SPRINGER BRIEFS IN MATHEMATICS

Marcelo Firer · Marcelo Muniz S. Alves
Jerry Anderson Pinheiro
Luciano Panek

# Poset Codes: Partial Orders, Metrics and Coding Theory

SBMAC
Sociedade Brasileira de Matemática Aplicada e Computacional

Springer

# SpringerBriefs in Mathematics



**SpringerBriefs in Mathematics** showcases expositions in all areas of mathematics and applied mathematics. Manuscripts presenting new results or a single new result in a classical field, new field, or an emerging topic, applications, or bridges between new results and already published works, are encouraged. The series is intended for mathematicians and applied mathematicians.

More information about this series at http://www.springer.com/series/10030

# SBMAC SpringerBriefs



The **SBMAC SpringerBriefs** series publishes relevant contributions in the fields of applied and computational mathematics, mathematics, scientific computing, and related areas. Featuring compact volumes of 50 to 125 pages, the series covers a range of content from professional to academic.

The Sociedade Brasileira de Matemática Aplicada e Computacional (Brazilian Society of Computational and Applied Mathematics, SBMAC) is a professional association focused on computational and industrial applied mathematics. The society is active in furthering the development of mathematics and its applications in scientific, technological, and industrial fields. The SBMAC has helped to develop the applications of mathematics in science, technology, and industry, to encourage the development and implementation of effective methods and mathematical techniques for the benefit of science and technology, and to promote the exchange of ideas and information between the diverse areas of application.

http://www.sbmac.org.br/

Marcelo Firer • Marcelo Muniz S. Alves
Jerry Anderson Pinheiro • Luciano Panek

# Poset Codes: Partial Orders, Metrics and Coding Theory

Springer

Marcelo Firer
IMECC – Department of Mathematics
University of Campinas
Campinas, São Paulo, Brazil

Jerry Anderson Pinheiro
IMECC – Department of Mathematics
University of Campinas
Campinas, São Paulo, Brazil

Marcelo Muniz S. Alves
Polytechnic Center – Department
of Mathematics
Federal University of Paraná
Curitiba, Paraná, Brazil

Luciano Panek
Center of Exact Sciences and Engineering
State University of West Paraná
Foz do Iguaçu, Paraná, Brazil

ISSN 2191-8198 ISSN 2191-8201 (electronic)
SpringerBriefs in Mathematics
ISBN 978-3-319-93820-2 ISBN 978-3-319-93821-9 (eBook)
https://doi.org/10.1007/978-3-319-93821-9

Library of Congress Control Number: 2018946658



This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

In 1995, Brualdi et al. [14] introduced the concept of a metric on a vector space determined by a partial order over a finite set, a *poset*. These are the so-called *poset metrics*, that generalize both the usual Hamming metric and the Niederreiter–Rosenbloom–Tsfasman metric ([84] and [96]). All the invariants and all the main questions of coding theory posed considering the traditional Hamming metric may be defined and asked when considering any metric determined by a poset. This is the subject of this text.

The Hamming metric (determined by an anti-chain) resembles (through many of its properties) an Euclidean metric, while the Niederreiter–Rosenbloom–Tsfasman metric (determined by a chain) is an ultrametric. A general poset mixes the properties of those two extremal cases: there are metric balls that have a unique center (a property of coordinate-wise additive metrics), but it is also possible to find metric balls for which each of its points is a center (an ultra-metric property). These mixed-properties of a general poset metrics make its study, as combinatorial and geometric object, an interesting and challenging area of research, with many questions suggested by coding theory, the original field where it first appeared. Moreover, general poset metric throws a new light into many of the classical invariants of coding theory (such as minimum distance, packing and covering radius) and many of its basic results (concerning perfect and MDS codes, MacWilliams' identity, syndrome decoding, and so on), in such a way that it contributes to a better understanding of these invariants and properties when considering the classical and nearly ubiquitous Hamming metric.

This is the first title to give a systematic approach of poset metrics and codes and it covers a significant part of the literature in the area and most of the research problems developed in recent years. It is motivated by the necessity to organize the knowledge cumulated in the last two decades, and this task became possible after some works that managed to give a clearer understanding of phenomena, which started with the study of particular cases. This is the case, for example, of a description of the group of linear isometries, a clear and deep understanding about the MacWilliams type identities and an overall understanding of the metrics determined by a hierarchical poset.

It is important to remark that in many cases we shortened the proofs of propositions, omitting some technical details and focusing on the main underlying ideas. Also, each chapter (with the exception of the last one) has a section of final notes which briefly points the past (and sometimes a possible future) development in the area and also a list of exercises, some of them are actually a guided script (broken into some items) to the proof of propositions stated in the main body of the text.

**An Overview of the Content** In order to make the presentation as self-contained as possible, we start with a brief recollection of basic concepts about coding theory (Chap. 1), which can be skipped (except for the notation) by anyone who has ever learned something about the subject. It is followed by an introduction to the most basic definitions about partially ordered sets—posets—and the definition of a metric determined by a poset (Chap. 2). In this chapter we also introduce two of the most emblematic families of posets, the hierarchical and the multi-chain posets. Chapter 3 is devoted to the study of the coding theory invariants when the space is endowed with a hierarchical poset metric. As we shall see, the hierarchical poset metric resembles the Hamming metric and most of the cornerstone propositions in coding theory have an amicable version in this context. This is what we call the "easy case." As expected, the easy case is followed, in Chap. 4, by the simplest "difficult case," the case of metrics determined by a multi-chain poset. It deserves a separate chapter since this case has a simple description of the minimum distance and admits a decomposition that is not as good as the canonical decomposition of hierarchical posets, but allows some advance in the study of MDS and perfect codes. Chapter 5 is devoted not to a specific poset metric, but to a specific topic, namely, duality. The highlight of the chapter is the deep understanding of the role of different equivalence relations between ideals in the relation between the weight distribution of a code and the dual code. This result (Theorem 5.9) was presented in a work that has not been published yet, we guess because it carries more new understandings than new results. Besides the importance of the MacWilliams identities for its sake, it is an opportunity to present some of the many different techniques of discrete mathematics that may be used to prove it. In Chap. 6 we approach the case of a general poset. We start discussing the difficulty of determining the packing radius of a code: it is as difficult as we can imagine. Considering this level of difficulty, we move on and use the simplest cases (hierarchical poset metrics) to produce bounds for this important invariant. Finally, Chap. 7 is devoted to many generalizations of a poset metric (poset-block metrics, metrics defined by a directed graph, partially ordered multisets) and one major variation (combinatorial metrics). The presentation in this chapter is in a different style, there are no proofs or examples, only the concepts and definitions necessary to understand the main results. This sharp writing is balanced by the novelty of many of those families of metrics (most of the advances were made after 2015), what, naturally, leaves a lot of open questions for those interested in the topics. Finally, an epilogue that can be seen as a brief sketch of a research statement is presented.

# Contents

# Chapter 1
# Basic Concepts of Coding Theory


We briefly introduce some very basic facts about the theory of *error correcting codes* or *coding theory*. To be more precise, we introduce those basic facts that are related to or dependent on the concept of metric. Many hypothesis are presented in a simplified form or in a restricted context, except for the concept of metric, which is considered in full generallity, not really specified at this point, since this is the main goal of these notes: to show the role that different metrics may have in the context of coding theory and to present some interesting questions and bias to the study of finite metric spaces that arouse from coding theory problems and invariants.

Coding theory considers a *set of possible input messages* $\mathcal{X}$, a *set of possible output messages* $\mathcal{Y}$ and a probabilistic *channel model* given by a *transition matrix* $\mathbb{P} = (P_{y,x})_{y\in\mathcal{Y},x\in\mathcal{X}}$, where

$$P_{y,x} := P(y = \text{received}|x = \text{sent})$$

is the conditional probability that the message $y \in \mathcal{Y}$ is received given that the message $x \in \mathcal{X}$ is sent. Since we are assuming those quantities to be conditional probabilities, we have that $P_{y,x} \geq 0$ for all $y \in \mathcal{Y}$, $x \in \mathcal{X}$ and $\sum_{y\in\mathcal{Y}} P_{y,x} = 1$, for all $x \in \mathcal{X}$.

In the context of coding theory, both the input and the output sets are considered to be finite (so that the transition matrix is indeed a matrix) and we may always consider $\mathcal{X} \subseteq \mathcal{Y}$.

The channel is seen as a physical media through which messages are transmitted and there is a non-zero probability of errors, that is, there is a probability that the message received is not equal to the message that has been sent. The basic goal of coding theory is to detect and correct such errors. Since the messages are considered to be meaningless (no semantic issues to be considered—this is a basic assumption stated by Shannon in the second paragraph of the seminal work [99]), every possible way of detecting errors demands adding some redundancy to the information we wish to transmit. To be precise, we consider a *code* to be proper



M. Firer et al., *Poset Codes: Partial Orders, Metrics and Coding Theory*, SpringerBriefs in Mathematics, https://doi.org/10.1007/978-3-319-93821-9_1

subset $C \subseteq \mathcal{X} \subseteq \mathcal{Y}$. It is considered as the set of *admissible messages* and its elements are called *codewords*. Whenever a message $y \in \mathcal{Y}$ is received, we check if $y \in C$ or not. If $y \notin C$ we can be sure there was an error and then, we can try to correct it. It means we need a map $\texttt{Dec} : \mathcal{Y} \to C$, called a *decoder*, that associates to a received message a codeword. The optimal way to do it (when no other information is available) is to maximize the probability of $y$ being received, that is, we should choose $\texttt{Dec}(y) \in C$ such that

$$P(y \mid \texttt{Dec}(y)) \geq P(y|x), \forall x \in C.$$

This is a decision criterion, and an algorithm which implements it is called a *maximum likelihood decoder* (ML-decoder). We remark that maximum likelihood decoding is not necessarily unique, in the sense that

$$\operatorname*{argmax}_{x \in C} P(y|x) := \{z \in C;\, P(y|z) \geq P(y|x), \forall x \in C\}$$

may be a set with more than one element. However, given the triple $(\mathcal{X}, \mathcal{Y}, \mathbb{P})$, an ML-decoder is a *universal* decoder, in the sense that it may be used for every code $C \subseteq \mathcal{X}$ and every $y \in \mathcal{Y}$. Another source of universal decoders are the metric decoders, available if $\mathcal{X} = \mathcal{Y}$. Given a metric $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, a *nearest neighbor decoder* (NN-decoder) is a map $\texttt{Dec}_d : \mathcal{X} \to C$ that associates to a received message a codeword that is as close as possible:

$$d(y, \texttt{Dec}_d(y)) \leq d(y, x), \forall x \in C,$$

or, in other words,

$$\texttt{Dec}_d(y) \in \operatorname*{argmin}_{x \in C} d(y, x) := \{z \in C; d(y, z) \leq d(y, x), \forall x \in C\}.$$

For a given metric and code, an NN-decoder is, in general, not unique, but as the ML-decoder, it is a family of universal decoders.

If we have a channel model $\mathbb{P}$ and a metric model $d(\cdot, \cdot)$, a pair of corresponding ML-decoder and NN-decoder does not need to coincide. In case we have that

$$\operatorname*{argmax}_{x \in C} P(y|x) = \operatorname*{argmin}_{x \in C} d(y, x), \forall C \subseteq \mathcal{X}, x \in \mathcal{X}$$

we say that $(\mathcal{X}, \mathbb{P})$ is a *matched pair*.

Given a channel $\mathbb{P}$, it is not always possible to find a metric $d$ such that $(\mathcal{X}, \mathbb{P})$ is a *matched pair* . However, the most common channel and the most used metric in the context of coding theory are matched to each other, as we shall see in the sequence.[1]

[1] The possibility of matching channels and metrics in a general setting is a question that, apparently, was first purposed by Jim Massey [79]. In 1980, it was proved in [98] that symmetric memoryless

## 1.1 Symmetric Channels and the Hamming Metric

Since we are interested in metric aspects, in these notes we assume that the input set and the output set coincide. Moreover, we consider an *alphabet* set $\mathcal{A}$ and the $n$-fold of this alphabet, which is denoted by $\mathcal{X} = \mathcal{A}^n$, that is, the elements of the input-output set $\mathcal{X}$ may be viewed as *words* $x = (x_1, x_2, \ldots, x_n)$ with $n$ *letters* taken from the alphabet $\mathcal{A}$.

The $n$-fold *Symmetric Channel* (SC) over an alphabet $\mathcal{A}$ with $q$ elements is a channel whose transition matrix is defined as follows: given $0 \leq p \leq 1$ and $x_i, y_i \in \mathcal{A}$ the error probability is uniform, equal to $p/(q-1)$, that is,

$$P(y_i|x_i) = \begin{cases} 1-p \text{ if } x_i = y_i \\ p/(q-1) \text{ if } x_i \neq y_i \end{cases}.$$

The channel is considered to be memoryless because the received symbols (letters in the alphabet) have no influence in the error probability of subsequent symbols, that is, given $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$, with $x, y \in \mathcal{A}^n$,

$$P(y|x) = \prod_{i=1}^{n} P(y_i|x_i).$$

Here, the quantity $p$ is called the *error probability* and it would be reasonable to assume that $p/(q-1) < 1-p$. This means that, in case there is no redundancy at all and there are no (non-semantic) tools to verify or correct errors, once a message is received (a sequence of letters in the alphabet), it is better to accept it than to pick any other sequence of letters. A stronger condition would be to impose that $p \leq 1-p$, or equivalently, to impose that $0 \leq p \leq 1/2$. We remark that both the conditions coincide when the alphabet is binary ($q = 2$): this is the *Binary Symmetric Channel* (BSC).

