller than k. Recall that the zero polynomial has degree $-\infty$ by definition and thus belongs to $\mathbb{F}_q[X]_{<k}$. Note moreover that $\mathbb{F}_q[X]_{<k}$ has dimension k over $\mathbb{F}_q$ with basis $\{1, X, X^2, ..., X^{k-1}\}$.

Definition 4.1. Let $0 \leq k \leq n$ be integers. Suppose $q \geq n$ and let $\mathscr{P} = (\alpha_1, ..., \alpha_n)$ be an n-tuple of distinct elements of $\mathbb{F}_q$. The **Reed-Solomon code** associated with $(q, n, k, \mathscr{P})$ is

$$\mathrm{RS}_q(n, k, \mathscr{P}) := \{(p(\alpha_1), ..., p(\alpha_n)) \mid p \in \mathbb{F}_q[X]_{<k}\} \leq \mathbb{F}_q^n.$$

It is important to note that Reed-Solomon codes only exist over sufficiently large fields. That's the main limitation of these codes.

Example 4.2. Let $\mathbb{F}_4 = \mathbb{F}_2[\alpha]$ be the finite field with 4 elements, with field equation $\alpha^2 + \alpha + 1 = 0$. Define the 4-tuple $\mathscr{P} = (0, 1, \alpha, \alpha^2) = (0, 1, \alpha, \alpha + 1)$. If $n = 4$ and $k = 2$, then $\mathrm{RS}_4(4, 2, \mathscr{P})$ is the set of the evaluations at $\mathscr{P}$ of the polynomials in $\mathbb{F}_4[X]_{<2}$. These are the sixteen polynomials in

$$\{\beta_0 + \beta_1 X \mid \beta_0, \beta_1 \in \mathbb{F}_4\}.$$

For example, evaluating $1 + \alpha X$ we obtain the codeword

$$x = (1, 1 + \alpha, 1 + \alpha^2, 1 + \alpha(\alpha + 1)) = (1, 1 + \alpha, \alpha, 0) \in \mathbb{F}_4^4.$$

Note that x has Hamming weight 3. Moreover, a generator matrix of our $\mathrm{RS}_4(4, 2, \mathscr{P})$ is

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

and a parity-check matrix is

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

The next result computes the dimension and minimum distance of Reed-Solomon codes. In particular, it shows that they are MDS.

Theorem 4.3. Let $(q, n, k, \mathscr{P})$ be as in Definition 4.1. Then $\mathrm{RS}_q(n, k, \mathscr{P})$ has dimension k and minimum distance $n - k + 1$.

Proof. Let $\varphi : \mathbb{F}_q[X]_{<k} \to \mathbb{F}_q^n$ be the evaluation map at the points of $\mathscr{P} = (\alpha_1, ..., \alpha_n)$, i.e., $\varphi(p) = (p(\alpha_1), ..., p(\alpha_n))$ for all p. It is easy to check that φ is $\mathbb{F}_q$-linear and that the code $\mathrm{RS}_q(n, k, \mathscr{P})$ is its image.

If $k = 0$ then $\mathrm{RS}_q(n, k, \mathscr{P})$ is the zero code and its minimum distance is $n + 1$, as desired. We henceforth assume $k \geq 1$. We will show that for every non-zero polynomial $p \in \mathbb{F}_q[X]_{<k}$ we have $\omega^{\mathrm{H}}(\varphi(p)) \geq n - k + 1$. This immediately implies that $\mathrm{RS}_q(n, k, \mathscr{P})$ has dimension k and minimum distance at least $n - k + 1$.

Suppose towards a contradiction that $\omega^{\mathrm{H}}(\varphi(p)) \leq n - k$ for a non-zero polynomial $p \in \mathbb{F}_q[X]_{<k}$. Then p has at least $n-(n-k) = k$ distinct roots and degree strictly smaller than k. Therefore it must be the zero polynomial, a contradiction.

To conclude that $\mathrm{RS}_q(n, k, \mathscr{P})$ has minimum distance *exactly* $n-k+1$ (and is therefore MDS) it suffices to apply the Singleton bound (Theorem 3.1). □

The proof of the previous theorem also shows the following.

Proposition 4.4. Let q, n and $\mathscr{P} = (\alpha_1, ..., \alpha_n)$ be as in Definition 4.1. The evaluation map $\varphi : \mathbb{F}_q[X]_{<n} \to \mathbb{F}_q^n$ defined by $\varphi(p) := (p(\alpha_1), ..., p(\alpha_n))$ for all $p \in \mathbb{F}_q[X]_{<n}$ is an $\mathbb{F}_q$-isomorphism.

4.2 The Berlekamp-Welch Algorithm

Reed-Solomon codes can be efficiently decoded. The algorithm that we now describe is due to Berlekamp and Welch, and attempts to reconstruct the polynomial $p \in \mathbb{F}_q[X]_{<k}$

that gives rise to the transmitted codeword via evaluation. In this section we follow the notation of the previous one.

Algorithm 4.5 (Berlekamp-Welch)**.** The inputs are:

- a prime power q, integers $n \geq k \geq 1$, $d := n - k + 1$;
- an integer $0 < \varepsilon < d/2$ (an estimate for the number of errors);
- the list of distinct evaluation points $\mathscr{P} = (\alpha_1, ..., \alpha_n) \in \mathbb{F}_q^n$;
- the received vector $y \notin \mathrm{RS}_q(n, k, \mathscr{P})$.

Proceed as follows:

1. Use linear algebra (solve a linear system) and look for a non-zero polynomial $E \in \mathbb{F}_q[X]$ of degree exactly ε and a polynomial $N \in \mathbb{F}_q[X]$ of degree at most $\varepsilon + k - 1$ such that
$$y_i E(\alpha_i) = N(\alpha_i) \quad \text{for all } 1 \leq i \leq n.$$
2. If such polynomials do not exist or E does not divide N, then return a failure message. Otherwise let $\tilde{p} := N/E$.
3. If $d^{\mathrm{H}}(y, (\tilde{p}(\alpha_1), ..., \tilde{p}(\alpha_n))) > \varepsilon$ then return a failure message, otherwise return the vector $(\tilde{p}(\alpha_1), ..., \tilde{p}(\alpha_n))$.

Next, we show that Algorithm 4.5 terminates correctly.

Theorem 4.6. Let $(q, n, k, \mathscr{P})$ be as in Definition 4.1. Assume $k \geq 1$ and define $d = n - k + 1$. Let $1 < \varepsilon < d/2$ be an integer, $x \in \mathrm{RS}_q(n, k, \mathscr{P})$ and $y \in \mathbb{F}_q^n$ with $1 \leq d^{\mathrm{H}}(x, y) \leq \varepsilon$. Then Algorithm 4.5 returns x.

Proof. By the definition of $\mathrm{RS}_q(n, k, \mathscr{P})$, there exists a polynomial $p \in \mathbb{F}_q[X]_{<k}$ such that $(p(\alpha_1), ..., p(\alpha_n)) = x$. Such a polynomial is in fact unique by Proposition 4.4.

1. We start by showing that there exist a non-zero polynomial $E \in \mathbb{F}_q[X]$ of degree exactly ε and a polynomial $N \in \mathbb{F}_q[X]$ of degree at most $\varepsilon + k - 1$ such that
$$y_i E(\alpha_i) = N(\alpha_i) \quad \text{for all } 1 \leq i \leq n \tag{4.1}$$
and $N/E = p$. Take
$$E := X^{\varepsilon - d^{\mathrm{H}}(y,x)} \prod_{\substack{j \in \{1,...,n\} \\ y_j \neq p(\alpha_j)}} (X - \alpha_j), \qquad N := p \cdot E.$$
We clearly have $N/E = p$. Now fix $1 \leq i \leq n$ and observe the following:
   - if $E(\alpha_i) = 0$, then $N(\alpha_i) = 0$ as well;
   - if $E(\alpha_i) \neq 0$, then $y_i = p(\alpha_i)$ by the definition of E and therefore $y_i E(\alpha_i) = N(\alpha_i)$ by the definition of N.

Therefore in any case we have that (4.1) holds.

2. Now we show that if (E, N) and (E', N') are pairs of polynomials that satisfy (4.1), then $N/E = N'/E'$. Consider the polynomial $R = NE' - N'E$. Since (4.1) holds, we have $R(\alpha_i) = 0$ for all $1 \leq i \leq n$. On the other hand, R has degree upper bounded by $2\varepsilon + k - 1 < d + k - 1 = n$. Since the α_i's are distinct, we conclude that R is the zero polynomial. Hence $N/E = N'/E'$, as claimed.

3. Combining the previous two steps we conclude that the algorithm must compute $\tilde{p} = p$ and thus returns $(p(\alpha_1), ..., p(\alpha_n)) = x$, as desired. □

Example 4.7. Let $\mathscr{C} = \mathrm{RS}_4(4, 2, \mathscr{P})$ be the code of Example 4.2. We use the Berlekamp-Welch algorithm to decode the vector $(1, 1+\alpha, \alpha, \alpha) \in \mathbb{F}_4^4$ to a codeword of $\mathscr{C}$. Note that $y \notin \mathscr{C}$, as $y \cdot H^\top \neq 0$. Take $\varepsilon = 1 < d^{\mathrm{H}}(\mathscr{C})/2$. We want to find polynomials $E = E_0 + E_1 X$ and $N = N_0 + N_1 X + N_2 X^2$ with $E_1 \neq 0$ and $y_i E(\alpha_i) = N(\alpha_i)$ for $1 \leq i \leq 4$. The latter conditions correspond to a homogeneous linear system with 5 variables and 4 equations:

$$\begin{cases} E_0 & = & N_0, \\ (1+\alpha)(E_0 + E_1) & = & N_0 + N_1 + N_2, \\ \alpha(E_0 + E_1\alpha) & = & N_0 + N_1\alpha + N_2(\alpha+1), \\ \alpha E_0 + E_1 & = & N_0 + N_1(\alpha+1) + N_2\alpha. \end{cases}$$

If we order the variables as $(N_0, N_1, N_2, E_0, E_1)$, the matrix of the system is

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

whose reduced row-echelon form is

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \alpha+1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \alpha \\ 0 & 0 & 0 & 1 & \alpha+1 \end{pmatrix}.$$

Therefore a solution with $E_1 \neq 0$ is $(\alpha+1, 0, \alpha, \alpha+1, 1)$, corresponding to the polynomials $E = \alpha + 1 + X$ and $N = \alpha + 1 + \alpha X^2$. One checks that E divides N and that $\tilde{p} = N/E = 1 + \alpha X$. Finally, the evaluation of $\tilde{p}$ at the points of $\mathscr{P}$ is $(1, 1+\alpha, \alpha, 0) \in \mathscr{C}$, which is at distance 1 from y.

We conclude this section with an analysis of the arithmetic complexity of the Berlekamp-Welch algorithm. We will follow the terminology and apply the results of Appendix B.

Remark 4.8. In Algorithm 4.5, finding N and E corresponds to a linear system in $(\varepsilon + k) + (\varepsilon + 1)$ unknowns and n equations. We therefore need to solve a system of n equations in $2\varepsilon + k + 1 < n + 2$ unknowns. We also impose the E has degree exactly ε,

and this corresponds to introducing a new equation. Therefore the final system has $n+1$ equations and at most $n+1$ variables. Using Gaussian elimination, such a system can be solved in $\mathscr{O}(n^3)$ operations.

The polynomial division N/E can definitely be performed in $\mathscr{O}(n^2)$ operations using the long division algorithm, as the degrees of N and E are both upper bounded by n.

Therefore the Berlekamp-Welch algorithm for decoding a Reed-Solomon code has arithmetic complexity in $\mathscr{O}(n^3)$.

4.3 Goppa Codes

Goppa codes form another important class of error-correcting objects. They admit several descriptions and here we give one based on congruences modulo polynomials. In the sequel, m and n are positive integers with $q^m \geq n$.

Definition 4.9. Let $\mathscr{L} = (\alpha_1, ..., \alpha_n)$ be an n-tuple of distinct elements of $\mathbb{F}_{q^m}$, and let $g \in \mathbb{F}_{q^m}[X]$ be a polynomial of degree at least 1 with $g(\alpha_i) \neq 0$ for all $i \in \{1, ..., n\}$. For each $i \in \{1, ..., n\}$ fix a polynomial $h_i \in \mathbb{F}_{q^m}[X]$ with $(X - \alpha_i)h_i \equiv 1 \mod g$. The **Goppa code** associated with $\mathscr{L}$ and g is

$$\Gamma_{q,m}(n, g, \mathscr{L}) := \left\{ x \in \mathbb{F}_q^n \mid \sum_{i=1}^{n} x_i h_i \equiv 0 \mod g \right\} \leq \mathbb{F}_q^n.$$

Remark 4.10. The Goppa code $\Gamma_{q,m}(n, g, \mathscr{L})$ is well-defined. Indeed, the polynomials h_i, for $1 \leq i \leq n$, in Definition 4.9 exist because g and $X - \alpha_i$ are coprime (explain how this shows the existence of $h_1, ..., h_n$). Moreover, each h_i is unique modulo g. Therefore if $h'_1, ..., h'_n$ are other polynomials with the same properties as $h_1, ..., h_n$, then

$$\left\{ x \in \mathbb{F}_q^n \mid \sum_{i=1}^{n} x_i h_i \equiv 0 \mod g \right\} = \left\{ x \in \mathbb{F}_q^n \mid \sum_{i=1}^{n} x_i h'_i \equiv 0 \mod g \right\}.$$

All of this shows that $\Gamma_{q,m}(n, g, \mathscr{L})$ is well-defined.

Exercise 4.11. Following the notation of Definition 4.9, show that for all $i \in \{1, ..., n\}$ we have

$$h_i = -\frac{g - g(\alpha_i)}{X - \alpha_i} g(\alpha_i)^{-1}.$$

Example 4.12. Take $q = 2$, $m = 2$ and $n = 4$. We construct $\mathbb{F}_4$ as $\mathbb{F}_2[\alpha]$ with $\alpha^2 = \alpha + 1$. Let $\mathscr{L} = (0, 1, \alpha, \alpha + 1)$ and $g = X^2 + X + \alpha$. We have $g(0) = g(1) = g(\alpha + 1) = \alpha$ and $g(\alpha) = \alpha + 1$. Therefore we can use $\mathscr{L}$ and g to construct $\Gamma_{2,2}(4, g, \mathscr{L})$. The inverses of $X - \alpha_i$ modulo g are (please check the computations yourself):

$$h_1 = h_4 = (\alpha + 1)X + (\alpha + 1), \quad h_2 = (\alpha + 1)X, \quad h_3 = (\alpha + 1)X + \alpha.$$

The next result gives lower bounds for the dimension and the minimum distance of Goppa codes. We state the result without proof.

Theorem 4.13. In the notation of Definition 4.9, the Goppa code $\Gamma_{q,m}(n, g, \mathscr{L})$ has dimension at least $n - ms$ over $\mathbb{F}_q$ and minimum distance at least $s + 1$, where $s \geq 1$ is the degree of the polynomial g.

Exercise **4.14.** Construct the finite field $\mathbb{F}_{16}$ as $\mathbb{F}_2[\alpha]$, where $\alpha^4 + \alpha^3 + 1 = 0$. Let $\mathscr{L} = \{\alpha^i \mid 2 \leq i \leq 13\}$ and $g = (x + \alpha)(x + \alpha^{14})$. Write down polynomials $h_1, ..., h_{12}$ defining the Goppa code $\Gamma_{2,4}(12, g, \mathscr{L})$. Apply Theorem 4.13 to obtain lower bounds for the dimension and minimum distance of $\Gamma_{2,4}(12, g, \mathscr{L})$.

4.4 Other Exercises

Exercise **4.15.** List all the codewords of the code $\mathscr{C} = \mathrm{RS}(4, 2, \mathscr{P})$ of Example 4.2. Verify that every non-zero codeword has Hamming weight at least 3.

Exercise **4.16.** Let $\mathscr{C} = \mathrm{RS}(4, 2, \mathscr{P})$ be the code of Example 4.2. We use the Berlekamp-Welch algorithm to decode the vectors

$$(1, 1, \alpha, \alpha + 1), \quad (\alpha, \alpha, \alpha, 1), \quad (1, \alpha, \alpha, 0).$$

Exercise **4.17.** Let $q = 5$ and $\mathscr{P} = (1, 3, 2, 4, 0)$. Write down a basis of the polynomials defining $\mathscr{C} = \mathrm{RS}(5, 3, \mathscr{P})$ and a generator matrix for $\mathscr{C}$. Choose a vector $x \in \mathscr{C}$ and change a single entry of your choice obtaining a vector y. Check that $y \notin \mathscr{C}$ and use the Berlekamp-Welch algorithm to decode y.

Exercise **4.18.** Repeat Exercise 4.17 with $\mathscr{C} = \mathrm{RS}(5, 2, \mathscr{P})$.

Chapter 5

Duality Theory

In this chapter we study how a code and its dual relate to each other. More precisely, we will see how information on a code $\mathscr{C}$ gives information on $\mathscr{C}^\perp$. In this chapter we only treat linear codes.

5.1 Preliminary Results

We start with a preliminary result that will play a key role for us. In the sequel, we follow the notation of Section 2.6.

Proposition 5.1. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code and let $S \subseteq \{1, ..., n\}$ be a set of cardinality s. We have

$$|\mathscr{C}(S)| = \frac{|\mathscr{C}|}{q^{n-s}} \, |\mathscr{C}^\perp(S^c)|.$$

Proof. We view $\mathbb{F}_q^n$ as a (trivial) code and write $\mathscr{C}(S) = \mathscr{C} \cap \mathbb{F}_q^n(S)$. Taking duals gives

$$\mathscr{C}(S)^\perp = (\mathscr{C} \cap \mathbb{F}_q^n(S))^\perp = \mathscr{C}^\perp + \mathbb{F}_q^n(S)^\perp,$$

where the last equality follows from part 6 of Theorem 2.64. By Exercise 2.66 we have $\mathbb{F}_q^n(S)^\perp = \mathbb{F}_q^n(S^c)$, from which $\mathscr{C}(S)^\perp = \mathscr{C}^\perp + \mathbb{F}_q^n(S^c)$. We now take dimensions in the previous identity using part 3 of Theorem 2.64, obtaining

$$n - \dim(\mathscr{C}(S)) = \dim(\mathscr{C}^\perp) + \dim(\mathbb{F}_q^n(S^c)) - \dim(\mathscr{C}^\perp(S^c)).$$

Since $\dim(\mathscr{C}^\perp) = n - \dim(\mathscr{C})$ and $\dim(\mathbb{F}_q^n(S^c)) = |S^c| = n - |S|$, we conclude

$$\dim(\mathscr{C}(S)) = \dim(\mathscr{C}) - (n - |S|) + \dim(\mathscr{C}^\perp(S^c)),$$

which is the desired identity. □

Another preliminary result that we will need is an inversion formula for functions

defined on the power set of $\{1, ..., n\}$. It is a specialization of the Möebius Inversion Formula for posets. The proof is omitted.

Lemma 5.2. Let $f : 2^{\{1,...,n\}} \to \mathbb{R}$ be any function. Define the function $g : 2^{\{1,...,n\}} \to \mathbb{R}$ by $g(S) := \sum_{T \subseteq S} f(T)$ for all $S \subseteq \{1, ..., n\}$. Then for all $S \subseteq \{1, ..., n\}$ we have

$$f(S) = \sum_{T \subseteq S} (-1)^{|S|-|T|} g(T).$$

As an application of Lemma 5.2, we prove the well-known inclusion-exclusion formula for sets.

Example 5.3. Let $A_1, ..., A_n \subseteq A$ be sets. We have

$$|A_1 \cup \cdots \cup A_n| = \sum_{\emptyset \neq S \subseteq \{1,...,n\}} (-1)^{|S|+1} \left| \bigcap_{i \in S} A_i \right|. \tag{5.1}$$

To see this, put $A := A_1 \cup \cdots \cup A_n$. For $a \in A$ we let $I(a) := \{1 \leq i \leq n \mid a \in A_i\}$. Define a function $f : 2^{\{1,...,n\}} \to \mathbb{R}$ by $f(S) := |\{a \in A \mid I(a)^c = S\}|$ for all $S \subseteq \{1, ..., n\}$. For $S \subseteq \{1, ..., n\}$ we let

$$\gamma(S) := \bigcap_{i \in S^c} A_i,$$

with the convention that $\gamma(\{1, ..., n\}) = A$.

Next, define $g : 2^{\{1,...,n\}} \to \mathbb{R}$ by $g(S) := \sum_{T \subseteq S} f(T)$. It easily follows from the definitions that for $a \in A$ and $S \subseteq \{1, ..., n\}$ one has

$$I(a)^c \subseteq S \text{ if and only if } a \in \gamma(S).$$

This implies that $g(S) = |\gamma(S)|$ for all $S \subseteq \{1, ..., n\}$.

Finally, observe that $f(\{1, ..., n\}) = 0$, as all elements of A belong to A_i for some i. On the other hand, by Lemma 5.2 we have

$$\begin{aligned} f(\{1, ..., n\}) &= \sum_{T \subseteq \{1,...,n\}} (-1)^{n-|T|} g(T) \\ &= g(\{1, ..., n\}) + \sum_{T \subsetneq \{1,...,n\}} \left| \bigcap_{i \in T^c} A_i \right| (-1)^{n-|T|} \\ &= |A| + \sum_{\emptyset \neq T \subseteq \{1,...,n\}} \left| \bigcap_{i \in T} A_i \right| (-1)^{|T|}, \end{aligned}$$

from which the identity in (5.1) follows.

5.2 The MacWilliams Identities

In this section we prove one of the most elegant result of coding theory, namely, the MacWilliams identities.

Recall from Section 2.5 that $W_i(\mathscr{C})$ denotes the number of weight i codewords in $\mathscr{C}$, and that $(W_0(\mathscr{C}), ..., W_n(\mathscr{C}))$ is called the weight distribution of $\mathscr{C}$. The MacWilliams identities establish a linear relation between the weight distribution of a code and the weight distribution of its dual code. They are among the most elegant results in coding theory.

Theorem 5.4 (MacWilliams identities)**.** Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code. For all $0 \leq j \leq n$ we have

$$W_j(\mathscr{C}^\perp) = \frac{1}{|\mathscr{C}|} \sum_{i=0}^{n} W_i(\mathscr{C}) \sum_{t=0}^{j} \binom{n-t}{j-t} \binom{n-i}{t} (-1)^{j-t} q^t.$$

Proof. For all $S \subseteq \{1, ..., n\}$ let $f(S) := |\{x \in \mathscr{C}^\perp \mid \sigma^{\mathrm{H}}(x) = S\}|$ and $g(S) := \sum_{T \subseteq S} f(T)$. By definition, we have $g(S) = |\mathscr{C}^\perp(S)|$ for all S. Using Lemma 5.2 we then see that for all $S \subseteq \{1, ..., n\}$ of cardinality j we have

$$f(S) = \sum_{t=0}^{j} \sum_{\substack{T \subseteq S \\ |T|=t}} |\mathscr{C}^\perp(T)| \, (-1)^{j-t}.$$

Applying Proposition 5.1 we obtain

$$f(S) = \sum_{t=0}^{j} \sum_{\substack{T \subseteq S \\ |T|=t}} \frac{|\mathscr{C}^\perp|}{q^{n-t}} \, |\mathscr{C}(T^c)| \, (-1)^{j-t} = \frac{1}{|\mathscr{C}|} \sum_{t=0}^{j} q^t (-1)^{j-t} \sum_{\substack{T \subseteq S \\ |T|=t}} |\mathscr{C}(T^c)|$$

for all $S \subseteq \{1, ..., n\}$ with $|S| = j$. Summing over all the sets S of size j we get

$$W_j(\mathscr{C}^\perp) = \frac{1}{|\mathscr{C}|} \sum_{t=0}^{j} q^t (-1)^{j-t} \sum_{\substack{S \subseteq \{1,...,n\} \\ |S|=j}} \sum_{\substack{T \subseteq S \\ |T|=t}} |\mathscr{C}(T^c)|. \tag{5.2}$$

Next, observe that for a fixed $0 \leq t \leq n$ we have (explain the single passages)

$$\sum_{\substack{S \subseteq \{1,...,n\} \\ |S|=j}} \sum_{\substack{T \subseteq S \\ |T|=t}} |\mathscr{C}(T^c)| = \sum_{\substack{S \subseteq \{1,...,n\} \\ |S|=j}} \sum_{\substack{T \supseteq S^c \\ |T|=n-t}} |\mathscr{C}(T)| = \binom{n-t}{j-t} \sum_{\substack{T \subseteq \{1,...,n\} \\ |T|=n-t}} |\mathscr{C}(T)|. \tag{5.3}$$

Finally, counting in two ways the elements of the set

$$\{(T, x) \mid T \subseteq \{1, ..., n\}, \ x \in \mathscr{C}, \ |T| = n - t, \ \sigma^{\mathrm{H}}(x) \subseteq T\}$$

one obtains

$$\sum_{\substack{T\subseteq\{1,\dots,n\}\\ |T|=n-t}} |\mathscr{C}(T)| = \sum_{i=0}^{n} W_i(\mathscr{C})\binom{n-i}{t}. \tag{5.4}$$

Combining equations (5.2), (5.3) and (5.4) we finally get

$$W_j(\mathscr{C}^\perp) = \frac{1}{|\mathscr{C}|}\sum_{t=0}^{j} q^t(-1)^{j-t}\binom{n-t}{j-t}\sum_{i=0}^{n} W_i(\mathscr{C})\binom{n-i}{t},$$

which is the desired expression up to re-arranging the terms. □

Example 5.5. Let $q=3$ and let $\mathscr{C}\leq\mathbb{F}_3^3$ be the 1-dimensional code generated by $(0,1,2)$. We have $W_0(\mathscr{C})=1$, $W_1(\mathscr{C})=W_3(\mathscr{C})=0$ and $W_2(\mathscr{C})=2$. We can compute $W_2(\mathscr{C}^\perp)$ using the MacWilliams identities as

$$W_2(\mathscr{C}^\perp) = \frac{1}{3}\left(\sum_{t=0}^{2}\binom{3-t}{2-t}\binom{3}{t}(-1)^{2-t}\,3^t \;+\; 2\sum_{t=0}^{2}\binom{3-t}{2-t}\binom{1}{t}(-1)^{2-t}\,3^t\right).$$

Exercise **5.6.** Use the MacWilliams identities to give a closed formula for the number of vectors $x\in\mathbb{F}_q^4$ of weight j and such that $x_1+x_2=x_3+x_4=0$.

5.3 Computation of Some Weight Distributions

In this section we show an important application of the MacWilliams identities, namely, the computation of the weight distribution of certain linear codes. The idea behind the approach is simple but quite powerful: for some codes $\mathscr{C}$ the weight distribution is difficult to explicitly compute, but the weight distribution of their dual codes $\mathscr{C}^\perp$ is instead easy to write down. Therefore the weight distribution of $\mathscr{C}$ can be computed from that of $\mathscr{C}^\perp$ via the MacWilliams identities.

Corollary 5.7. Let $r\geq 3$ and $n=(q^r-1)/(q-1)$. Let $\mathscr{C}\leq\mathbb{F}_q^n$ be a Hamming code of redundancy r. For all $0\leq i\leq n$ we have

$$W_i(\mathscr{C}) = \sum_{t=0}^{i}\binom{n-t}{i-t}\binom{n}{t}(-1)^{i-t}q^{t-r} + (q^r-1)\sum_{t=0}^{i}\binom{n-t}{i-t}\binom{n-q^{r-1}}{t}(-1)^{i-t}q^{t-r}.$$

Proof. The dual $\mathscr{C}^\perp$ is a simplex code; see Definition 3.29. We computed the weight distribution of $\mathscr{C}^\perp$ in Proposition 3.30 as $W_0(\mathscr{C}^\perp)=1$, $W_{q^{r-1}}(\mathscr{C}^\perp)=|\mathscr{C}^\perp|-1=q^r-1$, and $W_j(\mathscr{C}^\perp)=0$ for $j\notin\{0,q^{r-1}\}$. Therefore the formula follows from the MacWilliams identities (Theorem 5.4). □

If $\mathscr{C}\leq\mathbb{F}_2^n$ is the even weight code of Example 2.22, then $W_i(\mathscr{C})=0$ if i is odd, while $W_i(\mathscr{C})=\binom{n}{i}$ if i is even. The next result uses the MacWilliams identities to give a different formula for such a weight distribution.

Corollary 5.8. The even weight code $\mathscr{C} \leq \mathbb{F}_q^n$ of Example 2.22 was weight distribution given by

$$W_i(\mathscr{C}) = \binom{n}{i}(-1)^i + \sum_{t=1}^{i} \binom{n-t}{i-t}\binom{n}{t}(-1)^{i-t}\, 2^{t-1}, \qquad 0 \leq i \leq n.$$

Proof. It is easy to see (exercise) that the even weight code $\mathscr{C}$ is the dual of the n-times repetition code, i.e., $\mathscr{C}^\perp = \{(0,...,0),(1,...,1)\} \leq \mathbb{F}_2^n$. Therefore by Theorem 5.4 we have

$$W_i(\mathscr{C}) = \left(\sum_{t=0}^{i} \binom{n-t}{i-t}\binom{n}{t}(-1)^{i-t}\, 2^{t-1}\right) + \frac{1}{2}\binom{n}{i}(-1)^i$$

for all $i \in \{0,...,n\}$. This is the desired expression up to re-arranging the terms. □

Exercise **5.9.** Use the MacWilliams identities and the fact that $(\mathbb{F}_q^n)^\perp = \{0\}$ to give a formula for the number of vectors $x \in \mathbb{F}_q^n$ with Hamming weight i.

5.4 Duality and MDS Codes

In this section we show that the family of MDS codes is closed under duality, i.e., that the dual of an MDS code is MDS. We also prove that all MDS codes $\mathscr{C} \leq \mathbb{F}_q^n$ of the same dimension share the same weight distribution, and compute it explicitly.

Theorem 5.10. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be an MDS code of minimum distance d. Then $\mathscr{C}^\perp$ is MDS with minimum distance $n - d + 2$.

Proof. Let k denote the dimension of $\mathscr{C}$. The result is immediate if $k \in \{0, n\}$, so from now on we assume $1 \leq k \leq n-1$. By definition, the minimum distance of $\mathscr{C}$ is $d = n - k + 1$.

Fix any subset $S \subseteq \{1,...,n\}$ with $|S| = k$. We have $|S^c| = n - k = d - 1$, from which $\mathscr{C}(S^c) = \{0\}$. By Proposition 5.1 we have

$$|\mathscr{C}^\perp(S)| = \frac{q^{n-k}}{q^{n-k}}\,|\mathscr{C}(S^c)| = 1.$$

This implies $|\mathscr{C}^\perp(S)| = 1$. Since S was an arbitrary set of cardinality k, it must be that $d^{\mathrm{H}}(\mathscr{C}^\perp) \geq k + 1 = n - (n-k) + 1$. We conclude that $\mathscr{C}^\perp$ attains the Singleton bound of Theorem 3.1 and is therefore MDS. □

We now turn to the weight distribution of MDS codes.

Lemma 5.11. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be an MDS code of dimension k and minimum distance $d = n - k + 1$. Let $S \subseteq \{1,...,n\}$ be a subset of cardinality s. We have $|\mathscr{C}(S)| = 1$ if $0 \leq s \leq d-1$ and $|\mathscr{C}(S)| = q^{s-d+1}$ otherwise.

Proof. The result is immediate if $k = 0$ or if $k \geq 1$ and $s \leq d-1$. We henceforth assume $k \geq 1$ and $s \geq d$. By Theorem 5.4 we have

$$|\mathscr{C}(S)| = \frac{q^{n-d+1}}{q^{n-s}} |\mathscr{C}^{\perp}(S^c)| = q^{s-d+1},$$

where the last equality follows from the fact that $\mathscr{C}^{\perp}$ has minimum distance $n-d+2$ by Theorem 5.10. □

Exercise **5.12.** Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a non-zero MDS code of minimum distance d. Show that for every $S \subseteq \{1, ..., n\}$ with $|S| = d$ there exist exactly $q-1$ codewords $x \in \mathscr{C}$ with $\sigma^{\mathrm{H}}(x) = S$.

Finally, we compute the weight distribution of an MDS code. The proof is left to the reader; see Exercise 5.14.

Theorem 5.13. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be an MDS code of minimum distance d. For all $0 \leq i \leq n$ we have

$$W_i(\mathscr{C}) = \sum_{t=0}^{d-1} \binom{n}{i} \binom{i}{t} (-1)^{i-t} + \sum_{t=d}^{n} \binom{n}{i} \binom{i}{t} (-1)^{i-t} q^{t-d+1}.$$

Exercise **5.14.** Use Lemma 5.11 to prove Theorem 5.13. You can proceed as follows. For all $S \subseteq \{1, ..., n\}$ let $f(S)$ be the number of $x \in \mathscr{C}$ with $\sigma^{\mathrm{H}}(x) = S$. We can write

$$W_i(\mathscr{C}) = \sum_{\substack{S \subseteq \{1,...,n\} \\ |S|=i}} f(S).$$

Then let $g(S) = \sum_{T \subseteq S} f(T)$ for all $S \subseteq \{1, ..., n\}$ and observe that $g(S) = |\mathscr{C}(S)|$ for all S. Finally, combine Lemma 5.11 with Lemma 5.2.

Exercise **5.15.** Check that Theorem 5.13 returns the weight distribution of the zero code and of $\mathbb{F}_q^n$ for $d = n+1$ and $d = 1$, respectively.

As a corollary of Theorem 5.13, we can obtain strong constraints on the parameters of linear MDS codes. More precisely, the following result shows that MDS codes only exist over sufficiently large fields.

Corollary 5.16. Suppose that there exists a k-dimensional MDS code $\mathscr{C} \leq \mathbb{F}_q^n$ of minimum distance d. The following hold:

1. if $k \geq 2$, then $d \leq q$;
2. if $k \leq n-2$, then $k+1 \leq q$.

Proof. 1. Since $k \geq 2$ we have $d+1 = n-k \leq n$ and so $\binom{n}{d+1} > 0$. Therefore using Theorem 5.13 we compute

$$\frac{W_{d+1}^{\mathrm{H}}(\mathscr{C})}{\binom{n}{d+1}} = \sum_{t=0}^{d-1} \binom{d+1}{t} (-1)^{d+1-t} + \binom{d+1}{d} (-1) q + \binom{d+1}{d+1} (-1)^0 q^2$$

$$
\begin{aligned}
&= \sum_{t=0}^{d-1} \binom{d+1}{t} (-1)^{d+1-t} + (-(d+1)q + q^2) \\
&= \sum_{t=0}^{d+1} \binom{d+1}{t} (-1)^{d+1-t} - \sum_{t=d}^{d+1} \binom{d+1}{t} (-1)^{d+1-t} + (-(d+1)q + q^2) \\
&= 0 - \sum_{t=d}^{d+1} \binom{d+1}{t} (-1)^{d+1-t} + (-(d+1)q + q^2) \\
&= d - qd - q + q^2 \;=\; (1-q)(d-q).
\end{aligned}
$$

Since $W^{\mathrm{H}}_{d+1}(\mathscr{C}) \geq 0$, $\binom{n}{d+1} > 0$ and $q \geq 2$, we have $d \leq q$, as desired.