The symmetric channels (the binary ones in particular) have a singular importance in the context of coding theory, since assuming a channel to be memoryless and symmetric means to assume the worst possible scenario: if we have no real information about a channel, we should assume it is memoryless and symmetric, since it is the most difficult case, the one with more ambiguities.

We can understand the symmetric channels considering the *Hamming metric* $d_H : \mathcal{A}^n \times \mathcal{A}^n \to \{0, 1, \ldots, n\}$, which counts the number of different entries in the arrays of $\mathcal{A}^n$: for $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$, with $x, y \in \mathcal{A}^n$,

$$d_H(x, y) := |\{i \in [n]; x_i \neq y_i\}|,$$

channels can be matched to additive metrics. The question was left untouched for a long time and some interest on it have recently been raised [40, 92, 94].

where $[n] := \{1, 2, \ldots, n\}$ and $|X|$ is the cardinality of a finite set $X$. It is immediate to check that $d_H$ is indeed a metric.[2]

The probability $P(y|x)$ of receiving $y \in \mathcal{A}^n$ given that $x \in \mathcal{A}^n$ was sent may be expressed using the Hamming metric:

$$P(y|x) = \left(\frac{p}{q-1}\right)^{d_H(y,x)} (1-p)^{n-d_H(y,x)}.$$

Considering the function $f(t) = \left(\frac{p}{q-1}\right)^t (1-p)^{n-t}$, it is immediate to verify that its derivative is negative for $p < 1 - p$, hence, for $0 \leq p \leq 1/2$ we have that $P(y|x)$ increases *iff* $d_H(x, y)$ decreases, so that the $n$-fold SC and the Hamming metric on $\mathcal{A}^n$ constitute a matched pair.

From here on, we assume that $\mathcal{A}^n = \mathbb{F}_q^n$ is an $n$-dimensional vector space (over a finite field with $q$ elements $\mathbb{F}_q$) and denote vectors using boldface letters. This allows us to exchange the condition $x_i = y_i$ by the equivalent condition $x_i - y_i = 0$, so that the Hamming metric may be expressed as

$$\begin{aligned} d_H(\mathbf{x}, \mathbf{y}) &= |\{i \in [n]; x_i \neq y_i\}| \\ &= |\{i \in [n]; x_i - y_i \neq 0\}|. \end{aligned}$$

Due to the matching of the $q$-ary ($\mathcal{A} = \mathbb{F}_q$) $n$-fold SC and the Hamming metric, much of the (metric) coding theory is developed using the Hamming metric, and many properties that depend on this particular choice of metric are taken for granted. As we shall see, using other metrics may cause some discomfort to those that are used to general texts in coding theory. This discomfort will be explored in a significant part of the text and will, hopefully, lead the reader to a new comfort-zone.

Given a subset $A \subseteq \mathbb{F}_q^n$ we define the *support*

$$\text{supp}(A) = \{i \in [n]; x_i \neq 0 \quad \text{for} \quad \text{some} \quad \mathbf{x} = (x_1, \ldots, x_n) \in A\}$$

to be the set of non-null coordinate positions of the elements of $A$. If $A = \{\mathbf{x}\}$, we shall abuse of the notation and write $\text{supp}(A) = \text{supp}(\mathbf{x})$. With this definition we can express the Hamming metric as

$$d_H(\mathbf{x}, \mathbf{y}) = |\text{supp}(\mathbf{x} - \mathbf{y})| = d_H(\mathbf{0}, \mathbf{x} - \mathbf{y}).$$

The quantity $\varpi_H(\mathbf{x}) := d_H(\mathbf{0}, \mathbf{x})$ is called the *Hamming weight of* $\mathbf{x}$. It is easy to verify that it satisfies the following conditions

[2] For all $x, y, z \in \mathcal{X} = \mathcal{A}^n$, the following conditions are satisfied: *(1)* $d(x, y) \geq 0$ *(non-negativity)*; *(2)* $d(x, y) = 0$ *iff* $x = y$ *(identity of indicernibles)*; *(3)* $d(x, y) = d(y, x)$ *(symmetry)*; *(4)* $d(x, z) \leq d(x, y) + d(y, z)$ *(triangle inequality)*.

**Weight Conditions**

1. $\varpi_H(\mathbf{x}) \geq 0$ *(non-negativity)*;
2. $\varpi_H(\mathbf{x}) = 0$ *iff* $\mathbf{x} = \mathbf{0}$ *(identity of indicernibles)*;
3. $\varpi_H(\mathbf{x}) = \varpi_H(\lambda\mathbf{x}), \forall 0 \neq \lambda \in \mathbb{F}_q$ *(local homogeneity)* and in particular, for $\lambda = -1$, $\varpi_H(\mathbf{x}) = \varpi_H(-\mathbf{x})$ *(symmetry)*;
4. $\varpi_H(\mathbf{x} + \mathbf{y}) \leq \varpi_H(\mathbf{x}) + \varpi_H(\mathbf{y})$ *(triangle inequality)*.

A function satisfying the weight conditions is called a *weight over* $\mathbb{F}_q^n$ and it is straightforward to prove that, if $\varpi$ is a weight over $\mathbb{F}_q^n$, then the function defined by $d(\mathbf{x}, \mathbf{y}) = \varpi(\mathbf{y} - \mathbf{x})$ is a metric on $\mathbb{F}_q^n$. In this case, we say that $d$ is a *weight-defined metric*. We remark that a metric determined by a weight is *invariant by translations*, in the sense that $d(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z}) = d(\mathbf{x}, \mathbf{y})$, for every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}_q^n$. Another way to state it is to say that a translation $T_\mathbf{z}$ by a vector $\mathbf{z}$, defined as $T_\mathbf{z}(\mathbf{x}) = \mathbf{x} + \mathbf{z}$, is an *isometry* of the metric space $(\mathbb{F}_q^n, d)$, in the sense that $d(T_\mathbf{z}(\mathbf{x}), T_\mathbf{z}(\mathbf{y})) = d(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}_q^n$.

We also remark that the Hamming metric and weight assume only the integer values in $[n]$ and they *respect the support* of vectors, in the sense that if $\text{supp}(\mathbf{x}) \subseteq \text{supp}(\mathbf{y})$ then $\varpi_H(\mathbf{x}) \leq \varpi_H(\mathbf{y})$ or equivalently, $\text{supp}(\mathbf{x} - \mathbf{y}) \subseteq \text{supp}(\mathbf{x}' - \mathbf{y}')$ implies $d_H(\mathbf{x}, \mathbf{y}) \leq d_H(\mathbf{x}', \mathbf{y}')$. From here on, unless explicitly stated, we shall consider only metrics spaces $(\mathbb{F}_q^n, d)$ satisfying the following conditions:

- The metric $d$ is determined by some weight function $\varpi$;
- $d$ assumes only the integer values in $[n]$;
- $d$ respects the support of vectors.

## 1.2 Metric Invariants of Codes

Although most of the concepts we are concerned with can be defined for any code $C \subseteq \mathcal{A}^n$, when considering $\mathcal{A}^n = \mathbb{F}_q^n$ to be a vector space, it is natural and convenient to consider $C \subset \mathbb{F}_q^n$ to be a vector subspace.

We start with some definitions.

**Definition 1.1** An $(n, M)_q$ *(error correcting) code* is a subset $C \subset \mathbb{F}_q^n$ with $|C| = M$ elements. In case $C$ is a $k$-dimensional vector subspace of $\mathbb{F}_q^n$ we say it is an $[n, k]_q$ *linear code* (the square brackets indicate the linearity). In this situation, $M = q^k$. The elements of $C$ are called *codewords*.

**Definition 1.2** Given a metric $d$ over $\mathbb{F}_q^n$, the *minimum distance* $\delta_d(C)$ of a code is the minimum distance between distinct codewords:

$$\delta_d(C) := \min\{d(\mathbf{x}, \mathbf{y}); \mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}\}.$$

Let $C$ be an $(n, M)_q$ code with minimum distance $\delta_d(C)$. Suppose a message $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is sent and an error of size $r \leq \lfloor(\delta_d(C) - 1)/2\rfloor$ occurs, that is,

a message $\tilde{\mathbf{x}} = \mathbf{x} + \mathbf{e}$ is received and $\varpi(\mathbf{e}) = r$. Here, $\varpi(\mathbf{x}) = d(\mathbf{x}, \mathbf{0})$ is the weight determined by the distance and $\lfloor \cdot \rfloor$ is the floor function: $\lfloor t \rfloor$ is the greatest integer smaller or equal to $t$.

Under this circumstance, whatever is the metric to be considered over $\mathbb{F}_q^n$, using NN decoding will return the correct message. Indeed, let us denote by $B_d(\mathbf{x}, r)$ the (closed) metric ball with center at $\mathbf{x}$ and radius $r$:

$$B_d(\mathbf{x}, r) := \{\mathbf{y} \in \mathbb{F}_q^n; d(\mathbf{x}, \mathbf{y}) \leq r\}.$$

Then, the symmetry of the metric ensures that $\mathbf{x} \in B_d\left(\tilde{\mathbf{x}}, \lfloor (\delta_d(C) - 1)/2 \rfloor\right)$. Moreover, given another codeword $\mathbf{x} \neq \mathbf{y} \in C$, the triangular inequality ensures that $\mathbf{y} \notin B_d\left(\tilde{\mathbf{x}}, \lfloor (\delta_d(C) - 1)/2 \rfloor\right)$, hence $\operatorname{argmin}_{\mathbf{y} \in C} d(\tilde{\mathbf{x}}, \mathbf{y}) = \{\mathbf{x}\}$. So, an NN-decoder decides $\mathtt{Dec}_d(\mathbf{x} + \mathbf{e}) = \mathbf{x}$ if $\varpi(\mathbf{e}) \leq \lfloor (\delta_d(C) - 1)/2 \rfloor$. Ensuring that a code can correct errors up to some size is an important property, thus the minimum distance is many times incorporated to the code notation, and we write $(n, M, \delta_d)_q$ to say that the $(n, M)_q$ code $C$ has minimum distance $\delta_d = \delta_d(C)$. A symmetric way to look at this situation is to state that the balls centered at the codewords with radius $R = \lfloor (\delta_d(C) - 1)/2 \rfloor$ are all disjoint. Indeed, suppose there $\mathbf{z} \in B_d(\mathbf{x}, R) \cap B_d(\mathbf{y}, R)$. The triangular inequality ensures that $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y}) \leq \delta_d(C)$ and the minimality property of $\delta_d(C)$ ensures that $\mathbf{x} = \mathbf{y}$.

**Definition 1.3** Let $C$ be an $(n, M)_q$ code and consider a metric $d$ over $\mathbb{F}_q^n$. The *packing radius* $R_d(C)$ of $C$ relatively to the metric $d$ is

$$R_d(C) := \max\{r \in [n]; B_d(\mathbf{x}, r) \cap B_d(\mathbf{y}, r) = \emptyset, \text{ for all } \mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}\}.$$

**Proposition 1.4** *Let $d$ be a metric over $\mathbb{F}_q^n$. Given an $(n, M, \delta_d)_q$ code $C$, the packing radius $R_d(C)$ satisfies the inequalities*

$$\left\lfloor \frac{\delta_d(C) - 1}{2} \right\rfloor \leq R_d(C) \leq \delta_d(C) - 1.$$

*Proof* The first inequality was just proved. The second inequality follows from the fact that $d$ assumes only integer values and that there exist $\mathbf{x}, \mathbf{y} \in C$ such that $d(\mathbf{x}, \mathbf{y}) = \delta_d(C)$. ■

It is well known that, for the Hamming metric, the first inequality turns into an equality, that is,

$$\left\lfloor \frac{\delta_{d_H}(C) - 1}{2} \right\rfloor = R_{d_H}(C).$$

For those that are used to the Hamming or Lee metrics (see Exercise 5 for the definition of a Lee metric), it may be surprising that all the integer values between $\lfloor (\delta_d(C) - 1)/2 \rfloor$ and $\delta_d(C) - 1$ may occur as a packing radius of a code. More surprising may be the fact that the packing radius is not determined by the minimum

distance and even more surprising it may be to realize how difficult is to determine the packing radius, even to tiny codes. Those properties will be well explored in Chap. 2, but we introduce now a small teaser example.

*Example 1.5* Let $\mathcal{X} = \mathbb{F}_2^3$ and consider on it the function $\varpi$ defined in the table:

| $\mathbf{x}$ | 000 | 100 | 010 | 110 | 001 | 101 | 011 | 111 |
|---|---|---|---|---|---|---|---|---|
| $\varpi(\mathbf{x})$ | 0 | 1 | 1 | 2 | 2 | 2 | 3 | 3 |

We should check that $\varpi$ is a weight and the metric $d$ defined by $\varpi$ satisfies the conditions stated in the end of Sect. 1.1: it assumes only the integer values 0, 1, 2, 3 and it respects the support of vectors. Now, looking at the weight table we see that the vectors $\mathbf{x} = 110$ and $\mathbf{y} = 001$ have both weight 2. It follows that the codes $C = \{\mathbf{0}, \mathbf{x}\}$ and $C' = \{\mathbf{0}, \mathbf{y}\}$ have both minimum distance equal to 2. To determine the packing radii of the codes, we list the elements of the balls $B_d(\mathbf{0}, 1) = \{000, 100, 010\}$, $B_d(\mathbf{x}, 1) = \{110, 010, 100\}$ and $B_d(\mathbf{y}, 1) = \{001, 101, 011\}$ and we find that

$$B_d(\mathbf{0}, 1) \cap B_d(\mathbf{x}, 1) = \{010\} \Rightarrow R_d(C) = 0;$$
$$B_d(\mathbf{0}, 1) \cap B_d(\mathbf{y}, 1) = \emptyset \Rightarrow R_d(C') \geq 1.$$

Once we know the packing radius $R_d(C)$ of a code, we know that the balls of radius $R_d(C)$ centered at the codewords are all disjoint, so NN decoding gives us a unique decision criterion once the received message belongs to one of these balls. However, there may be elements $\mathbf{x} \in \mathbb{F}_q^n$ that do not belong to any of those balls. The *covering radius* $R_d^{cov}(C)$ of a code is the minimal radius that ensures any $\mathbf{x} \in \mathbb{F}_q^n$ will belong to one of these balls:

$$R_d^{cov}(C) := \min\{r \geq 0; \bigcup_{\mathbf{x} \in C} B_d(\mathbf{x}, r) = \mathbb{F}_q^n\}.$$

It is easy to see that

$$R_d(C) \leq R_d^{cov}(C) \leq n.$$

When the packing and the covering radius are equal, that is, when there is an $R = R_d(C) = R_d^{cov}(C)$ such that

$$B_d(\mathbf{x}, R) \cap B_d(\mathbf{y}, R) = \emptyset, \forall\, \mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y},$$

and

$$\bigcup_{\mathbf{x} \in C} B_d(\mathbf{x}, R) = \mathbb{F}_q^n,$$

we say that $C$ is a *perfect code*, or $d$*-perfect*, when it is necessary to stress the metric.

*Example 1.6* We consider on $\mathbb{F}_2^2$ the Hamming metric $d_H$ and the metric $d_T$ determined by the following weights:

$$\varpi_T(00) = 0,\ \varpi_T(10) = 1,\ \varpi_T(01) = \varpi_T(11) = 2.$$

Considering the code $C = \{00, 01\}$, we have that

$$B_{d_H}(00, 1) = \{00, 10, 01\}, \qquad B_{d_H}(01, 1) = \{01, 11, 00\}$$
$$B_{d_T}(00, 1) = \{00, 10\}, \qquad B_{d_T}(01, 1) = \{01, 11\}$$

and we see that $C$ is $d_T$-perfect but not $d_H$-perfect.

Considering the Hamming metric, perfect codes are very seldom, they are actually classified: Hamming Codes (which will be introduced in the next section) and Golay Codes are the unique non trivial $d_H$-perfect codes (see [77]). Since the property of being perfect is rare (at least in the usual Hamming distance framework), we measure how far or close a code is from being perfect by considering the difference $0 \leq R_d^{cov}(C) - R_d(C)$, called the *d-defect of* $C$. A code $C$ is said to be $(d, i)$*-quasi-perfect* if its $d$-defect is $i = R_d^{cov}(C) - R_d(C)$. In particular, we say that a 0-quasi-perfect and a 1-quasi-perfect codes are perfect and *quasi-perfect*, respectively.

Considering the Hamming metric, it is not difficult to see that a ball of radius $r$ has $|B_{d_H}(\mathbf{0}, r)| = \sum_{i=0}^{r} \binom{n}{i}(q-1)^i$ elements, where $\binom{n}{i}$ is the binomial coefficient. If we denote by $\mathcal{A}_q(n, \delta_{d_H})$ the maximum possible size of an $(n, M, \delta_{d_H})_q$ code, it is not difficult to see that

$$\mathcal{A}_q(n, \delta_{d_H}) \leq \frac{q^n}{\sum_{i=0}^{\lfloor (\delta_{d_H}-1)/2 \rfloor} \binom{n}{i}(q-1)^i}.$$

This is known as the *Hamming bound*, or sphere packing bound, and it is possible to say that a code is $d_H$-perfect *iff* it attains this bound. In general, if we denote by $\mathcal{A}_q(n, \delta_d)$ the maximum possible size of an $(n, M, \delta_d)_q$ code, relative to a weight-defined metric, it is bounded by

$$\frac{q^n}{|B_d(\mathbf{0}, \delta_d - 1)|} \leq \mathcal{A}_q(n, \delta_d) \leq \frac{q^n}{|B_d\left(\mathbf{0}, \lfloor (\delta_d - 1)/2 \rfloor\right)|},$$

what is a much weaker statement than in the Hamming case.

The Hamming bound is established once the minimum distance is given, but it says nothing about bounds to the minimum distance. The Singleton bound, proved by Richard C. Singleton in 1964 [100], states that the Hamming minimum distance

of a $q$-ary code of length $n$ with $q^k$ elements is at most $n-k+1$, that is, $\delta_{d_H} \leq n-k+1$. As a direct consequence of this bound we have

$$\mathcal{A}_q(n, \delta_{d_H}) \leq q^{n-\delta_{d_H}+1}.$$

Considering a general distance $d$, we say that an $(n, k, \delta_d)_q$ code is *maximum distance separable* ($d$-MDS) code if $\delta_d = n-k+1$. As the perfect codes, being an MDS code is also a very good property in what relates to correcting errors, and, for the binary case, the only $d_H$-MDS codes have parameter $k \in \{1, n-1, n\}$ and are determined as

$$C_1 = \{\mathbf{0}, 11\cdots 11\},$$

$$C_{n-1} = \{\mathbf{x} \in \mathbb{F}_q^n; \sum_{i-1}^{n} x_i = 0\} \quad \text{and} \quad C_n = \mathbb{F}_q^n,$$

for which $\delta_{d_H}$ is equal to $n$, 2 and 1, respectively. We remark that $d_H$-MDS codes are abundant for very large fields.

In a similar way we did for perfect codes, we can measure how far a code is from being an MDS code by considering the *$d$-Singleton defect* of an $(n, k)_q$ code $C$, defined as $n-k+1-\delta_d(C)$. Then, we say that a code is *$(d, i)$-almost MDS* (or just $(d, i)$-AMDS) if $i = n-k+1-\delta_d(C)$.

Other ways to measure how far a code is of being MDS (or $d$-MDS) are available considering the generalized weights of linear codes, to be introduced in the next section.

## 1.3 Linear Codes

Given an $[n, k, \delta_d]_q$ linear code $C$ and an ordered basis

$$\beta = \{(x_{11}, x_{12}, \ldots, x_{1n}), (x_{21}, x_{22} \ldots, x_{2n}), \ldots, (x_{k1}, x_{k2}, \ldots, x_{kn})\}$$

the matrix

$$G = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k1} & x_{k2} & \ldots & x_{kn} \end{pmatrix}$$

is called a *generator matrix* of $C$. Of course there are as many generator matrices as ordered basis. Among them, using the usual Gaussian row elimination and possible permutation of columns, we get a generator matrix in a row echelon form

$$G = \begin{pmatrix} 1\,0\,\ldots\,0\,x_{1,k+1}\,x_{1,k+2}\,\ldots\,x_{1,n} \\ 0\,1\,\ldots\,0\,x_{2,k+1}\,x_{2,k+2}\,\ldots\,x_{2,n} \\ \vdots\,\vdots\,\ddots\,\vdots\quad\vdots\quad\quad\vdots\quad\ddots\quad\vdots \\ 0\,0\,\ldots\,1\,x_{k,k+1}\,x_{k,k+2}\,\ldots\,x_{k,n} \end{pmatrix}$$

called a *systematic form* of the generator matrix. We remark that not every code has a generator matrix in the systematic form, but since permutation of coordinates is a linear isometry of the Hamming metric (see Exercise 9), we can say that any code is *Hamming-equivalent* to a code that admits a systematic generator matrix. Since $C$ is linear, every element of $C$ may be seen as the product $\mathbf{u}G$, where $\mathbf{u} = (u_1, u_2, \ldots, u_k) \in \mathbb{F}_q^k$.

Considering $C$ as the kernel of a linear map $f : \mathbb{F}_q^n \to \mathbb{F}_q^{n-k}$, it may be represented by an $(n-k) \times n$ matrix $H$ and, by definition, whenever $\mathbf{x} \in C$, we have that $H\mathbf{x}^T = \mathbf{0}$, so that, in particular for a generator matrix, we have that $HG^T = 0$. Any such matrix is called a *parity check matrix* and, if $G = (Id_{k\times k}|A_{k\times(n-k)})$ is a generator matrix in a systematic form, the matrix $H = (-A^T_{(n-k)\times k}|Id_{n-k})$ is a parity chek matrix, since $HG^T = -A^T + A^T = 0$.

Considering linear codes, the task of determining the minimum distances and the packing radius becomes a bit simplified, since we can consider the weight of vectors instead of the distance between pairs of vectors:

**Proposition 1.7** *Let $d$ be a metric on $\mathbb{F}_q^n$ defined by a weight $\varpi$ and let $C \subseteq \mathbb{F}_q^n$ be a linear code. Then,*

$$\delta_d(C) = \min\{\varpi(\mathbf{x}); \mathbf{0} \neq \mathbf{x} \in C\}$$

*and*

$$R_d(C) = \max\{r \geq 0;\, B_d(\mathbf{0}, r) \cap B_d(\mathbf{x}, r) = \emptyset, \forall \mathbf{0} \neq \mathbf{x} \in C\}.$$

*Proof* For every metric and any code, by definition of the minimum distance we have that $\delta_d(C) \leq \min\{d(\mathbf{0}, \mathbf{x}) = \varpi(\mathbf{x}); \mathbf{x} \in C\}$. Since $d$ is defined by a weight, it is invariant by translations and $d(\mathbf{y}, \mathbf{x}) = d(\mathbf{0}, \mathbf{x} - \mathbf{y}) = \varpi(\mathbf{x} - \mathbf{y})$. Assuming that $C$ is linear, given $\mathbf{x}, \mathbf{y} \in C$ we have that $\mathbf{x}' = \mathbf{x} - \mathbf{y} \in C$, so we get that $\min\{d(\mathbf{y}, \mathbf{x}); \mathbf{x} \in C\} \geq \min\{d(\mathbf{0}, \mathbf{x}') = w(\mathbf{x}'); \mathbf{x}' \in C\}$. It follows that $\delta_d(C) \geq \min\{d(\mathbf{x}_0, \mathbf{x}); \mathbf{x} \in C\}$ and hence $\delta_d(C) = \min\{d(\mathbf{x}_0, \mathbf{x}); \mathbf{x} \in C\}$. The claim about the packing radius follows in a similar manner. ■

Proposition 1.7 gives us a justification to call $\delta_d(C)$ also as the *minimal weight* of the code.

The maximality condition in the definition of the packing radius ensures that there is $\mathbf{x} \in C$ satisfying $B_d(\mathbf{0}, R_d(C) + 1) \cap B_d(\mathbf{x}, R_d(C) + 1) \neq \emptyset$. Any such vector $\mathbf{x} \in C$ is called a *packing vector*.

If we wish to compute the minimum distance or the packing radius of a code, Proposition 1.7 makes our life, in principle, easier: instead of looking on a table (the lookup table) with $|C| \times |C|$ entries and comparing those entries (that may be the distance between two elements or the packing radius considering only two elements), we have reduced our search to one single line of such a table. The problem is that such a line has $q^k = |C|$ elements and, in the context of coding theory, $k$ is typically very large (order of hundreds of thousands or more). In fact, determining the minimum distance is an intractable problem (an NP-Hard problem), as was proved by Vardy [107, in 1997]. As we shall see later, computing the packing radius of a code, for a general metric, is actually an intractable family of intractable problems!

Also, in the general context of coding theory, we aim to have the least possible redundancy, that is, it is desirable to have the *rate* $k/n$ of a linear code to be as large (close to 1) as possible, which makes it difficult to compute invariants depending on large $k$ and large rate. For this reason (among many other possible ones), we introduce the *dual* $C^\perp$ of a code $C$, which is defined as

$$C^\perp := \{\mathbf{x} \in \mathbb{F}_q^n;\ \mathbf{x} \cdot \mathbf{y} = 0, \forall\, \mathbf{y} \in C\},$$

where $\mathbf{x} \cdot \mathbf{y} := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$ is the formal inner product between vectors with coordinates $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$.