2. If $k \leq n-2$, then $n-k \geq 2$. We can therefore apply the first part to the dual code $\mathscr{C}^\perp$, which is also MDS by Theorem 5.10. We obtain $n-d+2 \leq q$, i.e., $k+1 \leq q$. □

5.5 Other Exercises

Exercise **5.17.** Use Exercise 3.34 to prove the following statement: If $\mathscr{C} \leq \mathbb{F}_2^n$ is an MDS code, then one of the following occurs:

- $\mathscr{C} = \{0\}$,
- $\mathscr{C} = \mathbb{F}_2^n$,
- $\mathscr{C}$ is the n-times repetition code,
- $\mathscr{C}$ is the even weight code (for this part you can use duality).

Exercise **5.18.** Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $k \geq 1$.

1. Show that $\pi_S(\mathscr{C})^\perp = \pi_S(\mathscr{C}^\perp(S))$ for all $S \subseteq \{1, ..., n\}$ (this is partially solved in Appendix C).
2. Show that a set $S \subseteq \{1, ..., n\}$ with $|S| = k$ is a minimal information set for $\mathscr{C}$ if and only if $\dim(\mathscr{C}^\perp(S)) = 0$.
3. Show that a set $S \subseteq \{1, ..., n\}$ with $|S| = k$ is a minimal information set for $\mathscr{C}$ if and only if S^c is a minimal information set for $\mathscr{C}^\perp$.

Chapter 6

Reed-Muller Codes

This chapter is devoted to one of the best known families of error-correcting codes, namely *Reed-Muller codes*. These are defined via multivariate polynomials and although their general theory can be developed over an arbitrary field $\mathbb{F}_q$, in these notes we restrict ourselves to the binary case. Therefore $q = 2$ in this chapter.

Reed-Muller codes have been extensively used, for example, for deep space exploration. In 1969, the *Mariner 6* spacecraft transmitted the first pictures of the surface of Mars using the Reed-Muller code $\mathrm{RM}(1, 5)$; see Definition 6.2.

6.1 Definition and First Properties

We start by establishing the notation for this chapter.

Notation 6.1. In the sequel, for integers $r \geq 0$ and $m \geq 1$ we denote by $\mathbb{F}_2[X_1, ..., X_m]^{\times}_{\leq r}$ the vector space of square-free polynomials in the variables $X_1, ..., X_m$ of total degree at most r, where the zero polynomial has degree $-\infty$. We also let $\mathscr{P}(m)$ be the list of vectors in $\mathbb{F}_2^m$ sorted in lexicographic order, with $0 < 1$. For example, for $m = 3$ we have

$$\mathscr{P}(3) = ((0,0,0),(0,0,1),(0,1,0),(0,1,1),(1,0,0),(1,0,1),(1,1,0),(1,1,1)). \tag{6.1}$$

Note moreover that

$$\mathbb{F}_2[X_1, ..., X_m]^{\times}_{\leq m} = \mathbb{F}_2[X_1, ..., X_m]^{\times}_{\leq m+1} = \mathbb{F}_2[X_1, ..., X_m]^{\times}_{\leq m+2} = \cdots,$$

as there is no square-free polynomial in m variables of degree strictly larger than m.

Reed-Muller codes are defined as follows.

Definition 6.2. Let $r \geq 0$ and $m \geq 1$ be integers. Let $n = 2^m$ and $\mathscr{P}(m) = (a_1, ..., a_n)$.

The **Reed-Muller code** of parameters (r, m) is

$$\mathrm{RM}(r,m) := \{(p(a_1), ..., p(a_n)) \mid p \in \mathbb{F}_2[X_1, ..., X_m]^{\times}_{\leq r}\} \leq \mathbb{F}_2^n.$$

Example 6.3. Take $r = 1$ and $m = 3$. Then $n = 8$ and $\mathrm{RM}(1,3)$ is generated by the evaluations of the polynomials in $\{1, X_1, X_2, X_3\}$ at the points listed in (6.1). Therefore a generator matrix of $\mathrm{RM}(1,3)$ is

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

The first properties of Reed-Muller codes are summarized in the following result.

Theorem 6.4. Let $r \geq 0$ and $m \geq 1$ be integers, and let $n = 2^m$. Then $\mathrm{RM}(r,m)$ is a linear code of dimension $\sum_{\ell=0}^{r} \binom{m}{\ell}$.

Proof. Let $V := \mathbb{F}_2[X_1, ..., X_m]^{\times}$ be the space of square-free polynomials in the variables $X_1, ..., X_m$ and let $\varphi : V \to \mathbb{F}_2^n$ be the evaluation map on $\mathscr{P}(m)$, i.e., $\varphi(p) = (p(a_1), ..., p(a_n))$ for all $p \in V$. Clearly, φ is linear. We claim that φ is surjective. To see this, let $1 \leq i \leq n$ be arbitrary and let e_i be the i-th element of the canonical basis of $\mathbb{F}_2^n$. Write $\mathscr{P}(m) = (a_1, ..., a_n)$ and $a_i = (a_{i1}, ..., a_{im})$. Define

$$p := (X_1 - a_{i1} + 1) \cdots (X_m - a_{im} + 1) \in V$$

and observe that $p(a_j) = 0$ unless $j = i$, in which case $p(a_i) = 1$ (here we heavily use the fact that we are working over $\mathbb{F}_2$). This shows that e_i is in the image of φ for all $i \in \{1, ..., n\}$. Therefore φ is surjective, as claimed.

Finally, we prove that φ is an isomorphism of vector spaces. As we already know that φ is linear and surjective, it suffices to show that V and $\mathbb{F}_2^n$ have the same dimension over $\mathbb{F}_2$. To prove this, observe that every element $p \in V$ can be uniquely written (why?) as

$$p = \sum_{\ell=0}^{m} \sum_{1 \leq i_1 < i_2 < \cdots < i_\ell \leq m} p_{i_1 i_2 \cdots i_\ell} \, X_1^{i_1} X_2^{i_2} \cdots X_\ell^{i_\ell}, \tag{6.2}$$

where $p_\emptyset$ is the constant term of p. From this description of the elements of V one sees that V has dimension $\sum_{\ell=0}^{m} \binom{m}{\ell} = 2^m = n$ over $\mathbb{F}_2$, as desired.

All of this implies that $\mathrm{RM}(r,m)$ is the image under the isomorphism φ of the linear space $\mathbb{F}_2[X_1, ..., X_m]^{\times}_{\leq r} \leq V$. In particular, $\mathrm{RM}(r,m)$ is linear of dimension

$$\dim(\mathrm{RM}(r,m)) = \dim\left(\mathbb{F}_2[X_1, ..., X_m]^{\times}_{\leq r}\right) = \begin{cases} n & \text{if } r \geq m, \\ \sum_{\ell=0}^{r} \binom{m}{\ell} & \text{if } r \leq m, \end{cases}$$

where the latter formula again follows from (6.2). Since $\binom{m}{\ell} = 0$ for $\ell > m$ we have that

$\mathrm{RM}(r,m)$ has dimension $\sum_{\ell=0}^{r}\binom{m}{\ell}$ for all $r \geq 0$ and $m \geq 1$. □

We conclude with two observations left as exercise.

Exercise 6.5.
- Show that, for all $m \geq 1$, $\mathrm{RM}(0,m)$ is the binary repetition code of length $n = 2^m$ defined in Example 2.7.
- Show that, for all $m \geq 1$ and $r \geq m$, $\mathrm{RM}(r,m) = \mathbb{F}_2^{2^m}$.

6.2 Structure of Reed-Muller Codes

In this section we show a structural property of Reed-Muller codes, connecting them with the Plotkin sum; see Definition 2.58. As a corollary we will obtain the minimum distance of Reed-Muller codes.

Theorem 6.6. Let $1 \leq r < m$ be integers. We have

$$\mathrm{RM}(r,m) = \mathrm{RM}(r,m-1) \oplus_{\mathrm{P}} \mathrm{RM}(r-1,m-1).$$

Proof. We split the terms of any $p \in \mathbb{F}_2[X_1,...,X_m]^{\times}_{\leq r}$ into those containing X_1 and those not containing X_1. We can then write $p = p_0 + X_1 p_1$ with $p_0 \in \mathbb{F}_2[X_2,...,X_m]^{\times}_{\leq r}$ and $p_1 \in \mathbb{F}_2[X_2,...,X_m]^{\times}_{\leq r-1}$. Note that this defines a bijection

$$\mathbb{F}_2[X_1,...,X_m]^{\times}_{\leq r} \to \mathbb{F}_2[X_2,...,X_m]^{\times}_{\leq r} \times \mathbb{F}_2[X_2,...,X_m]^{\times}_{\leq r-1}.$$

If $\mathscr{P}(m) = (a_1,...,a_n)$ and $\mathscr{P}(m-1) = (b_1,...,b_{n/2})$, then

$$\begin{aligned} p(a_1,...,a_n) &= (p_0(a_1),...,p_0(a_n)) + (a_{11}p_1(a_1),...,a_{n1}p_1(a_n)) \\ &= (p_0(b_1),...,p_0(b_{n/2}),p_0(b_1),...,p_0(b_{n/2})) + (0,...,0,p_1(b_1),...,p_1(b_{n/2})) \end{aligned}$$

(write down a small example and understand why this is the case). So far we have shown the inclusion $\mathrm{RM}(r,m) \subseteq \mathrm{RM}(r,m-1) \oplus_{\mathrm{P}} \mathrm{RM}(r-1,m-1)$. The equality follows from Theorem 6.4 and a dimension argument. □

We continue by discussing some of the consequences of Theorem 6.6. First, the theorem gives us an inductive construction of Reed-Muller codes as follows.

Corollary 6.7. For $r \geq 0$ and $m \geq 1$, the Reed-Muller code $\mathrm{RM}(r,m)$ is given by the recursive construction

$$\mathrm{RM}(r,m) = \begin{cases} \mathbb{F}_2^{2^m} & \text{if } r \geq m \geq 1, \\ \langle (1,...,1) \rangle \leq \mathbb{F}_2^{2^m} & \text{if } r = 0 \text{ and } m \geq 1, \\ \mathrm{RM}(r,m-1) \oplus_{\mathrm{P}} \mathrm{RM}(r-1,m-1) & \text{if } 1 \leq r < m. \end{cases}$$

Proof. Combine Exercise 6.5 with Theorem 6.6. □

Second, Theorem 6.6 and Proposition 2.61 tell us how to compute the minimum distance of a Reed-Muller code.

Corollary 6.8. Let $r \geq 0$ and $m \geq 1$ be integers with $m \geq r$. The minimum distance of $\mathrm{RM}(r, m)$ is 2^{m-r}.

Proof. We will show by induction on $m \geq 1$ the following statement: "For all $0 \leq r \leq m$ the code $\mathrm{RM}(r, m)$ has minimum distance 2^{m-r}".

If $m = 1$ the result easily follows from Corollary 6.7. Now suppose $m \geq 2$. If $r = 0$ or $r = m$ we can again apply Corollary 6.7. If $1 \leq r < m$, then by Theorem 6.6 we have $\mathrm{RM}(r, m) = \mathrm{RM}(r, m-1) \oplus_{\mathrm{P}} \mathrm{RM}(r-1, m-1)$. By the induction hypothesis, $\mathrm{RM}(r, m-1)$ has minimum distance 2^{m-1-r} and $\mathrm{RM}(r-1, m-1)$ has minimum distance $2^{m-1-r+1} = 2^{m-r}$. Therefore by Proposition 2.61 we conclude that the code $\mathrm{RM}(r, m)$ has minimum distance $\min\{2 \cdot 2^{m-1-r}, 2^{m-r}\} = 2^{m-r}$, as desired. □

6.3 The Dual of a Reed-Muller Code

In this short section we prove the following result showing that the dual of a Reed-Muller code is again a Reed-Muller code.

Theorem 6.9. For all r and m with $1 \leq r \leq m-1$ we have

$$\mathrm{RM}(r, m)^{\perp} = \mathrm{RM}(m-r-1, m).$$

The proof of Theorem 6.9 relies on the following preliminary fact.

Lemma 6.10. For all $m \geq 2$, $\mathrm{RM}(m-1, m)$ is the even weight code of length 2^m; see Example 2.22 for the definition.

Proof. Evaluating a monomial $\mu \in \mathbb{F}_2[X_2, ..., X_m]^{\times}_{\leq m-1}$ at the points of $\mathscr{P}(m)$ one always obtains a vector of even weight (explain why). Since these monomials span $\mathbb{F}_2[X_1, ..., X_m]^{\times}_{\leq r}$, by Example 2.22 all the codewords of $\mathrm{RM}(m-1, m)$ have even weight. On the other hand, by Theorem 6.4 the code $\mathrm{RM}(m-1, m)$ has dimension

$$\sum_{\ell=0}^{m-1} \binom{m}{\ell} = \sum_{\ell=0}^{m} \binom{m}{\ell} - 1 = 2^m - 1 = n - 1.$$

Therefore $\mathrm{RM}(m-1, m)$ must be the even weight code of length $n = 2^m$, again by Example 2.22 and a dimension argument. □

Proof of Theorem 6.9. By Theorem 6.4, the dimensions of the codes $\mathrm{RM}(m-r-1, m)$ and $\mathrm{RM}(r, m)^{\perp}$ are the same (check this doing the computation). Therefore it suffices to

show the inclusion

$$\mathrm{RM}(m-r-1,m) \subseteq \mathrm{RM}(m,r)^{\perp}. \tag{6.3}$$

To see this, fix arbitrary polynomials $p \in \mathbb{F}_2[X_1,...,X_m]^{\times}_{\leq m-r-1}$, $q \in \mathbb{F}_2[X_1,...,X_m]^{\times}_{\leq r}$ and write $\mathscr{P}(m) = (a_1,...,a_n)$. The degree of pq is upper bounded by $m-1$. Since $\mathrm{RM}(m-1,m)$ is the even weight code by Lemma 6.10, the vector $((pq)(a_1),...,(pq)(a_n))$ has even weight. In particular,

$$0 = \sum_{i=1}^{n} (pq)(a_i) = \langle (p(a_1),...,p(a_n)), (q(a_1),...,q(a_n)) \rangle.$$

This establishes the inclusion in (6.3) and concludes the proof. □

6.4 Other Exercises

Exercise **6.11.** Let $m \geq 3$.

1. Write down a general form for the generator matrix of $\mathrm{RM}(1,m)$.
2. Use Exercise 2.56 to show that $\mathrm{RM}(1,m)$ is equivalent to the dual of the extension of a Hamming code.

Exercise **6.12.** Find all the pairs (r,m) for which $\mathrm{RM}(r,m)$ is a self-dual code (recall that a code $\mathscr{C}$ is self-dual if $\mathscr{C}^{\perp} = \mathscr{C}$; see also Exercise 2.92).

Chapter 7

Distributed Storage and Locality

Error correcting codes can also be used to store files across servers (distributed storage), in such a way that if one or more servers fail, the file can be reconstructed entirely using the remaining servers. In this chapter we study the basics of distributed storage and the concept of locality.

7.1 Storage Strategies

A file is modeled as a vector $v \in \mathbb{F}_q^k$, an object that is naturally divided into k components. We assume that every vector in $\mathbb{F}_q^k$ is a valid file.

Observe that if we stored each component of v in a different server, and a server failed, then it would be in general impossible to recover v from the remaining components. We can solve this problem at the price of using more than k servers to store some redundant information. For example, if $v = (v_1, v_2) \in \mathbb{F}_q^2$, then we can construct the vector $x = (v_1, v_2, v_1 + v_2)$ and store each component of x in a different server. Now any of the three components of x can be recovered from the other two, and we can afford that a single server fails.

Another strategy would be to use four servers and store two copies of the file, i.e., store two copies of v_1 and two copies of v_2. This strategy, although very simple, is sometimes used in practice.

Exercise **7.1.** Assume $q \geq 3$, let $v = (v_1, v_2) \in \mathbb{F}_q^2$ and $x = (v_1, v_2, v_1 + v_2, v_1 - v_2)$. Show that any two components of x can be recovered from the other two.

The idea used to construct x from v in the previous exercise comes from coding theory. More precisely, we select the generator matrix G of a k-dimensional code $\mathscr{C} \leq \mathbb{F}_q^n$ and store the components of $x = v \cdot G$ in n different servers. In other words, we use G to create a vector with a certain redundancy from the original file v. This motivates the following definition.

Definition 7.2. Let $1 \leq s \leq n$ be an integer. A linear code $\mathscr{C} \leq \mathbb{F}_q^n$ is called **s-recoverable** if for every $S \subseteq \{1, ..., n\}$ with $|S| \leq s$ there are no codewords $x, y \in \mathscr{C}$ with $x \neq y$ but $x_i = y_i$ for all $i \in \{1, ..., n\} \setminus S$.

MDS codes are particularly useful to design storage schemes. At the time of writing these notes they are being used, for example, by *Facebook* to store information.

Proposition 7.3. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be an MDS code of dimension $1 \leq k < n$. Then $\mathscr{C}$ is $(n-k)$-recoverable.