**Proposition 1.8** *Let $C$ be an $[n, k]_q$ linear code. Let $G$ and $H$ be a generator and a parity check matrix of $C$. Then:*

1. *$C^\perp$ is an $[n, n-k]_q$ linear code;*
2. *$H$ is a generator matrix of $C^\perp$;*
3. *$G$ is a parity check matrix of $C^\perp$;*
4. *$(C^\perp)^\perp = C$.*

*Proof*

1. The formal inner product $\cdot : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{F}$ is a bilinear map and $C^\perp$ is the intersection of the kernels of the linear maps $\mathbf{x}^i(\mathbf{y}) = \mathbf{x}^i \cdot \mathbf{y}$ where $\beta = \{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k\}$ is an ordered base of $C$. Since $\beta = \{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k\}$ is a linearly independent set, it follows that $\dim(C^\perp) = n - k$.
2. Let $\alpha = \{\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^{n-k}\}$ be the set of vectors represented by the rows $H$. A general element of $C$ is represented in the basis $\beta$ (defined in the previous item) as $\mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{x}^i$ and, since

$$\mathbf{x} \cdot \mathbf{y}^i = \sum_{i=1}^n (\lambda_i \mathbf{x}^i) \cdot \mathbf{y}^i = \sum_{i=1}^n \lambda_i (\mathbf{x}^i \cdot \mathbf{y}^i) = 0,$$

   we have that $\alpha \subseteq C^\perp$, so it generates a subcode of $C^\perp$. Since the lines of $H$ are assumed to be linearly independent, we have that $H$ is a generator matrix of $C^\perp$.
3. Follows in a similar way as the previous item.

4. Let $\mathbf{x} \in C$. Given $\mathbf{y} \in C^{\perp}$, the definition of $C^{\perp}$ ensures that $\mathbf{y}\cdot\mathbf{x} = \mathbf{x}\cdot\mathbf{y} = 0$, hence $\mathbf{x} \in (C^{\perp})^{\perp}$ and we have that $C \subseteq (C^{\perp})^{\perp}$. Since $\dim\left((C^{\perp})^{\perp}\right) = n - \dim\left(C^{\perp}\right) = n - (n - \dim(C)) = \dim(C)$, it follows that $(C^{\perp})^{\perp} = C$.

■

We remark that for large rate $k/n$, the rate $(n-k)/n$ of the dual code is small, so computing invariants such as the minimum distance, the generalized weights and the packing radius, becomes a relatively simpler problem. As we will see in Sect. 1.3.1, when considering the Hamming metric, it is possible to relate the invariants of a code to the invariants of its dual.

The parity check matrix can also be used for estimating the minimum distance of the code.

**Proposition 1.9 ([48, Corollary 1.4.14])** *If $C$ is a linear code and $H$ is a parity check matrix of $C$, then $\delta_{d_H}(C) = d$ if and only if every set of $d-1$ columns of $H$ is linearly independent and there is a set of $d$ linearly dependent columns of $H$.*

*Proof* See Exercise 13. ■

We close this section with two important families of linear codes.

**Binary Hamming Codes**

Let $H_m$ be the matrix which has as columns all nonzero vectors of $\mathbb{F}_2^m$, arranged in "increasing order"[3] in the following sense: the $j$-th column of $H_m$ is the vector $\mathbf{h}^j \in \mathbb{F}_2^m$ which corresponds to the binary expansion of $j = \alpha_1 2^0 + \alpha_2 2^1 + \cdots + \alpha_m 2^{m-1}$. For $m = 3$, for instance,

$$H_3 = \begin{bmatrix} 1\,0\,1\,0\,1\,0\,1 \\ 0\,1\,1\,0\,0\,1\,1 \\ 0\,0\,0\,1\,1\,1\,1 \end{bmatrix}.$$

The Hamming code $\mathcal{H}(m)$ is the kernel of this matrix, and since $H_m$ is an $m \times (2^m - 1)$ matrix of rank $m$, this code has length $2^m - 1$ and dimension $2^m - m - 1$. Its Hamming minimum distance is 3, regardless of $m$, since any two columns of $H_m$ are linearly independent and there are (many) three linearly dependent columns: if $\mathbf{h}^1$ and $\mathbf{h}^2$ are two columns of $H_m$ then $\mathbf{h}^1 + \mathbf{h}^2 \neq \mathbf{0}$ and hence is a third column of $H_m$, and the set $\{\mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^1 + \mathbf{h}^2\}$ is linearly dependent. In particular, $R_{d_H}(\mathcal{H}(m)) = 1$. Since

$$|\mathcal{H}(m)|\,|B_{d_H}(\mathbf{0}, 1)| = 2^m = |\mathbb{F}_2^m|,$$

these codes are all $d_H$-perfect.

[3] In fact, this ordering is not relevant, since a different ordering would lead to an equivalent code, in the sense it yields a code with the same minimum distance (see Exercise 9).

**Reed-Solomon Codes**

There are several possible constructions of this family of codes; we present one that does not require any knowledge about the structure of finite fields. Let $\mathbb{F}_q$ be the finite field of $q$ elements, with $q > 2$, let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be distinct elements of $\mathbb{F}_q$ and let $r$ be an integer such that $1 \leq r < n$. Let $\mathbb{P}(r)$ be the vector space of polynomials in the variable $z$ with coefficients in $\mathbb{F}_q$ and degree at most $r$ (including here the zero polynomial), which is a vector space of dimension $r + 1$.

The Reed-Solomon code $C$ associated to this data is the image of the mapping $\phi : \mathbb{P}(r) \to \mathbb{F}_q^n$ given by $\phi(f) = (f(\alpha_1), \ldots, f(\alpha_n))$. This mapping is linear and injective, and hence $C$ is an $[n, r+1]_q$ linear code. Given that $\{1, z, z^2, \ldots, z^r\}$ is a basis of $\mathbb{P}(r)$, the matrix $G = ((\alpha_i)^j)_{\substack{i=1,\ldots,n \\ j=0,\ldots,r}}$ is a generator matrix for this code. With respect to its Hamming minimum weight, note that each $f \in \mathbb{P}(r)$ has at most $r$ roots, and therefore

$$\delta_{d_H}(C) = n - r.$$

It follows that $\delta_{d_H}(C) + k = (n - r) + (r + 1) = n + 1$ and $C$ is a $d_H$-MDS code.

### *1.3.1 Duality and Generalized Weights*

The minimal weight, or minimum distance, is just one (the first) among a set of relevant invariants of a (linear) code. Since we are considering only distances and weights assuming integer values in $[n]$, given $i \in [n]$ we denote by $A_i := A_{d,i}$ the number of codewords with weight $i$: $A_i := |\{\mathbf{x} \in C; \varpi(\mathbf{x}) = i\}|$. Being $C$ a linear code, we have that $A_0 = 1$, $A_i = 0$ for $0 < i < \delta_d(C)$ and $A_{\delta_d} \neq 0$. All this information is represented by the *weight enumerator*, a formal polynomial

$$W_\varpi(C, z) = \sum_{\mathbf{x} \in C} z^{\varpi(\mathbf{x})} = \sum_{i=0}^{n} A_i z^i.$$

We remark that the weight enumerator polynomial contains quite much information about a code, but it is not enough, in general, to characterize it: very different codes may have the same weight enumerator (see Exercise 7).

Considering $\varpi = \varpi_H$ to be the Hamming weight, the weight enumerator of a code and its dual are related by the *MacWilliams Identity*:

**Theorem 1.10** *Let $C$ be a linear code and $C^\perp$ its dual. Let*

$$W_{\varpi_H}(C, z) = \sum_{i=0}^{n} A_i z^i \quad \text{and} \quad W_{\varpi_H}(C^\perp, z) = \sum_{i=0}^{n} A_i^\perp z^i.$$

*Then,*

$$W_{\varpi_H}(C,z) = \frac{1}{|C^\perp|}\sum_{i=0}^{n} A_i^\perp (1+(q-1)z)^{n-i}(1-z)^i.$$

This theorem was proved by F.J. MacWilliams in 1961 [75] in her Ph.D. Thesis. The theorem was generalized by Delsarte [21] and MacWilliams et al. [78] in 1972, to non-linear codes, but in this case, instead of the weight enumerator, what is considered is the distance enumerator $W_{d_H}(C,z) = \sum_{\mathbf{x},\mathbf{y}\in C} z^{d_H(\mathbf{x},\mathbf{y})}$. It has many extensions to different contexts, including codes over rings and modules (see for example the lecture notes of Jay A. Wood [111]). The proof of this theorem is omitted at this point, since MacWilliams' identity and other duality questions will be developed in Chap. 5.

Considering now linear codes with the Hamming metric, we observe that, for every $0 \neq \lambda \in \mathbb{F}_q$ and every $\mathbf{x} \in \mathbb{F}_q^n$, $\varpi(\mathbf{x}) = \varpi(\lambda\mathbf{x})$. It follows that we may express the Hamming weight $\varpi_H(\mathbf{x})$ as the number of non-zero coordinates in the 1-dimensional space generated by $\mathbf{x}$, that is,

$$\varpi_H(\mathbf{x}) = |\mathrm{supp}(\llbracket\{\mathbf{x}\}\rrbracket)|,$$

where $\llbracket A\rrbracket$ is the vector subspace of $\mathbb{F}_q^n$ generated by the set $A$ and $\mathrm{supp}(\llbracket A\rrbracket)$ is the support of the set, as defined in Sect. 1.1. Out of this observation, in 1991, Victor K. Wei [108] generalized the concept of weight, as follows: Given a subset $A \subseteq \mathbb{F}_q^n$, its *norm* $\|A\|$ is the number of elements in its support, that is, $\|A\| := |\mathrm{supp}(A)|$. It follows that

$$\varpi_H(\mathbf{x}) = \|\llbracket\{\mathbf{x}\}\rrbracket\|$$

so that, for a linear code $C$, we have that

$$\begin{aligned}\delta_{d_H}(C) &= \min\{\varpi_H(\mathbf{x});\, \mathbf{0} \neq \mathbf{x} \in C\}\\ &= \min\{\|\llbracket\{\mathbf{x}\}\rrbracket\|;\, \mathbf{0} \neq \mathbf{x} \in C\}\\ &= \min\{\|D\|;\, D \subseteq C, \dim(D) = 1\},\end{aligned}$$

where $D \subseteq C$ is assumed to be a linear subcode. With this observation it becomes natural to define the *$i$-th generalized $d_H$-weight* of an $[n,k]_q$ code as

$$\delta_{i,d_H}(C) := \min\{\|D\|;\, D \subseteq C, \dim(D) = i\}.$$

As we had already remarked, the first generalized $d_H$-weight is just the minimum distance: $\delta_{1,d_H}(C) = \delta_{d_H}(C)$. It is not difficult to prove that the sequence of weights is strictly increasing:

**Theorem 1.11 ([108, Theorem 1])** *Let $C$ be a linear code. Then,*

$$\delta_{1,d_H}(C) < \delta_{2,d_H}(C) < \cdots < \delta_{k,d_H}(C).$$

*Proof* See Exercise 12. ■

We call $\{\delta_{1,d_H}(C), \delta_{2,d_H}(C), \ldots, \delta_{k,d_H}(C)\}$ the *weight hierarchy* of the code $C$. The weight hierarchy of a code and its dual are related by Wei's Duality Theorem:

**Theorem 1.12 ([108, Theorem 3])** *Let $C$ be a linear code. Then,*

$$\{\delta_{i,d_H}(C); i = 1, 2, \ldots, k\} = \{1, 2, \ldots, n\} \backslash \{n+1-\delta_{i,d_H}(C^{\perp}); i = 1, 2, \ldots, n-k\}.$$

Again, this theorem will not be proved at this moment, since many generalizations will be proved in subsequent chapters.

### *1.3.2 Syndrome Decoding*

When considering a metric structure $d$ over $\mathbb{F}_q^n$, we consider NN-decoding ($d$-NN-decoding, when needed to stress the role of the metric $d$): when a message $\mathbf{y}$ is received, we look for $\mathbf{z} \in C$ (not necessarily unique) such that $d(\mathbf{y}, \mathbf{z}) \leq d(\mathbf{y}, \mathbf{x})$, for all $\mathbf{x} \in C$. Implementing such a criterion as a search algorithm is possible, but it implies that, for a given $\mathbf{y} \in \mathbb{F}_q^n$, it is necessary to look and compare all the $|C|$ different entries $d(\mathbf{y}, \mathbf{x}), \mathbf{x} \in C$. If the metric is invariant by translations, a welcomed shortcut is available for "small" errors.

Let us assume $d$ is a metric defined by a weight $\varpi$ and $C$ is an $[n, k]_q$ linear code with packing radius $R_d(C)$. Since $C$ is linear, the cosets $\mathbf{y} + C := \{\mathbf{y} + \mathbf{x}; \mathbf{x} \in C\}, \mathbf{y} \in \mathbb{F}_q^n$, are either disjoint or coincide, that is, for every $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$, if $(\mathbf{y} + C) \cap (\mathbf{z} + C) \neq \emptyset$ then $\mathbf{y} + C = \mathbf{z} + C$. Moreover, since $C$ is closed by vectors sum, we have that $\mathbf{y} + C = \mathbf{z} + C$ if, and only if, $\mathbf{z} - \mathbf{y} \in C$. For each coset $\mathbf{y} + C$ there is a vector $\mathbf{e}_{\mathbf{y}}$ of minimal weight $\varpi\left(\mathbf{e}_{\mathbf{y}}\right)$ such that $\mathbf{y} + C = \mathbf{e}_{\mathbf{y}} + C$. Such a vector is called a *coset leader* and it is assured to be unique if $\varpi\left(\mathbf{e}_{\mathbf{y}}\right) \leq R_d(C)$. From the minimality of the coset leader, we have that, once a message $\mathbf{y} \in \mathbf{e} + C$ is received, NN criterion decodes $\mathtt{Dec}_d(\mathbf{y}) = \mathbf{y} - \mathbf{e}$ where $\mathbf{e} \in \mathbf{y} + C$ is a coset leader.

Suppose we have a table of coset leaders and we wish to implement such a decoding. Given a parity check matrix $H$ of the code $C$, we call the *syndrome* of a vector $\mathbf{y}$ the vector $s(\mathbf{y}) := H\mathbf{y}^T$. The syndromes characterize the cosets of a linear code.

**Proposition 1.13** *Given vectors $\mathbf{y}, \mathbf{z} \in \mathbb{F}_q^n$ we have that $\mathbf{y} + C = \mathbf{z} + C$ if, and only if, $s(\mathbf{y}) = s(\mathbf{z})$.*

*Proof* We have that $\mathbf{y} + C = \mathbf{z} + C$ if, and only if, $\mathbf{y} - \mathbf{z} \in C$ and this happens if, and only if,

$$\begin{aligned} 0 &= H(\mathbf{y} - \mathbf{z})^T \\ &= H\mathbf{y}^T - H\mathbf{z}^t = s(\mathbf{y}) - s(\mathbf{z}). \end{aligned}$$

■

So, we can implement NN decoding by using syndromes. First thing we should have a table with all syndromes $\left\{s(\mathbf{e}_1), s(\mathbf{e}_1), \ldots, s(\mathbf{e}_{q^{n-k}})\right\}$. When a message $\mathbf{y}$ is received, we compute its syndrome $s(\mathbf{y}) = H\mathbf{y}^T$, look in the table and find $\mathbf{e}_i$ such that $s(\mathbf{y}) = s(\mathbf{e}_i)$ and then decode $\mathtt{Dec}_d(\mathbf{y}) = \mathbf{y} - \mathbf{e}_i$. This process is called *syndrome decoding*, and it is actually an algorithm that implements NN-decoding.

We remark that the search table has $q^{n-k}$ elements and, in general, this may still be a very long list, even for rates $k/n$ approaching 1. In future chapters we shall see how different metric structures may help reducing the search list of syndrome decoding.

## 1.4 Chapter Notes

In this chapter we stressed the difference between the minimal distance $\delta_d(C)$ and the packing radius $R_d(C)$ of a code, calling the attention that, in general, those invariants are not a function one of the other (as it is in the case of the Hamming metric). It happens that the packing radius is the more relevant invariant for predicting the correction capability of a code. For this reason, it is surprising that an analogue of the weight enumerator was never explored for this invariant, the *packing radii enumerator*, which can be easily defined and deserves to be studied.

We should also make a disclaimer about decoding algorithms, a very important aspect of coding theory that is being nearly ignored in this text. Syndrome decoding is an algorithm of very large complexity, so many codes, such as cyclic and LDPC codes, were developed (and used in practice) not necessarily for their correction capability, but for the availability of efficient decoding algorithms.

## 1.5 Exercises

1. Prove that the Hamming metric is indeed a metric.
2. Prove that every weight over $\mathbb{F}_q^n$ determines a metric over $\mathbb{F}_q^n$.
3. Prove that not every metric over $\mathbb{F}_q^n$ is determined by a weight.
4. Given a metric $d : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{R}$ invariant by translations. Prove that $\varpi(\mathbf{x}) := d(\mathbf{0}, \mathbf{x})$ is a weight.
5. Consider $\mathbb{Z}_m = \{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}$ the ring of integers modulo $m$, with usual operations. The $m$-norm of $\overline{l} \in \mathbb{Z}_m$ is $|l|_m = \min\{l, m-l\}$. Given $\mathbf{x} = (x_1, x_2, \ldots x_n) \in \mathbb{Z}_m^n$, its *Lee weight* is defined as $\varpi_{\text{Lee}}(\mathbf{x}) = \sum_{i=1}^n |x_i|_m$. The *Lee metric* $d_{\text{Lee}}$ on $\mathbb{Z}_m^n$ is defined as $d_{\text{Lee}}(\mathbf{x}, \mathbf{y}) = \varpi_{\text{Lee}}(\mathbf{x} - \mathbf{y})$.

   (a) Prove that $d_{\text{Lee}}$ is a metric.
   (b) If $C \subset \mathbb{Z}_m^n$, endowed with the Lee metric. Prove that its minimum distance $\delta_{d_{\text{Lee}}}(C)$ and its packing radius $R_{d_{\text{Lee}}}(C)$ are related by $R_{d_{\text{Lee}}}(C) = \lfloor(\delta_{d_{\text{Lee}}}(C) - 1)/2\rfloor$.

6. Consider the $[3, 1]_2$ codes $C_1 = \{000, 111\}$, $C_2 = \{000, 011\}$ and $C_3 = \{000, 001\}$. Determine weights $\varpi, \varpi'$ (and corresponding metrics $d, d'$) such that: a) $R_d(C_1) = R_d(C_2) = R_d(C_3) = 1$, b) $R_{d'}(C_1) < R_{d'}(C_2) < R_{d'}(C_3)$.
7. Let $d$ be a metric determined by a weight that respect the supports of vectors. Show that the Singleton bound holds, that is, given a code $C \subseteq \mathbb{F}_q^n$, then $|C| \leq q^{n-\delta_d(C)+1}$.
8. Let $d$ be a metric defined by a weight on $\mathbb{F}_q^n$ and let $C \subseteq \mathbb{F}_q^n$ be a linear code. Prove that

$$R_d(C) = \max\{r \geq 0;\, B_d(\mathbf{x}_0, r) \cap B_d(\mathbf{x}, r) = \emptyset, \forall \mathbf{x}_0 \neq \mathbf{x} \in C\}$$

   where $\mathbf{x}_0 \in C$ is any arbitrary codeword. Show that this is not necessarily true if $C$ is not linear.
9. Given an $n \times n$ matrix $N$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{F}_q^n$, we define $N(\mathbf{x}) = \mathbf{x}N$ (the usual product of matrices). Prove that $N$ defines a linear isometry on $(\mathbb{F}_q^n, d_H)$, that is, $d_H(N(\mathbf{x}), N(\mathbf{y}) = d_H(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$, if, and only if, $N = AB$ where $A$ is a diagonal invertible matrix and $B$ is a matrix obtained from the identity by permuting its columns.
10. Consider on $\mathbb{F}_2^3$ the metrics $d_1$ and $d_2$ defined by the weights $\varpi_1$ and $\varpi_2$ described in the tables

| $\mathbf{x}$ | 000 | 100 | 010 | 110 | 001 | 101 | 011 | 111 |
|---|---|---|---|---|---|---|---|---|
| $\varpi_1(\mathbf{x})$ | 0 | 1 | 1 | 2 | 2 | 2 | 3 | 3 |

and

| $\mathbf{x}$ | 000 | 100 | 010 | 110 | 001 | 101 | 011 | 111 |
|---|---|---|---|---|---|---|---|---|
| $\varpi_2(\mathbf{x})$ | 0 | 1 | 2 | 2 | 3 | 3 | 3 | 3 |

respectively.

   (a) Describe explicitly all the seven different $[3, 2]_2$ linear codes $C_1, C_2, \ldots C_7$.
   (b) Compute $W_{d_j}(C_i, z)$ for $j = 1, 2$ and $i = 1, 2, \ldots, 7$.
   (c) Describe explicitly (as matrices) the 2 different linear isometries $T, S$ of the metric space $(\mathbb{F}_2^3, d_1)$. Show there are codes $C_i$ and $C_j$ such that $W_{d_1}(C_i, z) = W_{d_1}(C_j, z)$ but $T(C_1) \neq C_2$ and $S(C_1) \neq C_2$. Conclude that for the metric $d_1$ the weight enumerator is not enough to characterize a code (up to equivalence). **Hint:** *one of these isometries is the identity, all you need is to find a second one, looking at possible symmetries of the weight table.*
   (d) Describe explicitly (as matrices) all the 8 different linear isometries of the metric space $(\mathbb{F}_2^3, d_2)$. Show that if $W_{d_2}(C_i, z) = W_{d_2}(C_j, z)$ then there is one of those linear isometries $T$ that maps $C_1$ into $C_2$. Conclude that for the metric $d_2$ the weight enumerator is enough to characterize a code (up to equivalence). **Hint:** *take some time and many blank pages. It is a*

*hardwork to do it at this point, but you will start getting acquainted with the "environment" of the following chapters.*

11. Give an example of a metric $d$ and a linear code $C$ for which the weight of a packing vector is not minimal, that is, there is a packing vector $\mathbf{x}_0 \in C$ such that $B_d(\mathbf{0}, R_d(C)+1) \cap B_d(\mathbf{x}_0, R_d(C)+1) \neq \emptyset$ but $\varpi(\mathbf{x}_0) > \delta_d(C)$.
12. Given an $[n,k]_q$ code $C$, prove that $\delta_{1,d_H}(C) < \delta_{2,d_H}(C) < \ldots < \delta_{k,d_H}(C)$. **Hint:** *Given* $D \subset C$ *such that* $\|D\| = \delta_{r,d_H}(C)$, *take* $i \in \mathrm{supp}(D)$ *and consider* $D_i = \{\mathbf{x} \in D; x_i = 0\}$.
13. ([48, Cor. 1.4.14]) Prove Proposition 1.9: if $C$ is a linear code and $H$ is a parity check matrix of $C$, then $\delta_{d_H}(C) = d$ if and only if every set of $d-1$ columns of $H$ is linearly independent and there is a set of $d$ linearly dependent columns of $H$. **Hint:** *Consider the fact that* $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in C$ *iff* $H\mathbf{c}^T = 0$, *and if* $\mathbf{h}^1, \mathbf{h}^2, \ldots, \mathbf{h}^n$ *are the column vectors of* $H$ *then* $H\mathbf{c}^T = c_1\mathbf{h}^1 + \cdots + c_n\mathbf{h}^n$.

# Chapter 2
# Poset Metrics


Partially ordered sets is a central concept in set theory (hence in many branches of mathematics) and a subject of study on its own. In the context of coding theory it appears as an auxiliary structure that allows to determine different metrics over $\mathbb{F}_q^n$, metrics that satisfy the conditions stated at the end Sect. 1.1: they assume only integer values, are determined by a weight and respect the support of vectors. As we shall see, many properties concerning the metrics can be derived from properties of the poset.

In this chapter, we give a relatively detailed introduction to the main subject of this text, the poset metrics.

We start (Sect. 2.1) introducing the basic concepts about partial orders over finite set (the only instance of interest in this text), introducing the Hasse diagram (which allows to develop some intuition on the subject) and presenting two families of posets, the hierarchical posets and the multi-chain posets. These families of posets play a significant role in the subject of this book so that we devoted one chapter to each of them (Chaps. 3 and 4). The first of these families corresponds to the poset metrics that are a true generalization of the Hamming metric, and they are as understood as the Hamming metric. The second one is the most simple example of the difficulties we face in the general case.

In Sect. 2.2 we introduce the poset metrics and start exploring the two main invariants of coding theory, the minimum distance and the packing radius of a code, showing simple ways of how to determine the weight of a vector in the case of a hierarchical or multi-chain poset.

Finally, on Sect. 2.3 we give a description of the group of linear isometries of the space $\mathbb{F}_q^n$ endowed with a poset metric. Despite the fact that the proofs are relatively lengthy, we present it in some details, since some of the steps of the proof characterizes a "way of thinking" that are used repeatedly along the text. Again, we present in details the characterization of the isometries for our two main examples.



M. Firer et al., *Poset Codes: Partial Orders, Metrics and Coding Theory*, SpringerBriefs in Mathematics, https://doi.org/10.1007/978-3-319-93821-9_2

## 2.1 Partial Orders Over Finite Sets

We start this chapter with the most basic definitions that entitles this text: a *partial order relation* over a set $X$ is a binary relation $\preceq$ satisfying the following conditions:

1. (Reflexivity) $i \preceq i$ for all $i \in X$;
2. (Anti-symmetry) Given $i, j \in X$, if $i \preceq j$ and $j \preceq i$, then $i = j$;
3. (Transitivity) If $i \preceq j$ and $j \preceq k$, then $i \preceq k$.