Proof. Let $S \subseteq \{1, ..., n\}$ be any set of cardinality $n - k$. By Exercise 3.9, the complement $\{1, ..., n\} \setminus S$ is a minimal information set for $\mathscr{C}$. We then conclude by part 6 of Proposition 2.79. □

Therefore, in an MDS code $\mathscr{C}$ every $(n-k)$ components of $x \in \mathscr{C}$ can be recovered by the remaining ones.

Example 7.4. Let $\mathscr{C} \leq \mathbb{F}_5^5$ be the code generated by

$$G = \begin{pmatrix} 1 & 0 & 0 & 3 & 1 \\ 2 & 1 & 0 & 2 & 0 \\ 3 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

We want to find a codeword $x \in \mathscr{C}$ knowing that $x_2 = 2$, $x_3 = 1$, and $x_5 = 4$. Permuting the rows of G we arrive at

$$G' = \begin{pmatrix} 2 & 1 & 0 & 2 & 0 \\ 3 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 3 & 1 \end{pmatrix},$$

which generates the same code as G. Since x is a linear combination of the rows of G, it must be that

$$x = 2 \cdot (2, 1, 0, 2, 0) + 1 \cdot (3, 0, 1, 1, 0) + 4 \cdot (1, 0, 0, 3, 1).$$

Exercise **7.5.** Let $\mathscr{C}$ and G be as in the previous example.

1. Find the unique codeword $x \in \mathscr{C}$ with $x_1 = 4$, $x_3 = 2$, $x_4 = 3$.
2. Show that $\mathscr{C}$ is not 3-recoverable.

7.2 Locality

Suppose that $\mathscr{C} \leq \mathbb{F}_q^n$ is a 1-recoverable code of dimension $k \geq 1$. Let $x \in \mathscr{C}$, and let $i \in \{1, ..., n\}$ be the index of a lost entry. By definition, x_i can be recovered from the other $n - 1$ components. However, downloading them from the corresponding servers is costly, and a natural question is whether the lost entry x_i can be recovered from *fewer* known entries. This problem gives rise to the notion of locality.

Definition 7.6. Let $1 \leq r \leq n-1$ be an integer (in particular, $n \geq 2$). A code $\mathscr{C} \leq \mathbb{F}_q^n$ of dimension $k \geq 1$ has **locality** r if for every $i \in \{1, ..., n\}$ there exists a set S_i with the following properties:

1. $i \notin S_i$,
2. $|S_i| \leq r$,
3. if $x, y \in \mathscr{C}$ and $x_j = y_j$ for all $j \in S_i$, then $x_i = y_i$.

The S_i's are called **recovery sets**.

The above definition guarantees that the i-th entry of a codeword $x \in \mathscr{C}$ can be uniquely recovered by reading the entries indexed by S_i. In other words, x_i if a function (not necessarily linear) of $(x_j \mid j \in S_i)$.

Example 7.7. For $n \geq 2$, the n-times repetition code of Example 2.7 has locality 1. As the i-th recovery set we can take any subset $S_i \subseteq \{1, ..., n\} \setminus \{i\}$ of cardinality 1.

Example 7.8. The binary code $\mathscr{C} \leq \mathbb{F}_2^3$ generated by

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

has locality 2 with recovery sets $S_1 = \{2, 3\}$, $S_2 = \{3\}$, $S_3 = \{2\}$. The code $\mathscr{C}$ does not have locality 1 because the first component of $x \in \mathscr{C}$ cannot be reconstructed by any other component (please list all the four codewords and check).

Remark 7.9. The concept of locality only makes sense for codes of minimum distance 2 or more. Indeed, if $n \geq 2$ and $\mathscr{C} \leq \mathbb{F}_q^n$ is a code of minimum distance 1, then we can find $x, y \in \mathscr{C}$ and $i \in \{1, ..., n\}$ with $x_i \neq y_i$ and $x_j = y_j$ for all $j \in \{1, ..., n\} \setminus \{i\}$. This contradicts the definition of locality for any $1 \leq r \leq n-1$. In words, the i-th component of a codeword cannot be reconstructed from any subset of the remaining ones.

We can summarize the previous remark as follows.

Proposition 7.10. Let $1 \leq r \leq n-1$ be an integer and let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $k \geq 1$ having **locality** r. Then $d^{\mathrm{H}}(\mathscr{C}) \geq 2$.

On the other hand, all codes of minimum distance at least 2 have locality r for some integer $r \geq 1$.

Proposition 7.11. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $k \geq 1$ and minimum distance at least 2. Then $\mathscr{C}$ has locality $n-1$.

Proof. Exercise. □

By the previous proposition, the following concept is well-defined.

Definition 7.12. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $k \geq 1$ and minimum distance at least 2. The **minimum locality** of $\mathscr{C}$ is

$$\mathrm{loc}(\mathscr{C}) := \min\{1 \leq r \leq n-1 \mid \mathscr{C} \text{ has locality } r\}.$$

Exercise **7.13.** Compute the minimum locality of the ternary code $\mathscr{C} \leq \mathbb{F}_3^5$ defined by

$$\mathscr{C} = \{(a, b, a+b, a-b, a-2b) \mid a, b \in \mathbb{F}_3\}.$$

We can explicitly compute the minimum locality of MDS codes.

Exercise **7.14.** Let $\mathscr{C} \leq \mathbb{F}_q^n$ be an MDS code of dimension $k \geq 1$ and minimum distance at least 2. Then $\mathrm{loc}(\mathscr{C}) = k$. (*Hint*: use information sets and Exercise 3.9).

7.3 Bounds for Codes with Locality

The fact that a code $\mathscr{C}$ has locality r imposes some constraints on the other parameters (length, dimension and minimum distance). In this section we state two bounds for the parameters of a code having a certain locality.

Theorem 7.15. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code of dimension $k \geq 1$ and having locality $1 \leq r \leq n-1$. We have

$$\frac{k}{n} \leq \frac{r}{r+1}.$$

Note that the function $r \mapsto r/(r+1)$ is strictly increasing in r. This aligns with the intuition that (for a fixed n) the smaller the locality, the smaller the code dimension.

Proof of Theorem 7.15. Let $S_1, ..., S_n$ be recovery sets for $\mathscr{C}$ of cardinality at most r. We use the following algorithm to construct pairs of sets $(A_1, B_1), (A_2, B_2), ...$

1. Start with $A_1 := S_1$ and $B_1 := S_1 \cup \{1\}$.
2. For $i \geq 2$ do the following: If $|B_i| = n$, then terminate the algorithm. Otherwise pick $j \notin B_i$ and let $A_{i+1} := A_i \cup (S_j \setminus B_i)$, $B_{i+1} = B_i \cup S_j \cup \{j\}$.

Note that the algorithm terminates, as in the second step we always have $|B_{i+1}| > |B_i|$. Therefore the condition $|B_i| = n$ is met for some $\ell \geq 1$. Note moreover that $\ell \geq n/(r+1)$, as at every step the size of B_i increases at most by $r+1$. The algorithm gives us pairs of sets

$$(A_1, B_1), (A_2, B_2), ..., (A_\ell, B_\ell)$$

with $|B_\ell| = n$. The following properties can be checked (exercise) by induction on i:

- $A_i \subseteq B_i$ for all $1 \leq i \leq \ell$;
- $|B_{i+1}| - |A_{i+1}| = |B_i| - |A_i| + 1$ for all $1 \leq i \leq \ell - 1$;

- for all $1 \leq i \leq \ell$, if $x, y \in \mathscr{C}$ and $x_j = y_j$ for all $j \in A_i$, then $x_j = y_j$ for all $j \in B_i$.

The above properties imply that $|B_\ell| - |A_\ell| = \ell$ and that the projection $\pi_{A_\ell} : \mathscr{C} \to \mathbb{F}_q^{|A_\ell|}$ onto the coordinates in A_ℓ is injective. Therefore

$$k = \dim(\mathscr{C}) \leq |A_\ell| = |B_\ell| - \ell = n - \ell \leq n - n/(r+1) = rn/(r+1),$$

as desired. □

Example 7.16. We apply the previous bound to an MDS code $\mathscr{C} \leq \mathbb{F}_q^n$ with $k \geq 1$ and $d \geq 2$. By Exercise 7.14 we have $\mathrm{loc}(\mathscr{C}) = k$. Therefore

$$k/n \leq k/(k+1),$$

which is equivalent to $k \leq n-1$. Thus $\mathscr{C}$ meets the bound of Theorem 7.15 with equality if and only if $k = n-1$. For an MDS code, this happens if and only if $d = 2$.

Theorem 7.15 can be refined taking into account also the minimum distance of the code d. The proof of the following result is omitted.

Theorem 7.17. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a linear code of dimension $k \geq 1$, minimum distance d and having locality $1 \leq r \leq n-1$. Then

$$d \leq n - k - \lceil k/r \rceil + 2.$$

Example 7.18. An MDS code as in Example 7.16 meets the previous bound with equality. Indeed, since $d = n - k + 1$ we have $d = n - k - \lceil k/k \rceil + 2$.

Exercise **7.19.** For q sufficiently large, is there a code $\mathscr{C} \leq \mathbb{F}_q^{15}$ of dimension $k = 7$ and minimum distance $d = 9$? Is there a code with the same parameters and locality 3?

7.4 The Tamo-Barg Construction

In this section we illustrate a special case of a code construction discovered by Tamo and Barg. The resulting code meets the bound of Theorem 7.17.

Theorem 7.20. Let $n = q - 1$ and let $k, r \geq 1$ be integers with:

- $r \mid k$,
- $r + 1 \mid n = q - 1$,
- $k/r \leq (n+1)/(r+1)$.

Let α be a primitive element of $\mathbb{F}_q$ and let $\mathscr{P} = (\alpha_1, ..., \alpha_n)$ a list of the n non-zero elements of $\mathbb{F}_q$. Fix a bijection $\xi : \{0, ..., r-1\} \times \{0, ..., k/r - 1\} \to \{1, ..., k\}$ and let

$$\mathscr{C} := \left\{ \left(\sum_{t=0}^{r-1} \sum_{u=0}^{\frac{k}{r}-1} v_{\xi(t,u)} \alpha_1^{t+u(r+1)}, ..., \sum_{t=0}^{r-1} \sum_{u=0}^{\frac{k}{r}-1} v_{\xi(t,u)} \alpha_n^{t+u(r+1)} \right) \mid i \in \{1, ..., n\} \right\} \leq \mathbb{F}_q^n.$$

Then $\mathscr{C}$ is a linear code of dimension k, locality r and minimum distance $n-k-\lceil k/r\rceil+2$. In particular, $\mathscr{C}$ meets the bound of Theorem 7.17 with equality.

Proof. Define $\beta := \alpha^{(q-1)/(r+1)} \in \mathbb{F}_q$ and the polynomial $g := X^{r+1} \in \mathbb{F}_q[X]$. Moreover, for all $0 \leq \ell \leq (q-1)/(r+1) - 1$ let

$$A_\ell := \{\alpha^\ell \beta^m \mid 0 \leq m \leq r\}.$$

Claim A. The A_ℓ's partition $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$ into subsets of size $r+1$. Moreover, the polynomial g is constant on each A_ℓ with value $\alpha^{\ell(r+1)}$.

Proof of the claim. Let H be the multiplicatove subgroup of $\mathbb{F}_q^*$ generated by β. We have $|H| = r+1$, the order of β. The number of cosets of $\mathbb{F}_q^*$ modulo H is $(q-1)/|H| = (q-1)/(r+1)$. Each A_ℓ is a coset by definition. Moreover, $A_\ell \neq A_{\ell'}$ for all $0 \leq \ell, \ell' \leq (q-1)/(r+1)-1$ with $\ell \neq \ell'$. To see this, fix $\ell \neq \ell'$ as above and suppose without loss of generality that $\ell' > \ell$. We will show that $\alpha^\ell \in A_\ell$ but $\alpha^\ell \notin A_{\ell'}$. Towards a contradiction, suppose that there exists $0 \leq m \leq r$ with $\alpha^\ell = \alpha^{\ell'}\beta^m$. Then $1 = \alpha^{\ell'-\ell+m(q-1)/(r+1)}$. Since $\ell' > \ell$, we have

$$0 < \ell' - \ell + m(q-1)/(r+1) \leq (q-1)/(r+1) - 1 + r(q-1)/(r+1) = q-2 < q-1.$$

Since α is a primitive element, this is a contradiction. Since the A_ℓ's are distinct and they are cosets, they are all the cosets (as they are the right number). Moreover, since they are cosets they are pairwise disjoint. Finally, it is easy to see that for all $0 \leq \ell \leq (q-1)/(r+1) - 1$ and all $0 \leq m \leq r$ we have

$$g(\alpha^\ell \beta^m) = \alpha^{\ell(r+1)},$$

which only depends on ℓ. Thus g is constant on each A_ℓ, as claimed. ▲

Next, let $s := k + k/r - 1$ and define linear maps

$$\mathbb{F}_q^k \xrightarrow{\psi} \mathbb{F}_q[X]_{<s} \xrightarrow{\varphi} \mathbb{F}_q^n$$

as follows. For $v \in \mathbb{F}_q^k$, let

$$\psi(v) := \sum_{t=0}^{r-1} \sum_{u=0}^{\frac{k}{r}-1} v_{\xi(t,u)} X^{t+u(r+1)}.$$

For $p \in \mathbb{F}_q[X]_{<s}$, let $\varphi(p) := (p(\alpha_1), ..., p(\alpha_n))$. Note that the map ψ is well-defined because

$$r - 1 + (k/r - 1)(r+1) = k + k/r - 2 < s.$$

Moreover, $\mathscr{C}$ is the image of $\varphi \circ \psi$ by definition.

Claim B. The map ψ in injective.

Proof of the claim. The numbers $t + u(r+1)$ for $0 \leq t \leq r-1$ and $0 \leq u \leq k/r - 1$ are distinct. Therefore the polynomials $X^{t+u(r+1)}$ for $0 \leq t \leq r-1$ and $0 \leq u \leq k/r - 1\}$ are linearly independent. In particular, for $v \in \mathbb{F}_q^k$ we have

$$\sum_{t=0}^{r-1} \sum_{u=0}^{\frac{k}{r}-1} v_{\xi(t,u)} X^{t+u(r+1)} = 0$$

if and only if $v_{\xi(t,u)} = 0$ for all t and u. Since ξ is a bijection, this happens if and only if $v = 0$. Therefore ψ is injective. ▲

Claim C. For all $v \in \mathbb{F}_q^k$ with $v \neq 0$ we have $\omega^{\mathrm{H}}((\varphi \circ \psi)(v)) \geq n - k - k/r + 2 \geq 1$. In particular, $\varphi \circ \psi$ is injective and $\mathscr{C}$ has dimension k and minimum distance at least $n - k - k/r + 2$.

Proof of the claim. Let $v \in \mathbb{F}_q^k$ and suppose that $\omega^{\mathrm{H}}((\varphi \circ \psi)(v)) \leq n - k - k/r + 1$. Then the polynomial $\psi(v)$ has at least $k + k/r - 1 = s$ distinct roots and degree strictly smaller than s. Therefore $\psi(v)$ is the zero polynomial. In turn, this implies $v = 0$ because ψ is injective by Claim B. The fact that $n - k - k/r + 2 \geq 1$ follows from our assumption $k/r \leq (n+1)/(r+1)$.

Finally, observe that $\varphi \circ \psi$ is linear and injective by Claim C, and that $\mathscr{C}$ is its image by definition. Therefore $\mathscr{C}$ is a linear code of dimension k and minimum distance at least $n - k - k/r + 2$. ▲

Claim D. The code $\mathscr{C}$ has locality r.

Proof of the claim. For $0 \leq i \leq n$, let $S_i := \{1 \leq j \leq n \mid \alpha_j \in A_\ell\} \setminus \{i\}$, where ℓ is the unique integer with $\alpha_i \in A_\ell$. Note that the S_i's are well-defined by Claim A. Fix any $x \in \mathscr{C}$ and $i \in \{1, ..., n\}$. Define the polynomial

$$\delta := \sum_{h \in S_i} x_h \prod_{\gamma \in A_i \setminus \{\alpha_i, \alpha_h\}} \frac{X - \gamma}{\alpha_h - \gamma} \in \mathbb{F}_q[X].$$

We claim that $\delta(\alpha_i) = x_i$, showing that x_i can be retrieved from the entries indexed by $j \in S_i$. To see this, let $v \in \mathbb{F}_q^n$ be the unique vector with $x = (\varphi \circ \psi)(v)$. The fact that v is unique follows from the injectivity of $\varphi \circ \psi$ (Claim C). Define a second polynomial

$$\partial := \sum_{t=0}^{r-1} \sum_{u=0}^{\frac{k}{r}-1} v_{\xi(t,u)} \alpha_i^{u(r+1)} X^t \in \mathbb{F}_q[X].$$

For all $j \in S_i$ we have

$$\delta(\alpha_j) = x_j$$

and

$$\partial(\alpha_j) = \sum_{t=0}^{r-1} \sum_{u=0}^{\frac{k}{r}-1} v_{\xi(t,u)} g(\alpha_i)^u \alpha_j^t.$$

Since $j \in S_i$, α_i and α_j belong to the same coset A_ℓ. In particular, $g(\alpha_i) = g(\alpha_j)$ by Claim A and so

$$\partial(\alpha_j) = \sum_{t=0}^{r-1} \sum_{u=0}^{\frac{k}{r}-1} v_{\xi(t,u)} g(\alpha_j)^u \alpha_j^t = \psi(v)(\alpha_j) = (\varphi \circ \psi)(v)_j = x_j.$$

Therefore the polynomials δ and ∂ take the same value on r distinct elements of $\mathbb{F}_q$. Since these polynomials have degree upper bounded by $r-1$, we must have that $\delta = \partial$. Therefore

$$\delta(\alpha_i) = \partial(\alpha_i) = x_i,$$

as claimed. This shows that $\mathscr{C}$ has locality r with the S_i's as recovery sets. ▲

Combining all the claims we conclude that $\mathscr{C}$ has dimension k, locality r and minimum distance at least $n - k - k/r + 2$. Therefore, by Theorem 7.17, $\mathscr{C}$ has minimum distance exactly $n - k - k/r + 2$. □

Example 7.21. To be written.