The pair $P = (X, \preceq)$, consisting of a non-empty set $X$ and a partial order $\preceq$ over $X$, is called a *partially ordered set*, or, for short, a *poset*. We say that $i, j \in X$ are *comparable* (in $P$) if either $i \preceq j$ or $j \preceq i$.

In this work we are concerned with finite posets, that is, orders over finite sets. It follows that, up to isomorphism, we may assume that $X = [n] = \{1, 2, \ldots, n\}$ for some positive integer $n$. In this context, of posets over $X = [n]$, we will denote the elements of $[n]$ by the letters $a, b$ or $i, j, k, l$ and avoid the use of $u, v$ or $x, y, z$ that will be used to denote the coordinates of vectors. Also, we will denote subsets of $[n]$ by using the corresponding capital letters $A, B$ or $I, J, K, L$. This may be a little bit misleading, since many of the definitions and concepts we will introduce holds also for infinite posets, but this will prevent many notation confusions.

The prototype of a partial order is the inclusion order: given a set $X$, let $\mathcal{P}(X)$ be the set of all subsets of $X$ and consider the poset $(\mathcal{P}(X), \subseteq)$, where $\subseteq$ is the usual set inclusion relation. We remark that if $X$ is not a singleton, that is, if it contains more than one element, then, for $i \in X$, the sets $\{i\}$ and $X \setminus \{x\}$ are not comparable, in the sense that none is contained in the other. The adjective *partial*, in this context, relates exactly to this property, when elements are not necessarily comparable. A partial order under which every pair of elements is comparable is called a *total order* or *linear order*. The prototype of a total order is the usual order relation $\leq$ of the real numbers.

In what follows, whenever no confusion may arise, we will omit the index and write $\preceq$ for $\preceq_P$. The symbols $\leq$ and $\subseteq$ are used for the usual order relation over $\mathbb{R}$ and $\mathcal{P}(X)$ respectively, and the restriction of such orders to subsets.

An *order homomorphism* between two posets $P = (X, \preceq_P)$ and $Q = \left(Y, \preceq_Q\right)$ is a map $f : X \to Y$ that preserves the order relation, in the sense that $i \preceq_P j$ implies that $f(i) \preceq_Q f(j)$, for all $i, j \in X$. An *order isomorphism* between two posets $P = (X, \preceq_P)$ and $Q = \left(Y, \preceq_Q\right)$ is a bijective homomorphism $f : X \to Y$ such that the inverse $f^{-1} : Y \to X$ is also a homomorphism. In other words, a poset isomorphism is a bijection $f : X \to Y$ such that $i \preceq_P j$ if, and only if, $f(i) \preceq_Q f(j)$. A *poset automorphism* is an isomorphism $f : X \to X$, where $P = (X, \preceq_P)$ is a poset. The set of all poset automorphisms of a poset $P$ is a group, denoted by $\mathrm{Aut}(P)$. Sometimes we may denote an automorphism of poset by $f : P \longrightarrow P$, since in most part of the text the underlying set will be $X = \{1, 2, \ldots, n\}$.

Given a poset $P = (X, \preceq_P)$ and a subset $Y \subseteq X$, the restriction of $\preceq_P$ to $Y$ is called the *induced order* or *restricted order*. To emphasize this inclusion relation we

may write $P_X = (X, \preceq_P)$ and $P_Y = (Y, \preceq_P)$. In this situation, we say that $P_Y$ is a *subposet* of $P_X$, or shortly, that $Y$ is a subposet of $X$, and denote it as $P_Y \subseteq P_X$.

The most important characteristic a subposet may have is being closed by the order relation:

**Definition 2.1** A subposet $I \subseteq X$ is called an *ideal* of $P$ if it satisfies the *closeness property*: if $i \in I$ and $j \preceq i$, then $j \in I$.

Given a poset $P = (X, \preceq)$ and a subset $A \subset X$, the smallest ideal containing $A$ is called the *ideal generated by* $A$ and denoted by $\langle A \rangle_P$, the subindex $P$ will be omitted if no confusion may arise. If $A = \{i\}$ for some $i \in X$, $\langle A \rangle$ is called a *prime ideal*.

A *chain* in a poset $P = (X, \preceq_P)$ is a subset $A \subseteq X$ such that the subposet $P_A$ is totally ordered, that is, any two elements in $A$ are comparable. An *anti-chain* is a subset $A \subseteq X$ for which different elements are not comparable, that is, given $i, j \in A$ we have that $i \preceq_P j$ if and only if $i = j$.

We say that $j$ *covers* $i$ (in $P$) if $i \preceq_P j$ but there is no element between $i$ and $j$, in the sense that $i \preceq_P k \preceq_P j$ implies that $k = i$ or $k = j$. Given a chain $A \subseteq X$, we may label its elements as $a_1, a_2, \ldots, a_l$ in such a way that $a_i \preceq a_j$ *iff* $i \leq j$. We say that $l$ is the *length of the chain*. We say that the chain is *saturated* if there is no elements between $a_i$ and $a_{i+1}$, that is, if $a_{i+1}$ covers $a_i$, as elements in $X$, for every $1 \leq i < l$.

Given a poset $P = (X, \preceq_P)$, the *P-height* (or just *height*, also known in the literature as rank of the poset) $h(a)$ of an element $a \in X$ is the maximal length of a chain that has $a$ as a maximal element. The *height* $h(P)$ *of the poset* is the maximal height of its elements, that is, $h(P) = \max\{h(a); a \in X\}$. The *i-level* $H_i := H_i(P)$ of $P$ is the set of elements of height $i$:

$$H_i(P) := \{a \in X; h(a) = i\}.$$

It is clear from the definition that the levels of $P$ define a partition of $X$.

A finite poset may be described by its *Hasse diagram*, named after the German mathematician Helmut Hasse. Let us consider a directed graph that has $X$ as its set of vertices and an oriented edge $\overrightarrow{ba}$ connecting $b$ to $a$ *iff* $b$ covers $a$. The Hasse diagram is obtained from such a graph by considering a planar realization, where a vertex $a \in X$ corresponds to a point $(x_a, y_a) \in \mathbb{R}^2$ in such a way that points at the same level has the same $y$-coordinates and points at a higher levels are placed "above" the lower levels, that is, $y_a < y_b$ *iff* $h(a) < h(b)$. This last condition allows to represent the oriented edges of the graph as simple line segments, where the agreement is that $b$ "*is above*" $a$ means $a \preceq_P b$.

Drawing a Hasse diagram is a delicate task. A "good" diagram is a diagram that emphasizes interesting properties. Consider, for example, the poset $P = (\mathcal{P}([3]), \subseteq)$. In Fig. 2.1 we see three different Hasse diagrams of $P$. On the left side, we see a Hasse diagram that emphasizes the symmetries of the poset, what helps to determine the automorphism group $\mathrm{Aut}(P)$. In the middle we see a diagram that shows how two copies of the same diagram (considering only the

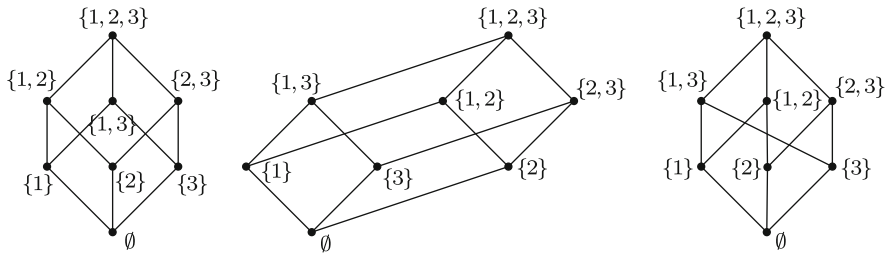


**Fig. 2.1** Three different Hasse diagrams of the poset $P = (\mathcal{P}([3]), \subseteq)$

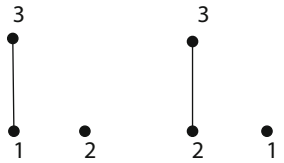


**Fig. 2.2** Two different natural labels to the same poset over [3]

vertices $\emptyset$, {1}, {3}, {1, 3} and {2}, {1, 2}, {2, 3}, {1, 2, 3}) are joined together. This diagram helps us to understand how we can draw a Hasse diagram for the poset $(\mathcal{P}([4]), \subseteq)$ out of $(\mathcal{P}([3]), \subseteq)$. The diagram on the right shows that, considering only the second and the third level, we have a crown poset (see Example 2.6).

A *natural labeling* of a Hasse diagram is an association $i \mapsto a_i$ between $[n]$ and the elements of $X$ such that $a_i \preceq a_j$ implies $i \leq j$. We remark that natural labeling is not unique (as one can see in Fig. 2.2). However, different labels (either natural or not) determine different but isomorphic posets.

As any mathematical structure, also for posets, it is important to determine substructures. In the case of a poset $P = ([n], \preceq)$, we can do it by restricting the underlying set $[n]$, as we did when defining subposets and ideals. Another approach is to maintain the underlying set $X$ unchanged and to consider different set of relations. Given two posets $P$ and $Q$ over the same set $X$, we say that $P$ is *coarser* then $Q$ (or equivalently, $Q$ is *finer* then $P$) if $i \preceq_P j$ implies $i \preceq_Q j$, that is, every relation in $P$ is also a relation in $Q$.

The easiest way to picture this relation between orders is to consider the incidence matrix of a poset.

**Definition 2.2** Given a poset $P = ([n], \preceq)$, the *incidence matrix* $M_P = (p_{ij})$ of $P$ is the $n \times n$ matrix with entries defined by

$$p_{ij} = \begin{cases} 1 \text{ if } i \preceq_P j \\ 0 \text{ otherwise} \end{cases}.$$

We define the *support of a poset* $P$ as $\text{supp}_P := \{(i, j) \in [n] \times [n];\ p_{ij} \neq 0\}$. It is clear that $P$ is coarser then $Q$ if, and only if, $\text{supp}_P \subseteq \text{supp}_Q$. Since the inclusion

is a partial order over any set, in particular for $[n] \times [n]$, we have that *being coarser* determines a poset structure on the set of all posets over $[n]$:

$$\mathcal{P}_n := \{P = ([n], \preceq); \preceq \text{ is a partial order over } [n]\}.$$

We denote by $P \leq Q$ the relation given by $\text{supp}_P \subseteq \text{supp}_Q$, so that $(\mathcal{P}_n, \leq)$ is itself a poset. This poset has a unique minimal element, the anti-chain poset, for which $\text{supp}_P = \{(i, i); i \in [n]\}$, and it has $n!$ different, but isomorphic (depending only on the labeling), maximal elements, the chain posets over $[n]$. If we consider the natural labeling of a chain, we have that $1 \preceq 2 \preceq \cdots \preceq n$, and $\text{supp}_P = \{(i, j); i, j \in [n], i \leq j\}$.

*Example 2.3* Let $n = 3$ and consider the anti-chain order and the chain order, determined by the incidence matrices

$$\begin{pmatrix} 1\,0\,0 \\ 0\,1\,0 \\ 0\,0\,1 \end{pmatrix} \text{ and } \begin{pmatrix} 1\,1\,1 \\ 0\,1\,1 \\ 0\,0\,1 \end{pmatrix}$$

respectively. The two sequences of incidence matrices

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

and

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

represent two different saturated chains in $\mathcal{P}_3$, intersecting only at the minimal and maximal elements.

### 2.1.1 Some Families of Finite Posets

In this section we introduce some relevant families of finite posets. In the first section of Chap. 2, we already considered two important families of posets over $[n]$, namely the anti-chain and the chain. We remark that the anti-chain order is unique, while the chain is not so: for any permutation $\sigma \in S_n$, the order $\preceq_\sigma$ defined by

$$\sigma(1) \preceq_\sigma \sigma(2) \preceq_\sigma \cdots \preceq_\sigma \sigma(n)$$

is a chain. It is worth to repeat that these orders are very particular since they represent a minimal and maximal orders on $[n]$.

We present now two families of posets over $[n]$, each of it includes the chains and the anti-chain as a particular case.

*Example 2.4 (Hierarchical Posets)* Consider a partition $[n] = \mathring{\bigcup}_{i=1,\ldots,l} H_i$, with $h_i := |H_i| > 0$, where $\mathring{\bigcup}$ denotes the union of disjoint sets. Define $\mathcal{H} = (H_1, \ldots, H_l)$ and $h = (h_1, \ldots, h_l)$ to be *hierarchy spectrum* and *hierarchy array*, respectively. We remark that $n = h_1 + h_2 + \ldots + h_l$. A *hierarchical poset* (also known as *weak order*) with hierarchy spectrum $\mathcal{H}$ is the poset $P_{\mathcal{H}} = ([n], \preceq_{\mathcal{H}})$, where

$$a \preceq_{\mathcal{H}} b \text{ iff } a \in H_i, b \in H_j \text{ and } i < j.$$

If we consider a natural labelling of the poset, we must have $H_1 = \{1, 2, \ldots, h_1\}$ and

$$H_j = \{(h_1 + \cdots + h_{j-1}) + 1, (h_1 + \cdots + h_{j-1}) + 2, \ldots, (h_1 + \cdots + h_{j-1}) + h_j\}$$

for every $1 < j \leq l$. Up to isomorphism, the hierarchy array completely determines the poset, and we say that $h$ is the *type* of $P_{\mathcal{H}}$ and denote this by writing $P_h = P_{\mathcal{H}}$ and $\preceq_h = \preceq_{\mathcal{H}}$. This notation is reflected in a simplified diagram, as one can see in Fig. 2.3.