7.5 Other Exercises

Exercise **7.22.** Suppose that $n \geq 2$ is even and let $\mathscr{C} \leq \mathbb{F}_q^{n/2}$ be an MDS code of dimension $k \geq 1$. Define $\mathscr{D} := \{(x, x) \mid x \in \mathscr{C}\} \leq \mathbb{F}_q^n$. Compute $\mathrm{loc}(\mathscr{D})$ and show that $\mathscr{D}$ achieves the bound of Theorem 7.17.

Chapter 8

Code-Based Cryptography

In a few words, cryptography is about protecting information in such a way that only those for whom information is intended can read it. The protection offered by cryptosystem is different from the protection offered by error-correcting codes: In coding theory we protect information from a noise by adding redundancy, while in cryptography we protect information from an eavesdropper. Nonetheless, error-correcting codes can also be used to construct cryptosystems.

8.1 The McEliece Cryptosystem

McEliece proposed a cryptosystem based on error-correcting codes in 1978. It is one of the oldest known *public key* cryptosystems. The system relies on the fact that decoding a linear code whose specific structure is unknown is computationally hard. The **McEliece scheme** works as follows:

- Alice picks a code $\mathscr{C} \leq \mathbb{F}_q^n$ for which the minimum distance decoder $D_{\mathscr{C}} : \mathbb{F}_q^n \to \mathscr{C} \cup \{\boldsymbol{f}\}$ can be efficiently implemented; see Definition 1.18. She then represents $\mathscr{C}$ via a generator matrix $G_{\mathscr{C}}$ and computes $G'_{\mathscr{C}} = A \cdot G_{\mathscr{C}} \cdot P$, where $A \in \mathbb{F}_q^{k \times k}$ is an invertible matrix and $P \in \mathbb{F}_q^{n \times n}$ is a permutation matrix; see Exercise 2.94. Note that, by Exercise 2.17, the matrix $A \cdot G_{\mathscr{C}}$ generates $\mathscr{C}$, while $G'_{\mathscr{C}}$ doesn't in general (but it generates a code that is equivalent to $\mathscr{C}$). Finally, Alice makes $G'_{\mathscr{C}}$ and $d = d^{\mathrm{H}}(\mathscr{C})$ public. This is the *key generation*.
- In order to send a message to Alice, Bob selects a vector $y \in \mathbb{F}_q^k$, computes $x = y \cdot G'_{\mathscr{C}}$ and adds to it a uniformly random vector $e \in \mathbb{F}_q^n$ of Hamming weight $\omega^{\mathrm{H}}(e) = \lfloor (d-1)/2 \rfloor$. This procedure is called *encryption*. The vector $x + e$ is then sent to Alice.
- In order to recover the message sent by Bob (*decryption*), Alice computes the vector

$$(x + e) \cdot P^{-1} = x \cdot P^{-1} + e \cdot P^{-1} = y \cdot A \cdot G_{\mathscr{C}} + e \cdot P^{-1}.$$

Observe that $y \cdot A \cdot G_{\mathscr{C}} \in \mathscr{C}$ and that $\omega^{\mathrm{H}}(e \cdot P^{-1}) = \omega^{\mathrm{H}}(e) < d/2$, because P^{-1} is a permutation matrix as well; see again Exercise 2.94. Therefore (explain why)

$$D_{\mathscr{C}}((x+e) \cdot P^{-1}) = y \cdot A \cdot G_{\mathscr{C}}.$$

Finally, since $A \cdot G_{\mathscr{C}}$ has full rank it has a right-inverse (that can be pre-computed by Alice). In particular, y can be efficiently recovered from $y \cdot A \cdot G_{\mathscr{C}}$.

The attacker Eve has access to $G'_{\mathscr{C}}$ and to $x + e$. Therefore she would be able to recover y if she could implement a decoder for the code generated by $G'_{\mathscr{C}}$. The security of the cryptosystem relies on the fact that decoding in a code whose structure is unknown is computationally hard. Note that Eve does not know the specific code $\mathscr{C}$ that was selected. Moreover, the matrices A and P have the function of disguising the matrix $G_{\mathscr{C}}$, making even more difficult to reveal the structure of $\mathscr{C}$.

8.2 A Note on Attack Strategies

The strategies to attack the McEliece cryptosystem (or other cryposystems based on error-correcting codes, such as the *Niederreiter scheme*) can be divided into two classes.

1. The first type of strategy attemps to recover the structure of the code $\mathscr{C}$ used by Alice starting from $G'_{\mathscr{C}}$. The attacker assumes that $\mathscr{C}$ belongs to a certain family of codes (for example, Reed-Solomon or Goppa codes) and tries to understand which code was picked from that family. This strategy uses the fact that the properties of a highly structured code cannot be entirely disguised by the matrices A and P.
2. The second type of strategy is based on finding efficient decoding algorithms that work nicely for sufficiently general classes of codes. The best attacks in this class are obtained by improving a general decoding technique called *information set decoding*.

8.3 Information Set Decoding

In this section we describe how to decode an arbitrary linear code using information sets; see Section 2.9. In the sequel, given a matrix $G \in \mathbb{F}_q^{k \times n}$ and a non-empty set $S \subseteq \{1, ..., n\}$ of cardinality s, we let $\pi_S(G) \in \mathbb{F}_q^{k \times s}$ be the matrix formed by the columns of G indexed by S (in the same order).

Lemma 8.1. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $k \geq 1$ and let $G \in \mathbb{F}_q^{k \times n}$ be a generator matrix of $\mathscr{C}$. Let $S \subseteq \{1, ..., n\}$ be a minimal information set for $\mathscr{C}$. Then $\pi_S(G)$ has size $k \times k$ and is invertible. Moreover, for all $x \in \mathscr{C}$ we have

$$x = \pi_S(x) \cdot \pi_S(G)^{-1} \cdot G.$$

Proof. The matrix $\pi_S(\mathscr{C})$ has size $k \times k$ by the definition of minimal information set and is invertible by Proposition 2.79, part 4. Write $x = z \cdot G$ for $z \in \mathbb{F}_q^k$. We trivially have

$$x = z \cdot \pi_S(G) \cdot \pi_S(G)^{-1} G. \tag{8.1}$$

Applying π_S to both sides of (8.1) we obtain

$$\pi_S(x) = z \cdot \pi_S(G) \cdot \pi_S(G)^{-1} \pi_S(G) = z \cdot \pi_S(G) \cdot I_k,$$

where I_k denotes the identity $k \times k$ matrix over $\mathbb{F}_q$. Therefore $z = \pi_S(x) \cdot \pi_S(G)^{-1}$. Substituting z into (8.1) gives the lemma. □

Proposition 8.2. Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a code of dimension $k \geq 1$ and minimum distance d. Let $G \in \mathbb{F}_q^{k \times n}$ be a generator matrix of $\mathscr{C}$. Let $x \in \mathscr{C}$ be a codeword and $e \in \mathbb{F}_q^n$ a vector of weight $\omega^{\mathrm{H}}(e) < d/2$. There exists a minimal information set $S \subseteq \{1, ..., n\}$ of $\mathscr{C}$ for which

$$x = \pi_S(x + e) \cdot \pi_S(G)^{-1} \cdot G.$$

In particular, there exists a minimal information set $S \subseteq \{1, ..., n\}$ of $\mathscr{C}$ with

$$d^{\mathrm{H}}\left(x + e, \pi_S(x + e) \cdot \pi_S(G)^{-1} \cdot G\right) < d/2.$$

Proof. Let $T := \{1, ..., n\} \setminus \sigma^{\mathrm{H}}(e)$. We claim that $\pi_T(G)$ has rank k. To see this, suppose towards a contradiction that $\mathrm{rk}(\pi_T(G)) \leq k - 1$. Then there exists $x \in \mathscr{C}$ with $x \neq 0$ and $\sigma^{\mathrm{H}}(x) \cap T = \emptyset$ (explain why). Therefore $1 \leq \omega^{\mathrm{H}}(x) < d/2$, contradicting the fact that $\mathscr{C}$ has minimum distance d.

Since $\pi_T(G)$ has rank k, there exists a minimal information set S for $\mathscr{C}$ with $S \subseteq T$. In particular, S is disjoint from $\sigma^{\mathrm{H}}(e)$ and so $\pi_S(x + e) = \pi_S(x)$. Therefore

$$\pi_S(x + e) \cdot \pi_S(G)^{-1} \cdot G = \pi_S(x) \cdot \pi_S(G)^{-1} \cdot G = x,$$

where the latter identity follows from Lemma 8.1. □

Proposition 8.2 suggests the following decoding algorithm for a linear code based on minimal information sets.

Algorithm 8.3 (Information set decoding)**.** The inputs are:

- the generator matrix $G \in \mathbb{F}_q^{k \times n}$ of a code $\mathscr{C} \leq \mathbb{F}_q^n$ of dimension $k \geq 1$ and $d^{\mathrm{H}}(\mathscr{C}) = d$;
- the collection $\mathscr{I}(\mathscr{C})$ of all minimal information sets of $\mathscr{C}$;
- the received vector $y \in \mathbb{F}_q^n$.

Proceed as follows:

1. For all $S \in \mathscr{I}(\mathscr{C})$ compute $x_S := \pi_S(y) \cdot \pi_S(G)^{-1} \cdot G \in \mathscr{C}$.
2. If $d^{\mathrm{H}}(y, x_S) < d/2$, then return x_S and terminate the algorithm.

3. If no such x_S is found, then return a failure message.

Remark 8.4. In the notation of Algorithm 8.3, Proposition 8.2 guarantees that if $y = x+e$ for some $x \in \mathscr{C}$ and $e \in \mathbb{F}_q^n$ with $\omega^{\mathrm{H}}(e) < d/2$, then information set decoding is successful and returns x. Note that, under these assumptions, x is the unique codeword of $\mathscr{C}$ that minimizes the Hamming distance from y.

We illustrate how information set decoding works with an example.

Example 8.5. Consider the 2-dimensional code $\mathscr{C} \leq \mathbb{F}_3^5$ generated by the matrix

$$G = \begin{pmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 1 \end{pmatrix} \in \mathbb{F}_3^{2\times 5}.$$

It can be checked that $\mathscr{C}$ has minimum distance $d = 3$. Let $x = (2,0,2,1,2) \in \mathscr{C}$ and $y = x + (0,0,2,0,0) = (2,0,1,1,2)$. Note that $d^{\mathrm{H}}(x,y) = 1 < d/2$.

We list the minimal information sets of $\mathscr{C}$ as:

$$S_1 = \{1,3\}, \quad S_2 = \{1,5\}, \quad S_3 = \{2,5\}, \quad S_4 = \{1,4\}, \quad S_5 = \{4,5\},$$
$$S_6 = \{2,4\}, \quad S_7 = \{1,2\}, \quad S_8 = \{2,3\}, \quad S_9 = \{3,5\}.$$

The algorithm starts with S_1 and computes

$$\begin{aligned} x_{S_1} = \pi_{S_1}(y) \cdot \pi_{S_1}(G)^{-1} \cdot G &= \begin{pmatrix} 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 & 1 & 2 & 0 \end{pmatrix}. \end{aligned}$$

We have $d^{\mathrm{H}}(y, x_{S_1}) = 3 \not< d/2$ and therefore the algorithm passes to the next minimal information set, $S_2 = \{1,5\}$, computing

$$\begin{aligned} x_{S_2} = \pi_{S_2}(y) \cdot \pi_{S_2}(G)^{-1} \cdot G &= \begin{pmatrix} 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 & 2 & 1 & 2 \end{pmatrix}. \end{aligned}$$

This time $d^{\mathrm{H}}(y, x_{S_2}) = 1 < d/2$ and therefore the algoritm returns $(2,0,2,1,2) = x$, terminating correctly.

It is natural to ask if the output of Algorithm 8.3 depends on the order in which the information sets are listed. The answer to this important question is negative.

Proposition 8.6. The output of Algorithm 8.3 does not depend on the order in which the elements of $\mathscr{I}(\mathscr{C})$ are tested.

Proof. Exercise. □

Remark 8.7. In the notation of Algorithm 8.3, suppose that $y = x + e$ where $x \in \mathscr{C}$, $e \in \mathbb{F}_q^n$, and $\omega^{\mathrm{H}}(e) \geq d/2$. Then information set decoding does not return x in general. The following exercise illustrates this point.

Exercise 8.8. Let $\mathscr{C}$, G and x be as in Example 8.5. Show that information set decoding returns a failure message if applied to $y = x + (2, 0, 2, 0, 0)$ and returns $(0, 1, 2, 1, 1) \neq x$ if applied to $y = x + (2, 1, 0, 0, 2)$.

Remark 8.9. In practice, it is convenient to select the information sets to be tested in Algorithm 8.3 uniformly at random. This yields the so-called *probabilistic* information set decoding algorithm.

8.4 Other Exercises

Exercise 8.10 (the answers can be found in Appendix C)**.** Find the minimum distance and all minimal information sets of the code generated by

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{2\times 5}.$$

Use information set decoding to decode the following vectors:

1. $(0, 1, 1, 1, 1) \in \mathscr{C}$;
2. $(1, 1, 1, 1, 1) \notin \mathscr{C}$;
3. $(1, 0, 1, 1, 1) \notin \mathscr{C}$;
4. $(1, 1, 1, 1, 0) \notin \mathscr{C}$;
5. $(0, 0, 0, 0, 0) \in \mathscr{C}$.

Exercise 8.11. Let $1 \leq d \leq n$ be an integer. In the notation of Algorithm 8.3, what is the cardinality of $\mathscr{I}(\mathscr{C})$ if $\mathscr{C}$ is an MDS code of minimum distance d?

Exercise 8.12 (the answers can be found in Appendix C)**.** Find the minimum distance and all minimal information sets of the code generated by

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 4 \end{pmatrix} \in \mathbb{F}_5^{2\times 4}.$$

Use information set decoding to decode the following vectors:

1. $(3, 4, 2, 4) \in \mathscr{C}$;
2. $(3, 4, 2, 1) \notin \mathscr{C}$;
3. $(0, 0, 0, 0) \in \mathscr{C}$;

4. $(0,0,0,2) \notin \mathscr{C}$;

5. $(0,0,1,2) \notin \mathscr{C}$.

Chapter 9

Network Coding

When discussing channels and codes, in Chapter 1 we implicitly concentrated on the scenarios where *one* source of information attempted to communicate with *one* terminal. *Network coding* focuses instead on the situation where one source $\mathbf{S}$ attempts to transmit several messages *simulteneously* to *multiple* terminals $\mathbf{T}_1, ..., \mathbf{T}_N$. The source and the receivers are connected via a *network* of intermediate vertices (or nodes), as in Figure 9.1. Multiple sources can also be allowed, leading to a rich theory. In these notes we only treat single-source networks.

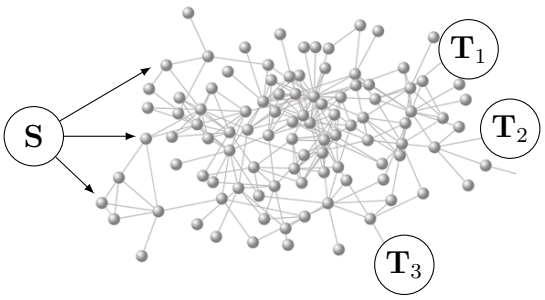

Figure 9.1: An example of network

9.1 Recombining Messages

In the context of network coding, the source $\mathbf{S}$ sends over the outgoing edges vectors that belong to an extension field $\mathbb{F}_{q^m}$, for some q and some m that depend on the edge capacity. We regard $\mathbb{F}_{q^m}$ as a linear space over $\mathbb{F}_q$. Each terminal needs to receive all such vectors (*multicast*). Moreover, the network is by assumption *delay free*, i.e., communication is instantaneous.

We assume that the network is error free, and ask ourselves how many messages can be transmitted with a single use of the network. This number is called the (**one-shot**) **capacity** of the network.

The traditional approach to network multicast is based on *routing* and works as follows. An intermediate vertex collects the inputs from the incoming edges, and forwards as many of these as possible towards the terminals. In the seminal paper [1], Ahlswede, Cai, Li, and Yeung discovered that there is a better approach, namely, *recombining* the inputs with each other before forwarding them. This strategy is called *network coding* and is efficiently illustrated with an example.

The network in Figure 9.2 has one source $\mathbf{S}$, two terminals $\mathbf{T}_1$ and $\mathbf{T}_2$, and four intermediate vertices connected as shown in the picture. It is called the *Butterfly network*. The source $\mathbf{S}$ attempts to transmit messages to both terminals $\mathbf{T}_1$ and $\mathbf{T}_2$ as efficiently as possible.

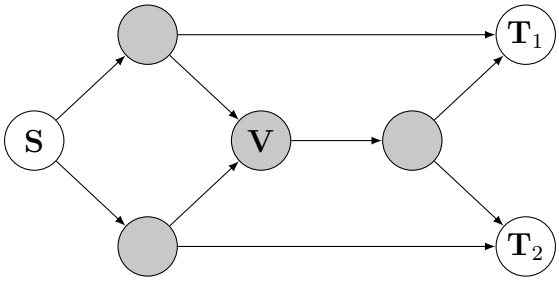

Figure 9.2: The Butterfly network

It is not difficult to see that, using a classical routing strategy, the source $\mathbf{S}$ cannot transmit more than 1 message to both terminals in a single channel use. Intuitively, the reason behind this lies in the fact that the vertex $\mathbf{V}$ acts as a "bottleneck", as it has two incoming edges but only one outgoing edge. Therefore, when performing routing, only one of the input messages can be forwarded by $\mathbf{V}$. Please try yourself to transmit e.g. two messages in a single channel use using routing, and make sure you don't manage! ;-)

If one uses the network multiple times (which is not the model we are focusing on in these notes, but still very interesting to think about) it can be shown that routing cannot deliver more than 1.5 messages per channel use in average. A strategy that delivers 3 messages in 2 time slots in depicted in Figures 9.3 and 9.4.

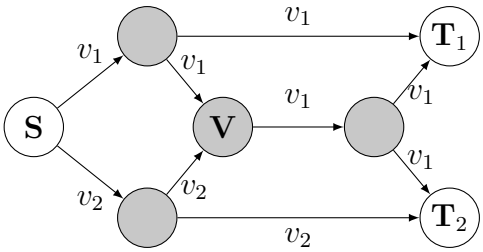

Figure 9.3: Routing, time slot 1.

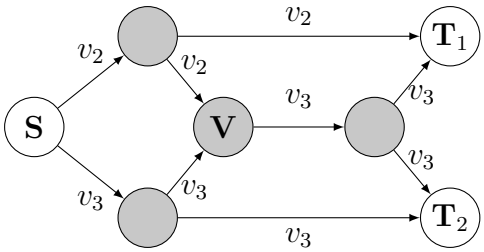

Figure 9.4: Routing, time slot 2.

In Figure 9.5 a network coding strategy is presented that delivers instead two messages, achieving a rate of 2 transmitted messages per channel use. This time the vertex $\mathbf{V}$ is allowed to transmit the sum of the two incoming messages v_1 and v_2, instead of routing only one of the two. Terminal $\mathbf{T}_1$ obtains v_1 and $v_1 + v_2$, and terminal $\mathbf{T}_2$ obtains v_2 and

$v_1 + v_2$. So both terminals can easily compute v_1 and v_2. It is possible to show that 2 is the maximum number of messages that can be transmitted in average with *any* strategy; see also the following Theorem 9.7.

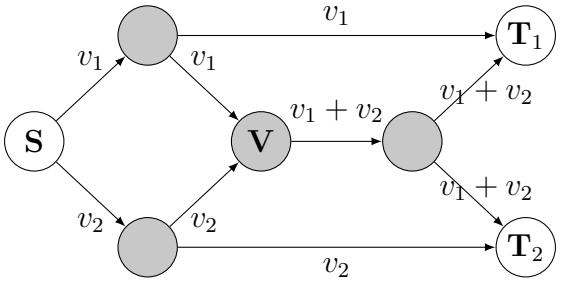

Figure 9.5: A network coding scheme.

A fundamental result in coding theory states that network coding can be applied to any network, resulting in a capacity-achieving communication strategy under certain assumptions.

DISCLAIMER For the sake of clarity in the exposition, we will need to treat some network coding concepts a bit informally. For example, we do not give a rigorous mathematical definition of one-shot capacity of a network, and simply define it as "the maximum number of messages that can be transmitted to all terminals in a single channel use". This is *not* a rigorous mathematical definition, but it is good enough to convey the right idea. In general, when discussing network coding one always needs to find a balance between formalism and clarity.

To reassure the purists, all what we treat can be made fully rigorous, at the price of a quite heavy notation and terminology. The interested reader is referred to [7] for more details about this. We instead refer to [3] for a general introduction to network coding.

9.2 Multicast Networks and the Edge-Cut Bound

In this section we define communication networks and state an upper bound for their capacity. We start by establishing the notation for the remainder of the chapter.

Notation 9.1. In the sequel, q is a prime power and m is a positive integer, unless otherwise stated.

A communication network can be mathematically modeled as follows (as already mentioned, we only treat networks with one source of information).

Definition 9.2. A **network** is a 4-tuple $\mathscr{N} = (\mathscr{V}, \mathscr{E}, \mathbf{S}, \mathscr{T})$ where:

1. $(\mathscr{V}, \mathscr{E})$ is a finite, directed, acyclic multigraph;
2. $\mathbf{S} \in \mathscr{V}$ is a vertex called **source**;

3. $\mathscr{T} \subseteq \mathscr{V}$ is a set of vertices called **terminals**.

We also assume that the following hold:

4. $|\mathscr{T}| \geq 1$ and $\mathbf{S} \notin \mathscr{T}$;
5. For every $\mathbf{T} \in \mathscr{T}$ there exists a directed path from $\mathbf{S}$ to $\mathbf{T}$.
6. The source does not have incoming edges, and terminals do not have outgoing edges.
7. For every vertex $\mathbf{V} \in \mathscr{V}$ there exists a direct path from $\mathbf{S}$ to $\mathbf{V}$ and from $\mathbf{V}$ to some terminal $\mathbf{T} \in \mathscr{T}$.

Note that property 7 guarantees that in $\mathscr{N}$ there are no "isolated" vertices.

Notation 9.3. We denote the one shot capacity (see page 83) of $\mathscr{N}$ by $C(\mathscr{N})$.

The main result of this section states that the capacity cannot exceed a certain graph-theoretic invariant of the underlying network $\mathscr{N}$. The latter is defined as follows.

Definition 9.4. Let $\mathscr{N} = (\mathscr{V}, \mathscr{E}, \mathbf{S}, \mathscr{T})$ be a network. An (**edge**) **cut** between $\mathbf{S}$ and a terminal $\mathbf{T} \in \mathscr{T}$ is a subset $\mathscr{E}' \subseteq \mathscr{E}$ with the property that every directed path from $\mathbf{S}$ to $\mathbf{T}$ has an edge from $\mathscr{E}'$. We denote by min-cut$(\mathbf{S}, \mathbf{T})$ the minimum cardinality of an edge-cut between $\mathbf{S}$ and $\mathbf{T}$. Finally, we let

$$\mu(\mathscr{N}) := \min\{\text{min-cut}(\mathbf{S}, \mathbf{T}) \mid \mathbf{T} \in \mathscr{T}\}.$$

Example 9.5. Consider the Butterfly network $\mathscr{N}$ in Figure 9.2. It is easy to see that $\mu(\mathscr{N}) = 2$.

Example 9.6. Consider the network $\mathscr{N}$ in Figure 9.6, with edges labeled as in the picture. It is obtained from the Butterfly network by changing the direction of one edge.

We have that $\{e_2, e_5\}$ is an edge cut between $\mathbf{S}$ and $\mathbf{T}_1$ (while the same edges in the Butterfly network are not an edge cut between $\mathbf{S}$ and $\mathbf{T}_1$). Moreover, there is no edge cut between $\mathbf{S}$ and $\mathbf{T}_1$ with cardinality 1. Thus min-cut$(\mathbf{S}, \mathbf{T}_1) = 2$. On the other hand, $\{e_2\}$ is an edge cut between $\mathbf{S}$ and $\mathbf{T}_2$, hence min-cut$(\mathbf{S}, \mathbf{T}_2) = 1$ (in the Butterfly network we instead have min-cut$(\mathbf{S}, \mathbf{T}_2) = 2$). We therefore conclude that $\mu(\mathscr{N}) = \min\{2, 1\} = 1$.

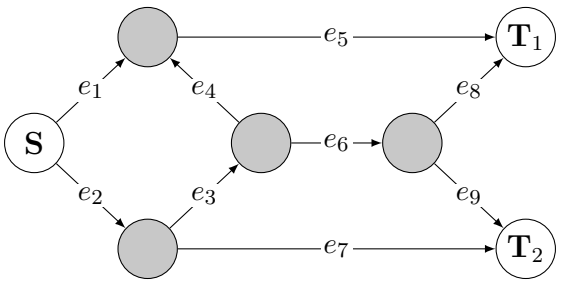

Figure 9.6

The following result gives an upper bound on the number of messages that can be transmitted in terms of $\mu(\mathscr{N})$. Note that the bound does not depend on q and m. A rigorous proof goes beyond the scope of this course and is therefore omitted.

Theorem 9.7. Let $\mathscr{N}$ be a network. We have $C(\mathscr{N}) \leq \mu(\mathscr{N})$.

For example, over the Butterfly network we cannot transmit more than 2 messages in a single channel use (try yourself to do better). In particular, the scheme outlined in Section 9.1 is capacity-achieving.

A natural question is whether the bound of Theorem 9.7 is sharp or not. In this chapter we prove that this is the case under certain assumptions. In fact, we prove that if the messages have entries from a sufficiently large field, then capacity can be achieved by performing only *linear* operations at the intermediate vertices of the underlying network. This is called the *max-flow-min-cut theorem* (not to be confused with the homonymous result from graph theory). Before proving the theorem we need to establish the notation and introduce some network information theory concepts.

9.3 Communication Schemes

In the sequel, $\mathscr{N} = (\mathscr{V}, \mathscr{E}, \mathbf{S}, \mathscr{T})$ denotes a fixed network. We start by observing that the edges of $\mathscr{N}$ form a partially ordered set. More precisely, for $e, e' \in \mathscr{E}$ we write $e \preceq e'$ if there exists a directed path in $\mathscr{N}$ whose initial edge is e and its final edge is e'. It is easy to see (exercise) that $\preceq$ is an order relation on $\mathscr{E}$, which is not total in general.

Example 9.8. In the network of Figure 9.6 we have $e_2 \preceq e_8$, $e_3 \preceq e_5$ and $e_2 \preceq e_5$. On the other hand, e_7 and e_3 are not comparable.

Next, we use a powerful result from order theory, stating that any order can be extended to a total (also called *linear*) order.

Theorem 9.9. Let $(P, \preceq)$ be a finite partially ordered set. There exists a total order $\leq$ on P with the property that $a \preceq b$ implies $a \leq b$.

When working with a network $\mathscr{N}$, its edges are generally labeled with numbers (e_1, e_2, etc). Moreover, the numbers represent an extension of the partial order defined on the edges. This convention has been followed in all networks discussed so far. A "bad" example of labeling is instead the one in Figure 9.7, where $e_4 \preceq e_3$.

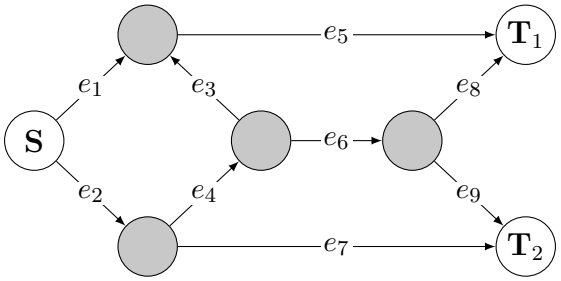

Figure 9.7

Remark 9.10. The choice of a total order $\leq$ for the edges of a network $\mathscr{N}$ allows us to interpret vertex operations without ambiguity. For example, in the network of Figure 9.5, where the partial order was extended according to the indices of the edges, any operation performed by $\mathbf{V}$ is uniquely specified by a function $f : \mathbb{F}_{q^m} \times \mathbb{F}_{q^m} \to \mathbb{F}_{q^m}$. In general, the incoming vectors are first sorted according to the total order chosen for the edges and processed via f. They are then transmitted over the outgoing edges, again according to the chosen total order.

In the sequel, we fix an extension of the partial order on the edges of $\mathscr{N}$. All the results of this chapter do not depend on the specific choice of the order extension.

Definition 9.11. For a vertex $\mathbf{V} \in \mathscr{V}$, denote by $\mathrm{in}(\mathbf{V})$ the set of incoming edges of $\mathbf{V}$ and by $\mathrm{out}(\mathbf{V})$ the set of its outgoing edges. Then a **network code** for $\mathscr{N}$ is a collection of maps

$$F = \left\{ f_{\mathbf{V}} : \mathbb{F}_{q^m}^{|\mathrm{in}(\mathbf{V})|} \to \mathbb{F}_{q^m}^{|\mathrm{out}(\mathbf{V})|} \ : \ \mathbf{V} \in \mathscr{V} \setminus (\{\mathbf{S}\} \cup \mathscr{T}) \right\}.$$

Thanks to Remark 9.10, these uniquely define the operations to be performed by the network vertices. The network code F is called **linear** if each $f_{\mathbf{V}}$ is a map of the form

$$f_{\mathbf{V}} : \begin{pmatrix} v_1 \\ \vdots \\ v_\ell \end{pmatrix} \mapsto G_{\mathbf{V}} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_\ell \end{pmatrix}$$

for some matrix G with entries in $\mathbb{F}_q$. Equivalently, a linear network code is defined as a collection of matrices

$$F = \left\{ f_{\mathbf{V}} \in \mathbb{F}_q^{|\mathrm{out}(\mathbf{V})| \times |\mathrm{in}(\mathbf{V})|} \ : \ \mathbf{V} \in \mathscr{V} \setminus (\{\mathbf{S}\} \cup \mathscr{T}) \right\}.$$

Finally, in order to fully specify a communication strategy on $\mathscr{N}$, we need to describe how it can be initialized by the source.

Definition 9.12. An **outer code** for $\mathscr{N}$ is a non-empty subset $\mathscr{C} \subseteq \mathbb{F}_{q^m}^{|\mathrm{out}(\mathbf{S})|}$.

The set $\mathscr{C}$ is to be interpreted as the collection of messages that the source can attempt to transmit to the terminals; see also the following Example 9.14. Finally, a communication scheme is specified by an outer code and a network code.

Definition 9.13. A **communication scheme** for $\mathscr{N}$ is a pair $(\mathscr{C}, F)$, where $\mathscr{C}$ and F are an outer code and a network code for $\mathscr{N}$, respectively. We say that the communication scheme is **linear** if F is linear. Moreover, $(\mathscr{C}, F)$ is called **unambiguous** if every terminal $\mathbf{T}$ of $\mathscr{N}$ can recover every transmitted vector $x \in \mathscr{C}$ (here we assume that $\mathscr{C}$ and F are known to the terminals).

Example 9.14. We revisit the communication scheme of Figure 9.5. We first extend the partial order on the edges and label the vertices of the network as in Figure 9.8. The

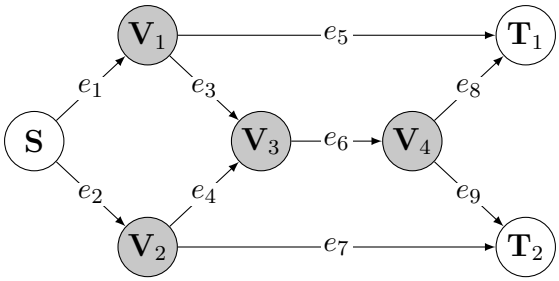

Figure 9.8

outer code is simply $\mathscr{C} = \mathbb{F}_{q^m}^2$. The network code F consists of four functions, one for each intermedite vertex. These functions are linear and are given by the following matrices:

$$f_{\mathbf{V}_1} = f_{\mathbf{V}_2} = f_{\mathbf{V}_4} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad f_{\mathbf{V}_3} = \begin{pmatrix} 1 & 1 \end{pmatrix}.$$

We have already shown at the beginning of the chapter that the communication scheme $(\mathscr{C}, F)$ is unambiguous, as each terminal can recover the transmitted message.

We now give an example of communication scheme that is not unambiguous.

Example 9.15. Take the same setup as in Example 9.14 and remove edge e_5 from the network. Then re-define $f_{\mathbf{V}_1}$ as the identity 1×1 matrix. The pair $(\mathbb{F}_{q^m}^2, F)$ is not anymore an unambiguous communication scheme, as $\mathbf{T}_1$ cannot distinguish between $(0, 0)$ and $(1, -1)$, which are both elements of the outer code $\mathbb{F}_{q^m}^2$ (check all of this yourself).

Remark 9.16. The capacity of $\mathscr{N}$ can now be reinterpreted as the largest real number $C \geq 0$ for which there exists an unambiguous communication scheme $(\mathscr{C}, F)$ for the network $\mathscr{N}$ with $C = \log_{q^m}(|\mathscr{C}|)$.

9.4 The Max-Flow-Min-Cut Theorem

In this section we prove the "main" theorem of network coding. The result states that the bound of Theorem 9.7 can be achieved with linear network coding over a sufficiently large finite field $\mathbb{F}_q$. In the sequel we follow the notation of the previous sections and work with a fixed network $\mathscr{N} = (\mathscr{V}, \mathscr{E}, \mathbf{S}, \mathscr{T})$.

Theorem 9.17 (Max-Flow-Min-Cut)**.** Up to removing some edges from $\mathscr{N}$ and assuming that $q > |\mathscr{T}|$, there exists an unambiguous linear communication scheme $(\mathscr{C}, F)$ for $\mathscr{N}$ with $\mu(\mathscr{N}) = \log_{q^m}(|\mathscr{C}|)$.

The proof of Theorem 9.17 relies on two fundamental classical results from graph theory and algebra, respectively. We state them without proof directly in the context of our interest, starting from the graph thory one.

Theorem 9.18 (Menger)**.** For all $\mathbf{T} \in \mathscr{T}$, the integer min-cut$(\mathbf{S}, \mathbf{T})$ is the maximum number of edge-disjoint directed paths in $(\mathscr{V}, \mathscr{E})$ connecting $\mathbf{S}$ and $\mathbf{T}$.

The next theorem shows that a certain multivariate polynomial always has a non-zero over a sufficiently large field $\mathbb{F}_q$.