We remark that, in case $l = 1$, we have only one level and the poset is an anti-chain. On the other hand, in case $l = n$, we have that each $h_i = 1$ and so, we have a hierarchical poset with hierarchy array $h = (1, 1, \ldots, 1)$, that is, a total order.

If $P_{\mathcal{H}}$ is a hierarchical poset with hierarchy array $h = (h_1, \ldots, h_l)$, the automorphism group of $P_{\mathcal{H}}$ is given by the direct sum

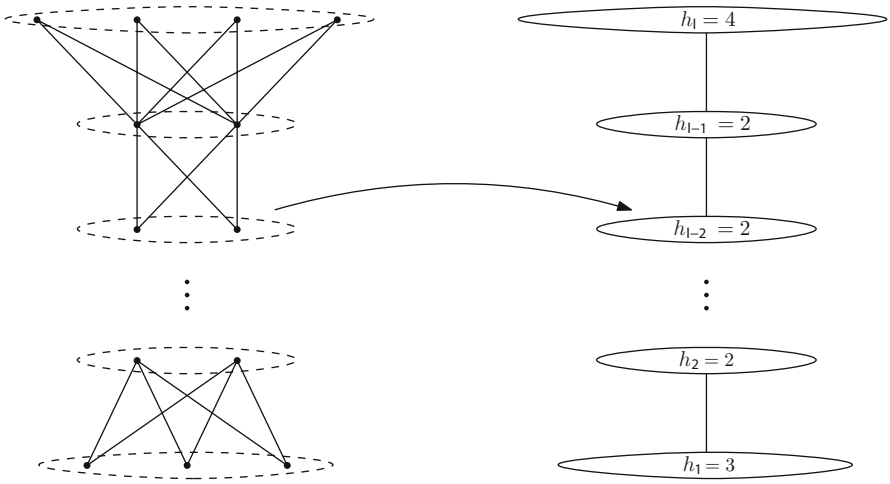


**Fig. 2.3** Hasse diagram and simplified representation for a hierarchical poset with hierarchy array $h = (3, 2, \ldots, 2, 2, 4)$

$$\mathrm{Aut}(P_{\mathcal{H}}) = S_{h_1} \oplus S_{h_2} \oplus \cdots \oplus S_{h_l},$$

where $S_{h_i}$ acts as a permutation of the elements of the level $H_i$.

*Example 2.5 (Disjoint Chains or Multi-Chain)* We consider, as in Example 2.4, a partition

$$[n] = \overset{\circ}{\bigcup_{i=1,\ldots,k}} R_i$$

and define $s_i := |R_i| > 0$. We call $\mathcal{R} = (R_1, \ldots, R_r)$ and $\mathcal{S} = (s_1, \ldots, s_r)$ the *chain spectrum* and *chain array*, respectively. We remark that $n = s_1 + s_2 + \cdots + s_r$. In the hierarchical case we imposed a chain relation between the parts and left each part containing non-comparable elements. Now, we do the opposite way: we impose a chain relation within each part and leave different parts unrelated. To be more explicit, we write

$$R_1 = \{x_{11}, x_{12}, \ldots, x_{1s_1}\}, \ldots, \ R_r = \{x_{r1}, x_{r2}, \ldots, x_{rs_r}\}$$

and define the relation $\preceq_{\mathcal{R}}$ by

$$x_{ij} \preceq_{\mathcal{R}} x_{i'j'} \text{ iff } i = i' \text{ and } j \leq j'.$$

We remark that, in case $r = 1$ we have only one chain, so a multi-chain poset with chain array $\mathcal{S} = (n)$ is a chain. On the other hand, in case $r = n$, we have that each $s_i = 1$ and so, a multi chain poset with chain array $\mathcal{S} = (1, 1, \ldots, 1)$, that is, an anti-chain.

In the special case when $n = r \cdot s$ and each $s_i = s$, we say $P_{\mathcal{R}}$ is a *Niederreiter-Rosenbloom-Tsafsmann poset*, or simply *NRT poset*, for historical reasons that will be exposed in the next section. In this case, we shall denote it by $\mathcal{R}(r, s) = ([r \cdot s], \preceq_{\mathcal{R}})$. Let us consider a multi-chain poset $P_{\mathcal{R}}$ with chain array $\mathcal{S} = (s_1, \ldots, s_r)$. We may assume, without loss of generality, that the chains are ordered in non-decreasing order, that is, $s_1 \leq s_2 \leq \cdots \leq s_r$. Some of those chains may have equal length, so there are $j_1, j_2, \ldots, j_l$ such that

$$s_1 = \cdots = s_{j_1} < s_{j_1+1} = \cdots = s_{j_1+j+2} < \cdots < s_{j_1+\cdots+j_{l-1}+1} = \cdots = s_{j_1+\cdots+j_l}.$$

Considering the Hasse diagram, it is not difficult to realize that the group of automorphisms of $P_{\mathcal{R}}$ is given by the direct product

$$\mathrm{Aut}(P_{\mathcal{R}}) = S_{j_1} \oplus S_{j_2} \oplus \cdots \oplus S_{j_l},$$

where each $S_{j_i}$ acts permuting the chains of length $s_{j_1+j_2+\cdots+j_i}$. In the case of an NRT poset with $r$ chains, all of length $s$, we find that $\mathrm{Aut}(P_{\mathcal{R}}) = S_r$. In the two

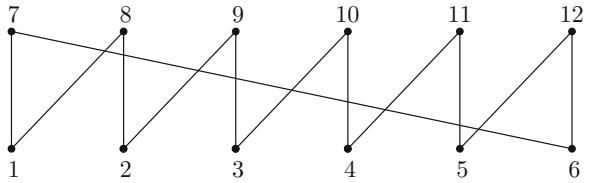


**Fig. 2.4** A crown poset with 12 elements ($n = 6$)

special cases, of a chain and an anti-chain, we get that $\mathrm{Aut}(\mathcal{R}(1,n)) = S_1$ and $\mathrm{Aut}(\mathcal{R}(n,1)) = S_n$, respectively.

*Example 2.6 (Crown Poset)* The crown poset is a very particular instance, that was explored, in the context of coding theory, in [59].

It is a poset $Cr = ([2n], \preceq_{Cr})$, where the only relations are

$$n \preceq_{Cr} 2n,\ n \preceq_{Cr} n+1 \text{ and } i \preceq_{Cr} n+i,\ i \preceq_{Cr} n+i+1, \forall 1 \leq i \leq n-1.$$

The classes of hierarchical and multi-chain posets are the most relevant ones in the context of coding theory. The crown poset is a convenient poset to produce examples and counterexamples (Fig. 2.4).

## 2.2 Metrics Defined by Posets

The starting point to realize how does a poset $P = ([n], \preceq)$ over $[n]$ determines a metric over $\mathbb{F}_q^n$ is to consider the Hamming case. The Hamming weight $\varpi_H$ of a vector $\mathbf{x} \in \mathbb{F}_q^n$ is determined by counting its support: $\varpi_H(\mathbf{x}) = |\mathrm{supp}(\mathbf{x})|$. If we consider the anti-chain order on $H := ([n], \preceq_H)$, we have that every subset $I \subseteq [n]$ is an ideal, since $i \in \left\langle \mathrm{supp}_H(j) \right\rangle_H$ means $i = j$. It follows that $\langle A \rangle_H = A$ for any $A \subset [n]$, so we have that

$$\varpi_H(\mathbf{x}) = |\, \langle \mathrm{supp}(\mathbf{x}) \rangle_H \,|$$

and for this reason we may, sometimes, refer to the anti-chain as the *Hamming order*. From this observation, the following definition arises naturally:

**Definition 2.7** Let $P = ([n], \preceq)$ be a poset over $[n]$. Given a vector $\mathbf{x} \in \mathbb{F}_q^n$ we define the *$P$-weight* of $\mathbf{x}$ by

$$\varpi_P(\mathbf{x}) := |\, \langle \mathrm{supp}(\mathbf{x}) \rangle_P \,|.$$

The first step we need is to prove that $\varpi_P$ is indeed a weight, in the sense that it satisfies the weight conditions in Sect. 1.1.

**Theorem 2.8** *Given a poset $P = ([n], \preceq)$, the $P$-weight $\varpi_P$ over $\mathbb{F}_q^n$ is a weight that respects the support of vectors.*

*Proof* See Exercise 11. ■

As a direct consequence of $\varpi_P$ being a weight over $\mathbb{F}_q^n$, we have that $d_p(\mathbf{x}, \mathbf{y}) := \varpi_P(\mathbf{x} - \mathbf{y})$ is a metric invariant by translations, called the *P-metric* or *P-distance*:

**Theorem 2.9** *Given a poset $P$ over $[n]$, $d_P$ is a metric over $\mathbb{F}_q^n$ which respects the support of vectors and is invariant by translations.*

*Proof* The fact that $d_P$ is a metric and that it respects the support of vectors follows straight from Theorem 2.8. The metric is invariant by translations since

$$\begin{aligned} d_P(\mathbf{x}, \mathbf{y}) &= \varpi_P(\mathbf{x} - \mathbf{y}) \\ &= \varpi_P((\mathbf{x} + \mathbf{z}) - (\mathbf{z} + \mathbf{y})) = d_P(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z}) \end{aligned}$$

for every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}_q^n$. ■

In the literature, the metric invariants of a code are defined considering the Hamming weight and metric. In this text, we have defined those invariants considering a general metric, so that they apply, in particular, to poset metrics. Just to recall, we summarize those definitions considering a poset metric.

**Definition 2.10** Given a code $C \subseteq \mathbb{F}_q^n$ and fixed a poset metric $d_P$, we define

- minimum distance: $\delta_{d_P}(C) := \min\{d_P(\mathbf{x}, \mathbf{y}); \mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}\}$;
- $P$-norm of a subspace $D \subseteq \mathbb{F}_q^n$: $\|D\|_P := |\langle \mathrm{supp}(D) \rangle_P|$;
- $i$-th generalized weight: $\delta_{i,d_P}(C) := \min\{\|D\|_P; D \subseteq C, \dim(D) = i\}$;
- metric ball centered in $\mathbf{x} \in \mathbb{F}_q^n$ *with radius* $r \in \mathbb{N}$:

$$B_{d_P}(\mathbf{x}, r) := \{\mathbf{y} \in \mathbb{F}_q^n; d_P(\mathbf{x}, \mathbf{y}) \leq r\};$$

- packing radius: $R_{d_P}(C) := \max\{r \in \mathbb{N}; B_{d_P}(\mathbf{x}, r) \cap B_{d_P}(\mathbf{y}, r) = \emptyset, \forall \mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}\}$;
- covering radius: $R_{d_P}^{cov}(C) := \min\{r \in \mathbb{N}; \mathbb{F}_q^n = \bigcup_{\mathbf{c} \in C} B_{d_P}(\mathbf{c}, r)\}$;
- weight enumerator: $W_{\varpi_P}(C, z) := \sum_{\mathbf{c} \in C} z^{\varpi_P(\mathbf{c})}$.

Along the text, in case the notation becomes too heavy, we may omit the symbol $P$ in the index of the notation.

The study of these metric invariants of a code considering a general poset metric or special families of poset metrics, is the main subject of research in the area and our goal is to present the main results and understandings on the subject. Before we proceed, we give explicit calculation of these invariants (the most relevant ones) for five different posets over [3].

*Example 2.11* We consider now five different posets over [3] and compute the main code invariants of each code. These are all the possible orders (up to isomorphism) on a set with three elements. The posets we consider are the following:

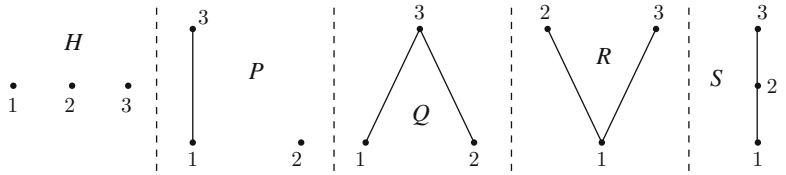


**Fig. 2.5** Hasse diagrams of posets $H$, $P$, $Q$, $R$ and $S$

- $H$: the Hamming poset;
- $P$: the poset determined by $1 \preceq_P 3$;
- $Q$: the poset determined by $1 \preceq_Q 3$ and $2 \preceq_Q 3$;
- $R$: the poset determined by $1 \preceq_R 3$ and $1 \preceq_R 2$;
- $S$: the poset determined by $1 \preceq_S 3$, $2 \preceq_S 3$ and $1 \preceq_S 2$.

The Hasse diagrams of these posets are represented in Fig. 2.5.