Theorem 9.19. Let $p \in \mathbb{F}_q[X_1, ..., X_L]$ be a non-zero polynomial in $L \geq 1$ variables with coefficients in $\mathbb{F}_q$. Suppose that the degree of p in each variable is upper bounded by d. If $q > d$, then there exists $(\alpha_1, ..., \alpha_L) \in \mathbb{F}_q^L$ with $p(\alpha_1, ..., \alpha_L) \neq 0$.

We are now ready to establish the max-flow-min-cut theorem.

Proof of Theorem 9.17. Let $\mu := \mu(\mathscr{N})$ and $t := |\mathscr{T}|$. Write $\mathscr{T} = \{\mathbf{T}_1, ..., \mathbf{T}_t\}$. By Theorem 9.18, for all $i \in \{1, ..., t\}$ there exists μ edge-disjoint directed paths connecting $\mathbf{S}$ and $\mathbf{T}_i$. We delete from $\mathscr{N}$ all edges that do not lie on any of these paths. Note that after these deletions we have $|\text{in}(\mathbf{T}_i)| = \mu$ for all $i \in \{1, ..., t\}$.

We look for a linear communication scheme that is not unambiguous. We will find an outer code of the form

$$\mathscr{C} := \{M(\mathbf{S}) \cdot v^\top \mid v \in \mathbb{F}_{q^m}^{\mu}\} \subseteq \mathbb{F}_{q^m}^{|\text{out}(\mathbf{S})|},$$

where $M(\mathbf{S})$ is an $|\text{out}(\mathbf{S})| \times \mu$ matrix over $\mathbb{F}_q$. Each intermediate vertex $\mathbf{V} \in \mathscr{V} \setminus (\{\mathbf{S}\} \cup \mathscr{T})$ will process information via a linear function $f_{\mathbf{V}}$, represented by a matrix over $\mathbb{F}_q$ of size $|\text{out}(\mathbf{V})| \times |\text{in}(\mathbf{V})|$.

Our goal is to show that there exists a choice for the matrices $\{M(\mathbf{V}) \mid \mathbf{V} \in \mathscr{V} \setminus \mathscr{T}\}$ that produces an unambiguous communication scheme $(\mathscr{C}, F)$. Before continuing with the proof, we briefly illustrate how the approach works over the Butterfly network of Figure 9.2: the matrices we are looking for are visually depicted in Figure 9.9 and their entries are denoted by $X_1, X_2, ..., X_{12}$.

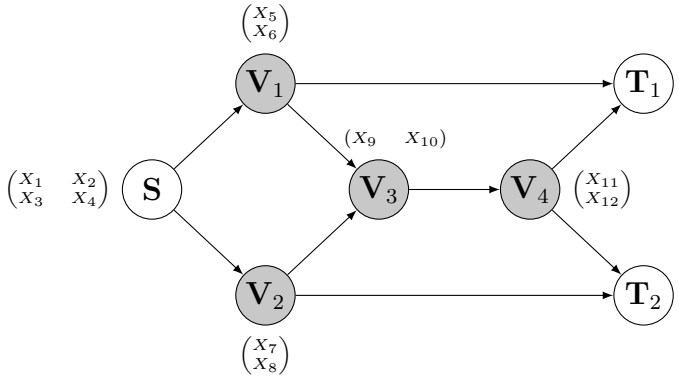

Figure 9.9

We now go back to the general proof. Since every intermediate node processes information according to linear functions, each terminal receives on the incoming edges an

$\mathbb{F}_q$-linear combination of the messages $v_1, ..., v_\mu$ emitted by the source. These linear combinations are uniquely determined by the choice of the matrices in $\{M(\mathbf{V}) \mid \mathbf{V} \in \mathscr{V} \setminus \mathscr{T}\}$. For example, in the network of Figure 9.9 terminal $\mathbf{T}_1$ receives

$$X_5(X_1 v_1 + X_2 v_2)$$

on the top edge if the source emits $(v_1, v_2) \in \mathbb{F}_{q^m}^2$. In other words, whenever $\mathbf{S}$ emits $(v_1, ..., v_\mu) \in \mathbb{F}_{q^m}^\mu$, terminal $\mathbf{T}_i$ receives over the incoming edges the rows of

$$G(\mathbf{T}_i) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_\mu \end{pmatrix},$$

where $G(\mathbf{T}_i)$ is a matrix of size $\mu \times \mu$ and entries in $\mathbb{F}_q$. We will show that the matrices $G(\mathbf{T}_1), ..., G(\mathbf{T}_t)$ can be made *simultaneously* invertible. In turn, this implies that $M(\mathbf{S})$ has rank μ (explain why) and that $(\mathscr{C}, F)$ is unambiguous (explain why), concluding the proof of the theorem.

The entries of $G(\mathbf{T}_1), ..., G(\mathbf{T}_t)$ will be denoted by $(X_\ell \mid 1 \leq \ell \leq L)$ and treated as variables. For the network in Figure 9.9 we had $L = 12$. Note that $G(\mathbf{T}_i)$ is invertible if and only if its determinant is nonzero (we already observed that each of these matrices has size $\mu \times \mu$). Therefore the matrices $G(\mathbf{T}_1), ..., G(\mathbf{T}_t)$ are all invertible if and only if

$$p(X_1, ..., X_L) := \prod_{i=1}^{t} \det(G(\mathbf{T}_i)) \neq 0.$$

We regard p as a polynomial in the L variables $X_1, ..., X_L$. Note that p is not the zero polynomial because, for all $i \in \{1, ..., t\}$, $\det(G(\mathbf{T}_i))$ is a non-zero polynomial. In order to show this, it suffices to observe that one can set the values of the variables to achieve the routing solution from $\mathbf{S}$ to $\mathbf{T}_i$ along the edge-disjoint paths we selected at the very beginning of the proof. These values make $G(\mathbf{T}_i)$ the identity $\mu \times \mu$ matrix, whose determinant is 1. This shows that $\det(G(\mathbf{T}_i))$ is not the zero polynomial.

Another remarkable property of p, whose proof we omit (see [3, Theorem 4]) is that its degree in each variable is upper bounded by t. Therefore, by Theorem 9.19, if $q > t$ then the polynomial p has a non-zero $(\overline{X}_1, ..., \overline{X}_L)$. This concludes the proof.

The communication scheme shown in Figure 9.5 can be obtained with the choice of variables illustrated in Figure 9.10. This choice gives the transfer matrices

$$G(\mathbf{T}_1) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad G(\mathbf{T}_2) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Note that these are both invertible over any field. Therefore the terminals can recover any transmitted messages $(v_1, v_2) \in \mathbb{F}_{q^m}^2$ by inverting them. □

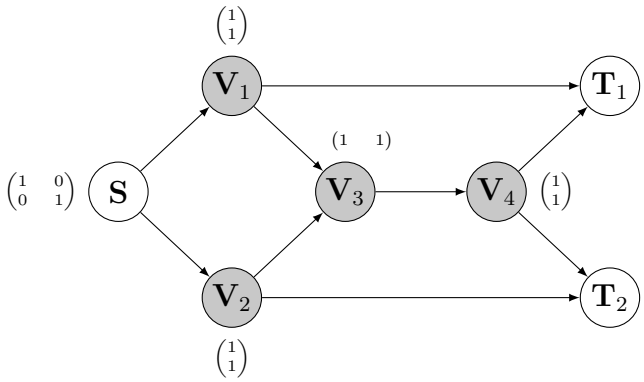

Figure 9.10

Exercise **9.20.** Consider the network $\mathscr{N}$ in Figure 9.11.

1. Identify a linear extension of the partial order defined on the network edges (i.e., label the edges in a way that is compatible with the partial order).
2. Compute $\mu(\mathscr{N})$.
3. Find an unambiguous linear communication scheme $(\mathscr{C}, F)$ that delivers $\mu(\mathscr{N})$ messages to both terminal in a single channel use, for some q.
4. For an edge e of $\mathscr{N}$, let $\mathscr{N}_e$ be the network obtained from $\mathscr{N}$ by removing e (all the rest remains unchanged). Show that $\mu(\mathscr{N}_e) < \mu(\mathscr{N})$ for any edge e.
5. Is there an edge e of $\mathscr{N}$ and an unambiguous (possibly non-linear) communication scheme $(\mathscr{C}, F)$ for $\mathscr{N}_e$ with $\log_{q^m}(|\mathscr{C}|) = 3$? Justify your answer.

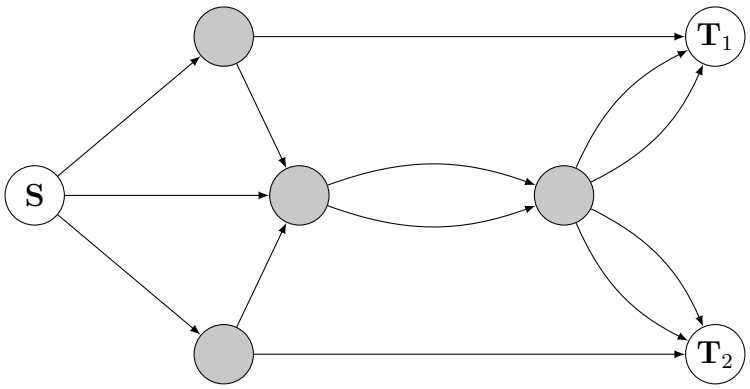

Figure 9.11

Appendix A

Finite Fields

In this appendix we briefly recall the main results of the theory of finite fields. The subject is purposely treated concisely and quite informally. We refer to [4] for a more complete discussion.

Definition A.1. A **finite field** is a field (intended as an algebraic structure) whose cardinality is finite.

The cardinality of a finite field cannot be any number.

Theorem A.2. If $\mathbb{F}$ is a finite field, then there exists a prime number p and an integer $e \geq 1$ with $|\mathbb{F}| = p^e$.

Proof. Define the map $\varphi : \mathbb{Z} \to \mathbb{F}$ by $\varphi(m) = m \cdot 1$. Then φ is a ring homomorphism. Since $\mathbb{F}$ is finite, φ cannot be injective and there exists $m \in \mathbb{Z}$ with $m \neq 0$ and $m \cdot 1 = 0$. We can assume $m \geq 2$ without loss of generality. There is then a ring isomorphism $\mathbb{Z}/m\mathbb{Z} \to \varphi(\mathbb{Z})$. Since $\mathbb{F}$ is a field, $\varphi(\mathbb{Z})$ must be a domain and therefore m has to be prime, say $m = p$. But then $\mathbb{Z}/p\mathbb{Z}$ is a field and thus $\mathbb{F}$ has a subfield $\mathbb{F}'$ isomorphic to $\mathbb{Z}/p\mathbb{Z}$. As $\mathbb{F}$ is a vector space over $\mathbb{F}'$ (whose dimension is finite as $|\mathbb{F}| < +\infty$), we conclude that $|\mathbb{F}| = p^e$ for some $e \geq 1$. □

It turns out that for every prime power p and every $e \geq 1$ there exists a finite field with cardinality (also called **order**) p^e. This result is not trivial and a classical proof relies on a little bit of Galois theory. In words, a finite field with the desired cardinality can be obtained as the splitting field of the polynomial $X^{p^e} - X \in \mathbb{Z}/p\mathbb{Z}[X]$. In these notes we obtain the existence of finite fields as a corollary of the following theorem, which we state without proof.

Theorem A.3. Let p be a prime power and let $e \geq 1$ be an integer. There exists a monic irreducible polynomial $f \in \mathbb{Z}/p\mathbb{Z}[X]$ of degree e.

The previous theorem gives us a concrete way of constructing a finite field with the desired cardinality using the following result from commutative algebra.

Lemma A.4. Let R be a commutative ring with 1 and let $I \subseteq R$ be a maximal ideal. Then the quotient ring R/I is a field.

Constructing a finite field with $q = p^e$ elements is now easy: Take an irreducible monic polynomial $f \in \mathbb{Z}/p\mathbb{Z}[X]$ of degree e and define the quotient $\mathbb{F} := \mathbb{Z}/p\mathbb{Z}[X]/(f)$. Since f is irreducible and $\mathbb{Z}/p\mathbb{Z}[X]$ is a principal ideal ring, the ideal (f) generated by f is maximal. Therefore $\mathbb{F}$ is a field by Lemma A.4. By definition, an element of $\mathbb{F}$ is the equivalence class of a polynomial $g \in \mathbb{Z}/p\mathbb{Z}[X]$ modulo f. These are in bijection with the polynomials of degree up to $e-1$ in $\mathbb{Z}/p\mathbb{Z}[X]$. As the number of such polynomials is p^e, we have $|\mathbb{F}| = p^e$ as well.

If $\pi : \mathbb{Z}/p\mathbb{Z}[x] \to \mathbb{Z}/p\mathbb{Z}[x]/(f)$ denotes the projection on the quotient and $\alpha = \pi(X)$ is the class of X, then

$$\mathbb{F} = \mathbb{Z}/p\mathbb{Z}[\alpha] := \{a_0 + a_1\alpha + \cdots + a_{e-1}\alpha^{e-1} \mid (a_0, ..., a_{e-1}) \in (\mathbb{Z}/p\mathbb{Z})^e\}.$$

In other words, $\mathbb{F}$ can be seen as the set of polynomial expressions in α, where α satisfies the equation $f(\alpha) = 0$.

Example A.5. We want to construct a finite field with $q = 2^3$ elements, so we take $p = 2$ and $e = 3$. The polynomial $f = X^3 + X + 1 \in \mathbb{Z}/2\mathbb{Z}[X]$ is irreducible. Let α be the image of X in the quotient $\mathbb{Z}/2\mathbb{Z}[X]/(X^3 + X + 1)$. Note that α is a root of (the image of) f, i.e., $\alpha^3 = \alpha + 1$. Then the polynomial expressions in α and coefficients in $\mathbb{F}_2$ are a finite field with 2^3 elements. Apart from the 0 element, we will obtain these by taking the powers of α and writing them over the basis $\{1, \alpha, \alpha^2\}$ over $\mathbb{Z}/2\mathbb{Z}$.

Powers of α	Polynomial expression
α^0	1
α^1	α
α^2	α^2
α^3	$\alpha + 1$
α^4	$\alpha^2 + \alpha$
α^5	$\alpha^3 + \alpha^2 = \alpha^2 + \alpha + 1$
α^6	$\alpha^2 + 1$

Note that $\alpha^7 = \alpha(\alpha^2 + 1) = \alpha^3 + \alpha = 1$. Therefore $\mathbb{F} = \{0\} \cup \{\alpha^i \mid 0 \leq i \leq q-2\}$. and all the non-zero elements of $\mathbb{F}$ can be written as a power of α.

If $\mathbb{F}$ is a finite field and $\alpha \in \mathbb{F}$, then the powers of α do not necessarily generate all the non-zero elements of $\mathbb{F}$. However there is always such an α.

Proposition A.6. The multiplicative group of a finite field is cyclic.

An element $\alpha \in \mathbb{F}$ that generate the multiplicative group $\mathbb{F}^*$ is called **primitive**. A monic irreducible polynomial $f \in \mathbb{Z}/p\mathbb{Z}[X]$ is called **primitive** if the class of X (usually

denoted by α) generates the multiplicative group of $\mathbb{F} = \mathbb{Z}/p\mathbb{Z}/(f)$. It turns out that there is always at least one primitive polynomial.

Theorem A.7. For all primes p and all integers $e \geq 1$ there exists a primitive polynomial $f \in \mathbb{Z}/p\mathbb{Z}[X]$ of degree e.

The previous result tells us that a finite field $\mathbb{F}$ with $q = p^e$ elements can always be constructed using the recipe illustrated in Example A.5:

1. identify a primitive polynomial $f \in \mathbb{Z}/p\mathbb{Z}[X]$,
2. define α to be the class of X in the quotient $\mathbb{Z}/p\mathbb{Z}[X]/(f)$,
3. describe $\mathbb{F}$ as the powers of α and the zero element.

A natural question is whether the choice of the primitive polynomial f has a huge impact on the structure of the finite field $\mathbb{F}$. Another one is whether all finite fields can be constructed via primitive polynomials. The answers to this questions are no and yes, respectively. Indeed, a fundamental result from Galois theory tells us that the finite field with $q = p^e$ elements is unique up to isomorphisms.

Theorem A.8. Finite fields $\mathbb{F}$ and $\mathbb{F}'$ are isomorphic if and only if they have the same cardinality.

Note that the writing $q = p^e$ (where p is a prime power and $e \geq 1$) is unique.

Remark A.9. If p_1 and p_2 are prime numbers and $e_1, e_2 \geq 1$ are integers with $p_1^{e_1} = p_2^{e_2}$, then $p_1 = p_2$ and $e_1 = e_2$.

Theorem A.8 ensures that there is essentially a unique finite field with q elements, for every prime power q. We denote it by $\mathbb{F}_q$ and call it *the* finite field with q elements. When $q = p$ is a prime we have $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ up to isomorphism.

Let us see another example.

Example A.10. Let $f = X^2 + 6X + 3 \in \mathbb{F}_7[X]$. We claim that f is irreducible. To see this, observe that f has degree 2. In particular, if f was reducible, it would have a root in $\mathbb{F}_7$. However, we have

$$f(0) =, \quad f(1) = 3, \quad f(2) = 5, \quad f(3) = 2, \quad f(4) = 1, \quad f(5) = 2, \quad f(6) = 5.$$

Thus f is irreducible. Let α be the image of X in the quotient $\mathbb{Z}/7\mathbb{Z}[X]/(X^2 + 6X + 3)$. Then the polynomial expressions in α and coefficients in $\mathbb{F}_7$ are a finite field with 7^2 elements. These can be uniquely written over the basis $\{1, \alpha\}$ using the fact that $\alpha^2 = -6\alpha - 3 = \alpha + 4$.

We conclude the appendix with some exercises.

Exercise **A.11.** Show that the polynomial $f = X^2 + 2X + 2 \in \mathbb{F}_3[X]$ is irreducible. Use f to construct the finite field $\mathbb{F}_9$.

Exercise **A.12.** Is there a finite field with 15 elements? Explain why.

Exercise **A.13.** Find all the irreducible monic polynomials of degree 3 in $\mathbb{F}_2[X]$.

Exercise **A.14.** Let $q = p^e$ be a prime power and let $f = X^q - X \in \mathbb{Z}_p[X]$. Let F be a field containing the splitting field of f. Prove that the roots of f form a subfield $R \subseteq F$. Deduce that R is a splitting field of f.

Exercise **A.15.** Show that the polynomial $f = X^2 + 4X + 2 \in \mathbb{F}_5[X]$ is irreducible. Use f to construct the finite field $\mathbb{F}_{25}$.

Exercise **A.16.** Find all the irreducible monic polynomials in $\mathbb{F}_2[X]$ of degree 3. Construct the finite field $\mathbb{F}_8$ using two of them.

Appendix B

Complexity Essentials

When a code is used in digital communication, decoding is performed by a machine implementing a *decoding algorithm*. As decoding needs to be performed several times in very short time intervals, the corresponding algorithm must be efficient.

In this appendix we briefly describe how the efficiency of an algorithm is measured. In order to do this, we introduce the standard *Landau notation* to describe the growth rate of functions.

Notation B.1. Let $S \subseteq \mathbb{N}$ be an infinite subset and let $f, g : S \to \mathbb{R}$ be functions that only take non-negative values. We write:

$$f \in \mathscr{O}(g) \text{ as } s \to +\infty$$

when there exist numbers $c \in \mathbb{R}$ and $s_0 \in S$ with $c > 0$ and $0 \leq f(s) \leq c \cdot g(s)$ for all real numbers $s \geq a_0$.

Suppose that an algorithm $\mathscr{A}$ has a numerical value, say s, among its inputs (in coding theory, that's typically the code length n). Suppose moreover that $\mathscr{A}$ performs operations over $\mathbb{F}_q$. The (**arithmetic**) **complexity** of $\mathscr{A}$ in base q with respect to the input s is the function that associates to s the number of elementary operations over $\mathbb{F}_q$ needed to terminate the algorithm when s is the input, in the worst-case scenario with respect to all the other inputs. We take as **elementary operations** over $\mathbb{F}_q$ addition and multiplication.

The efficiency of an algorithm is often measured by providing an asymptotic estimate of its complexity with respect to one of the numerical inputs going to infinity, treating the others as constants. We therefore write sentences of the form:

$$\text{“}\mathscr{A} \text{ has complexity in } \mathscr{O}(s^3)\text{”} \quad \text{or} \quad \text{“}\mathscr{A} \text{ terminates in } \mathscr{O}(s^3) \text{ operations”}$$

to say that the complexity of $\mathscr{A}$ with respect to s, say $s \mapsto C(s)$, satisfies $C \in \mathscr{O}(s^3)$. Note that this only provides an asymptotic estimate and does not specify the *exact* number of

needed operations.

In the remainder of the appendix we compute (or state) the complexity of some algorithms. We let n be a positive integer and q be a prime power.

Example B.2. The inner product of $x, y \in \mathbb{F}_q^n$ can be computed by performing n multiplications and $n-1$ additions over $\mathbb{F}_q$ via

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

The complexity in base q of this strategy with respect to n is therefore $2n - 1 \in \mathscr{O}(n)$, and is *linear* in n.

Example B.3. A system of n equations in n unknowns over $\mathbb{F}_q$ can be solved in $\mathscr{O}(n^3)$ operations over $\mathbb{F}_q$ using Gaussian elimination.

Example B.4. The division of polynomials $p, q \in \mathbb{F}_q[X]$ of degree at most n can be performed in $\mathscr{O}(n^2)$ operations over $\mathbb{F}_q$ using the long division algorithm.

Appendix C

Solutions to Some of the Exercises

Proof of Proposition 1.15. We only show the third property. For $x, y \in \mathbb{F}_q^n$ define the set $D(x,y) := \{1 \leq i \leq n \mid x_i \neq y_i\}$. Then for any $x, y, z \in \mathbb{F}_q^n$ we have

$$D(x,y) \subseteq D(x,z) \cup D(z,y).$$

Computing the cardinalities in the above inclusion we obtain

$$d^{\mathrm{H}}(x,y) = |D(x,y)| \leq |D(x,z) \cup D(z,y)| \leq |D(x,z)| + |D(z,y)| = d^{\mathrm{H}}(x,z) + d^{\mathrm{H}}(z,y),$$

as desired. □

Proof of Lemma 1.20. For a vector y, let $S_y := \{1 \leq i \leq n \mid y_i \neq x_i\}$. The number of vectors y at distance i from x is

$$\sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=i}} |\{y \in \mathbb{F}_q^n \mid S_y = S\}| = \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=i}} (q-1)^i = \binom{n}{i}(q-1)^i,$$

which is the desired formula. □

Solution to Exercise 1.29. 1. Take $\mathscr{C} = \{(0,0,0,0), (1,1,1,0)\}$.

2. Suppose by contradiction that there exists a code $\mathscr{C} \subseteq \mathbb{F}_2^4$ with $|\mathscr{C}| \geq 3$ and $d^{\mathrm{H}}(\mathscr{C}) \geq 3$. Let $x, y, z \in \mathscr{C}$ be distinct codewords. Define $D(z,x) := \{1 \leq i \leq n \mid z_i \neq x_i\}$ and $D(z,y) := \{1 \leq i \leq n \mid z_i \neq y_i\}$. We have $|D(z,x)| \geq 3$ and $|D(z,y)| \geq 3$, hence $D(z,x) \cap D(z,y) \geq 6 - 4 = 2$. In particular, there exist $i, j \in \{1, ..., n\}$ with:
   - $i \neq j$;
   - $z_i \neq x_i$ and $z_i \neq y_i$;
   - $z_j \neq x_j$ and $z_j \neq y_j$.

   Since x, y, z have entries in $\mathbb{F}_2$, the last two properties force $x_i = y_i$ and $x_j = y_j$, from which (since $i \neq j$) $d^{\mathrm{H}}(x,y) \leq 4 - 2 = 2 < 3$, a contradiction.

Note: alternatively, one can use the fact that d^{H} is translation invariant (Exercise 1.26) and assume $x = (0, 0, 0, 0)$, simplifying the proof.

3. Take for example $\mathscr{C} = \{(0, 0, 0, 0), (1, 1, 1, 0), (0, 2, 2, 2)\}$. □

Proof of Proposition 2.20. Since $\mathscr{C}$ is linear, $0 \in \mathscr{C}$ and therefore

$$\{d^{\mathrm{H}}(x, y) \mid x, y \in \mathscr{C},\, x \neq y\} \supseteq \{d^{\mathrm{H}}(x, 0) \mid x \in \mathscr{C},\, x \neq 0\} = \{\omega^{\mathrm{H}}(x) \mid x \in \mathscr{C},\, x \neq 0\}.$$

It follows that $d^{\mathrm{H}}(\mathscr{C}) \leq \min\{\omega^{\mathrm{H}}(x) \mid x \in \mathscr{C},\, x \neq 0\}$. Vice versa, take $x, y \in \mathscr{C}$ with $x \neq 0$ and $d^{\mathrm{H}}(x, y) = d^{\mathrm{H}}(\mathscr{C})$. We have $x - y \in \mathscr{C}$ and $d^{\mathrm{H}}(\mathscr{C}) = d^{\mathrm{H}}(x, y) = d^{\mathrm{H}}(x - y, 0) = \omega^{\mathrm{H}}(x - y)$. Therefore

$$\min\{\omega^{\mathrm{H}}(x) \mid x \in \mathscr{C},\, x \neq 0\} \leq d^{\mathrm{H}}(\mathscr{C}),$$

concluding the proof. □

Solution to Exercise 2.55.

1. The result is immediate if $|\mathscr{C}| = 1$. Now suppose $|\mathscr{C}| \geq 2$. Define the linear map $\varphi : \mathbb{F}_q^n \to \mathbb{F}_q^{n+1}$ by $\varphi : (x_1, ..., x_n) \mapsto (x_1, ..., x_n, -x_1 - x_2 - \cdots - x_n)$. The code $\mathscr{C}^{\mathrm{ext}}$ is the image of $\mathscr{C}$ under φ. Since φ is injective, $\mathscr{C}^{\mathrm{ext}}$ has the same dimension as $\mathscr{C}$. Furthermore, for all $x \in \mathscr{C}$ we have $\omega^{\mathrm{H}}(\varphi(x)) \geq \omega^{\mathrm{H}}(x)$. Thus $d^{\mathrm{H}}(\mathscr{C}^{\mathrm{ext}}) \geq d^{\mathrm{H}}(\mathscr{C})$.

2. If $q = 2$, $\mathscr{C}$ is non-zero and $d = d^{\mathrm{H}}(\mathscr{C})$ is odd, then every minimum weight codeword $x \in \mathscr{C}$ satisfies $\omega^{\mathrm{H}}(\varphi(x)) = d + 1$. Therefore $\mathscr{C}^{\mathrm{ext}}$ has minimum distance $d + 1$. □

Proof of Proposition 2.61. Let $x \in \mathscr{C}_1$ and $y \in \mathscr{C}_2$ be codewords of weight d_1 and d_2 respectively. Then $(x, x), (0, y) \in \mathscr{C}_1 \oplus_{\mathrm{P}} \mathscr{C}_2$ and therefore $d^{\mathrm{H}}(\mathscr{C}) \leq \min\{2d_1, d_2\}$.

Now take any $x \in \mathscr{C}_1$ and $y \in \mathscr{C}_2$ such that $(x, x + y) \neq 0$. If $y = 0$ then $x \neq 0$ and therefore $(x, x + y)$ has weight at least $2d_1$. If $y \neq 0$, observe that for all i with $y_i \neq 0$ we have either $x_i \neq 0$ or $(x + y)_i \neq 0$ (or both). Therefore the weight of $(x, x + y)$ is at least $\omega^{\mathrm{H}}(y) \geq d_2$. To conclude, in any case the weight of $(x, x + y)$ is at least $\min\{2d_1, d_2\}$. □

Solution to Exercise 2.74. Let $\Lambda_n := \{\lambda \in \mathbb{F}_q^n \mid \lambda_i \neq 0 \text{ for all } 1 \leq i \leq n\}$ and let S_n denote the symmetric group on the set $\{1, ..., n\}$. Finally, let G_n be the group of linear Hamming metric isometries $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$.

Next, define a map $\varphi : \Lambda_n \times S_n \to G_n$ by $\varphi(\lambda, \tau) := f_\lambda \circ f_\tau$ for all $\lambda \in \Lambda_n$ and $\tau \in S_n$. The map φ is well-defined an surjective by Theorem 2.73. We claim that φ is also injective. Indeed, suppose that $\varphi(\lambda, \tau) = \varphi(\lambda', \tau')$ for some $\lambda, \lambda' \in \Lambda_n$ and $\tau, \tau' \in S_n$. In particular, for all $1 \leq i \leq n$ we have

$$\varphi(\lambda, \tau)(e_i) = \varphi(\lambda', \tau')(e_i),$$

i.e.,

$$\lambda_{\tau(i)} e_{\tau(i)} = \lambda'_{\tau'(i)} e_{\tau'(i)}.$$

This implies $\tau(i) = \tau'(i)$ and $\lambda_i = \lambda'_i$ for all $1 \leq i \leq n$. Therefore $\tau = \tau'$ and $\lambda = \lambda'$, which shows that φ is a bijection. In particular, the cardinality of G_n is $|\Lambda_n \times S_n| = (q-1)^n n!$. □

Solution to Exercise 2.84. 1. If the i-th and the j-th column of H are equal, with $i \neq j$, then $(e_i - e_j) \cdot H^\top = 0$. Therefore $e_i - e_j \in \mathscr{C}$ and $\mathscr{C}$ has minimum distance at most 2, a contradiction.

2. If $x \in \mathscr{C}$ was sent, $y \in \mathbb{F}_2^5$ is received and only one error occurred, then $y = x + e_i$ for some $1 \leq i \leq n$. Therefore $H \cdot y^\top = H \cdot (x + e_i)^\top = H \cdot x^\top + H \cdot e_i^\top = H \cdot e_i^\top$, where the latter product is precisely the i-th column of H. Since the columns of H are distinct, i can be retrieved by looking at H and at $H \cdot e_i^\top$. □

Solution to Exercise 2.89. The sum of the coefficients of the weight enumerator of a code $\mathscr{C}$ is precisely $|\mathscr{C}|$. The sum of the coefficients of $1Y^9 + 14X^3Y^6 + 16X^4Z^5 + 5X^7Y^2 + 10X^8Y + 9X^9$ is 55, which cannot be written in the form q^k, with q a prime power. Therefore there is no linear code with such a weight enumerator. □

Solution to Exercise 2.91. Denote by k the dimension of $\mathscr{C}$ and let $V := \{x \in \mathbb{F}_2^n \mid x_1 + x_2 + \cdots + x_n = 0\} \leq \mathbb{F}_q^n$. Note that V is a vector space of dimension $n-1$ over $\mathbb{F}_2$. The codewords of even weight in $\mathscr{C}$ are precisely the elements of $\mathscr{C} \cap V$. Moreover,

$$\dim(\mathscr{C} \cap V) = \dim(\mathscr{C}) + \dim(V) - \dim(\mathscr{C} + V) \geq k + n - 1 - n = k - 1.$$

Therefore $\dim(\mathscr{C} \cap V) \in \{k, k-1\}$. In particular, the number of codewords $x \in \mathscr{C}$ of even weight is either $2^k = |\mathscr{C}|$, or $2^{k-1} = |\mathscr{C}|/2$. □

Solution to Exercise 3.36. Since $k \geq 2$ and $d = 3$, by the Griesmer bound we have

$$n \geq \sum_{i=0}^{k-1} \lceil 3/2^i \rceil = 3 + 2 + \sum_{i=2}^{k-1} \lceil 3/2^i \rceil \geq 5 + (k-2) = k - 3,$$

from which $k \leq n - 3$. The Singleton bound gives instead $k \leq n - 2$. □

Partial answer to Exercise 5.18. To show the inclusion $\pi_S(\mathscr{C})^\perp \supseteq \pi_S(C^\perp(S))$, we fix $x \in \pi_S(C^\perp(S))$ and show that $\langle x, \pi_S(y) \rangle = 0$ for all $y \in \mathscr{C}$. Since $x \in \pi_S(C^\perp(S))$, there exists $z \in \mathscr{C}^\perp(S)$ with $x = \pi_S(z)$. Now observe that

$$\langle x, \pi_S(y) \rangle = \langle \pi_S(z), \pi_S(y) \rangle = \langle z, y \rangle,$$

where the latter equality follows from the fact that the Hamming support of z is contained in S. Since $z \in \mathscr{C}^\perp$ and $y \in \mathscr{C}$, we have $\langle x, \pi_S(y) \rangle = 0$, as desired. □

Answers to Exercise 8.10. 1. $(0,1,1,1,1)$ decodes to $(0,1,1,1,1)$;

2. $(1,1,1,1,1)$ decodes to $(0,1,1,1,1)$;

3. $(1, 0, 1, 1, 1)$ decodes to $(1, 0, 1, 1, 0)$;

4. $(1, 1, 1, 1, 0)$ decodes to $(1, 0, 1, 1, 0)$;

5. $(0, 0, 0, 0, 0)$ decodes to $(0, 0, 0, 0, 0)$.

□

Answers to Exercise 8.12. 1. $(3, 4, 2, 4)$ decodes to $(3, 4, 2, 4)$;

2. $(3, 4, 2, 1)$ decodes to $(3, 4, 2, 4)$;

3. $(0, 0, 0, 0)$ decodes to $(0, 0, 0, 0)$;

4. $(0, 0, 0, 2)$ decodes to $(0, 0, 0, 0)$;

5. $(0, 0, 1, 2)$ gives to a decoding failure.

□

Bibliography

[1] R. Ahlswede, N. Cai, S.-Y.R. Li, R.W. Yeung, *Network Information Flow*, IEEE Transactions on Information Theory 46 (2000), 4, pp. 1204 – 1216.

[2] V. Guruswami, A. Rudra, M. Sudan, *Essential Coding Theory*, in preparation (`https://cse.buffalo.edu/faculty/atri/courses/coding-theory/book/`).

[3] F. R. Kschischang, *An Introduction to Network Coding*, In *Network Coding: Fundamentals and Applications* (eds: M. Médard, A. Sprintson), Elsevier 2012.

[4] R. Lidl, H. Niederreiter, *Finite Fields*, Cambridge University Press 1997.

[5] F. J. MacWilliams, N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North Holland Mathematical Library 1977.

[6] R. Pellikaan, X.-W. Wu, S. Bulygin, R. Jurrius, *Codes, Cryptology and Curves with Computer Algebra*, Cambridge University Press 2018.

[7] A. Ravagnani, F. R. Kschischang, *Adversarial Network Coding*, IEEE Transactions on Information Theory 65 (2019), 1, pp. 198–219.

[8] C. Shannon, *A Mathematical Theory of Communication*, Bell System Technical Journal, vol. 27, no. 3, pp. 379 – 423, 1948.

[9] H. van Tilborg, *Coding Theory – A First Course*, Kluwer Academic Publishers (not printed anymore).