We consider two different codes, the 1-dimensional code $C_1 = \{000, 001\}$ and the 2-dimensional code $C_2 = \{000, 010, 001, 011\}$. The minimum distance, the packing radius and the metric ball with center at 000 and radius equals the packing radius of the code are presented in the following table:

| | $H$ | $P$ | $Q$ | $R$ | $S$ |
|---|---|---|---|---|---|
| $\delta_{d_*}(C_1)$ | 1 | 2 | 3 | 2 | 3 |
| $\delta_{d_*}(C_2)$ | 1 | 1 | 1 | 2 | 2 |
| $R_{d_*}(C_1)$ | 0 | 1 | 2 | 1 | 2 |
| $R_{d_*}(C_2)$ | 0 | 0 | 0 | 1 | 1 |
| $B_{d_*}\big(000, R_{d_*}(C_1)\big)$ | 000 | 000, 100, 010 | 000, 100, 010, 110 | 000, 100 | 000, 100, 010, 110 |
| $B_{d_*}\big(000, R_{d_*}(C_2)\big)$ | 000 | 000 | 000 | 000, 100 | 000, 100 |

It is easy to check (see Exercise 12 for a hint) that these codes may or may not be perfect, depending on the metric taken into consideration. Moreover, we clearly see that, depending on the code and the metric, the packing radius can attain all the values between $\lfloor(\delta_{d_*}(C) - 1)/2\rfloor$ and $\delta_{d_*}(C) - 1$. This is the case, for example, for $C_1$: the packing radius assume the values 0, 1 and 2, depending on the metric.

It is also worth to remark that, depending on the poset, the packing radius of a code does not necessarily determine its minimum distance. If we consider $C_3 = \{000, 110\}$, it is easy to check that $\delta_{d_P}(C_1) = 2 = \delta_{d_P}(C_3)$ but $R_{d_P}(C_1) = 1$ and $R_{d_P}(C_3) = 0$.

All these aspects will be explored in details in the next chapters and we will find, for example, necessary and sufficient conditions on a poset $P$ to ensure that $R_{d_P}(C) = \lfloor(\delta_{d_P}(C) - 1)/2\rfloor$ or $R_{d_P}(C) = \delta_{d_P}(C) - 1$ for every code $C$, or conditions to determine that $R_{d_P}(C)$ is completely determined by $\delta_{d_P}(C)$ (details in Chap. 3).

In the next examples we show how to explicitly compute the weight of a vector when considering our two main examples of posets.

*Example 2.12 (Hierarchical Poset Metric)* We consider the hierarchical poset $P_{\mathcal{H}} = ([n], \preceq_{\mathcal{H}})$ with hierarchy spectrum $\mathcal{H} = (H_1, H_2, \ldots, H_l)$, as in Example 2.4. Given $\mathbf{x} \in \mathbb{F}_q^n$, we write $\mathbf{x} = \mathbf{x}^1 + \mathbf{x}^2 + \cdots + \mathbf{x}^l$ where $\text{supp}(\mathbf{x}^i) \subseteq H_i$. Then, if $M(\mathbf{x}) = \max\left\{i; \mathbf{x}^i \neq \mathbf{0}\right\}$ we have that

$$\langle \text{supp}(\mathbf{x}) \rangle = \left(\text{supp}(\mathbf{x}) \cap H_{M(\mathbf{x})}\right) \overset{\circ}{\bigcup} \left( \bigcup_{i=1}^{M(\mathbf{x})-1} H_i \right)$$

and the disjoint union ensures that

$$\varpi_{\mathcal{H}}(\mathbf{x}) = \left|\text{supp}\left(\mathbf{x}^{M(\mathbf{x})}\right)\right| + \sum_{i=1}^{M(\mathbf{x})-1} h_i$$

where $(h_1, h_2, \ldots, h_l)$ is the hierarchy array of $\mathcal{H}$. In other words, we look for the higher level that the non-zero coordinates of $\mathbf{x}$ intersects and then we count how many non-zero coordinates we have in that level and sum all the positions in the lower levels.

*Example 2.13 (NRT Poset Metric)* Let us consider the NRT poset consisting of $r$ disjoint chains, each having length equal to $s$, as exposed in Example 2.5. We write $n = rs$ and express

$$\mathbf{x} = \left(x_1^1, x_2^1, \ldots, x_s^1; x_1^2, x_2^2, \ldots, x_s^2; \ldots; x_1^r, x_2^r, \ldots, x_s^r\right) \in \mathbb{F}_q^n$$

where $x_1^i, x_2^i, \ldots, x_s^i$ are the coordinates corresponding to the $i$-th chain. For short, we may write $\mathbf{x} = \left(\mathbf{x}^{(1)}; \mathbf{x}^{(2)}; \ldots; \mathbf{x}^{(r)}\right)$ and $\mathbf{x}^{(i)} = \left(x_1^i, x_2^i, \ldots, x_s^i\right)$. We recall that $x_j^i \preceq_{\mathcal{R}} x_{j'}^{i'}$ if, and only if, $i = i'$ and $j \leq j'$. We define $M\left(\mathbf{x}^{(i)}\right) := \max\left\{j; x_j^i \neq 0\right\}$ and we have that $\varpi_{\mathcal{R}}(\mathbf{x}) = \sum_{i=1}^{r} M\left(\mathbf{x}^{(i)}\right)$. In other words, we look for the maximal non-zero coordinates in each chain and sum these maximal levels.

## 2.3 Linear Isometries and Characterization of Orbits

An *isometry* $T$ of the metric space $\left(\mathbb{F}_q^n, d_P\right)$ (or a $P$-isometry) is a map $T : \mathbb{F}_q^n \to \mathbb{F}_q^n$ such that $d_P(T(\mathbf{x}), T(\mathbf{y})) = d_P(\mathbf{x}, \mathbf{y})$, for every $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$. Since we are considering mainly linear codes, it is interesting to consider isometries that also preserve linearity: a *linear isometry* $T$ of the metric space $\left(\mathbb{F}_q^n, d_P\right)$ is a linear transformation $T : \mathbb{F}_q^n \to \mathbb{F}_q^n$ that is also a $P$-isometry. It is not difficult to see that a

linear transformation $T$ is a $P$-isometry if, and only if, it preserves the weight, that is, $\varpi_P(T(\mathbf{x})) = \varpi_P(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{F}_q^n$. We say that two codes $C$ and $C'$ are *$P$-equivalent* (or just equivalent, where no confusion may arise) if there is a linear $P$-isometry $T$ such that $T(C) = C'$.

The *group of linear isometries* of the poset space $\left(\mathbb{F}_q^n, d_P\right)$ is denoted by $GL_P\left(\mathbb{F}_q^n\right)$. Knowing and understanding the structure of this group is crucial in our setting, since we are concerned with metric invariants. For instance, if $T$ is a $P$-isometry and $C \subseteq \mathbb{F}_q^n$ is a code, then the code $C$ and its image $T(C)$ have the same metric invariants: $\delta_{d_P}(C) = \delta_{d_P}(T(C))$, $R_{d_P}(C) = R_{d_P}(T(C))$ and so on, since the isometry $T$ is defined on $\mathbb{F}_q^n$ and not only on $C$.

Understanding the structure of $GL_P\left(\mathbb{F}_q^n\right)$ is one of the first subjects to be approached by researchers studying codes with poset metrics. The first work to deal with this question in 2003 [68] considered the special case of NRT posets consisting of $k$ disjoint chains, each having length equal to $m$. The key point to describe $GL_P\left(\mathbb{F}_q^n\right)$ is to understand two different types of isometries, already identified in [68]. We shall describe and explain these two types of isometries with some details. We shall start with some definitions and notations.

We let $\mathbf{e}_i \in \mathbb{F}_q^n$ be the vector such that $\operatorname{supp}(\mathbf{e}_i) = \{i\}$ and its $i$-th coordinate is equal to 1. We denote by $\beta = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ the so-called canonical basis of $\mathbb{F}_q^n$. Given a vector $\mathbf{x} \in \mathbb{F}_q^n$, we denote by $M(\mathbf{x}) := M_P(\mathbf{x})$ the set of *maximal elements* of $\operatorname{supp}(x)$ (according to $P$) and by

$$\widehat{\mathbf{x}} = (\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_n) = \sum_{i \in M(X)} x_i \mathbf{e}_i$$

the *cleared out form of* $\mathbf{x}$. It is clear that $\varpi(\widehat{\mathbf{x}}) = \varpi(\mathbf{x})$. We now introduce two distinguished types of isometries:

**(i) Operating on Ideals:** We map each *principal ideal* (an ideal $\langle i \rangle_P$ generated by a single element) onto itself, in the sense that, $\langle \operatorname{supp}(\mathbf{e}_i) \rangle = \langle \operatorname{supp}(T(\mathbf{e}_i)) \rangle$ for every $\mathbf{e}_i \in \beta$. By saying so, we mean that $T$ is the linear map determined by

$$T(\mathbf{e}_i) = \sum_{j \in \langle i \rangle} \alpha_{ij} \mathbf{e}_j$$

with $\alpha_{ii} \neq 0$. It is simple to verify that $T$ is an isometry. Indeed, let $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$. Since we are assuming $T$ to be linear, we have that

$$T(\mathbf{x}) = T\left(\sum_{i=1}^n x_i \mathbf{e}_i\right) = \sum_{i=1}^n x_i T(\mathbf{e}_i) = \sum_{i=1}^n x_i \left(\sum_{j \in \langle i \rangle}^n \alpha_{ij} \mathbf{e}_j\right).$$

If $i \in M(\mathbf{x})$, we have that $\mathbf{e}_i$ appears only once in the last summand, so the cleared out form of $T(\mathbf{x})$ is $\widehat{T(\mathbf{x})} = \sum_{i \in M(\mathbf{x})} x_i \alpha_{ii} \mathbf{e}_i$. Since $\alpha_{ii} \neq 0$, we have that $\operatorname{supp}(\widehat{T(\mathbf{x})}) = \operatorname{supp}(\widehat{\mathbf{x}})$ and

$$\varpi(\mathbf{x}) = \varpi(\widehat{\mathbf{x}}) = \varpi(\widehat{T(\mathbf{x})}) = \varpi(T(\mathbf{x})).$$

Considering the basis $\beta$, each such isometry is determined by a matrix $A = (a_{ij})_{i,j \in [n]}$ where $a_{ii} \neq 0$ and $a_{ij} = 0$ if $j \preceq_P i$ and $j \neq i$. We denote by $G_P$ the set of all such matrices (or linear isometries):

$$G_P = \left\{ T_A;\ A = \left(a_{ij}\right) \in M_{n \times n}\left(\mathbb{F}_q\right) \text{ and } a_{ij} \in \begin{vmatrix} \mathbb{F}_q & \text{if } i \preceq_P j, i \neq j \\ \mathbb{F}_q^* & \text{if } \quad i = j \\ 0 & \text{otherwise} \end{vmatrix} \right\}.$$

It is lengthy but straightforward to prove that $G_P$ is a group.

**(ii) Induced by Poset Automorphism:** Given $\sigma \in Aut(P)$, we denote by $T_\sigma$ its action on $\mathbb{F}_q^n$: $T_\sigma(x_1, x_2, \ldots, x_n) = (x_{\sigma(1)}, x_{\sigma(n)}, \ldots, x_{\sigma(n)})$. It follows that $\operatorname{supp}(T_\sigma(\mathbf{x})) = \sigma(\operatorname{supp}(\mathbf{x}))$ and since $\sigma \in \operatorname{Aut}(P)$ we get that $\varpi(\mathbf{x}) = \varpi(T_\sigma(\mathbf{x}))$ and $T_\sigma$ is also an isometry.

Since both $G_P$ and $\operatorname{Aut}(P)$ are groups of linear isometries, it follows that the product $G_P \cdot \operatorname{Aut}(P)$ is a set of linear isometries. As we shall see in Theorem 2.19, we actually have that $GL_P\left(\mathbb{F}_q\right)$ is the semi-direct product $GL_P\left(\mathbb{F}_q\right) = G_P \rtimes \operatorname{Aut}(P)$. The key point of the proof is to show that any $T \in GL_P(\mathbb{F}_q^n)$ induces an automorphism of poset $\sigma_T \in \operatorname{Aut}(P)$. To construct $\sigma_T$ we need the following:

**Lemma 2.14** *Let $T \in GL_P\left(\mathbb{F}_q^n\right)$ and $i \in \{1, 2, \ldots, n\}$. If $j \in \operatorname{supp}(T(\mathbf{e}_i))$, then $|\langle j \rangle| \leq |\langle i \rangle|$.*

*Proof* By assumption $\langle j \rangle \subseteq \langle \operatorname{supp}(T(\mathbf{e}_i)) \rangle$ so that $|\langle j \rangle| \leq |\langle \operatorname{supp}(T(\mathbf{e}_i)) \rangle|$. Since $T$ preserves the $P$-weight, it follows that $|\langle \operatorname{supp}(T(\mathbf{e}_i)) \rangle| = |\langle \operatorname{supp}(\mathbf{e}_i) \rangle|$ and, by definition, $|\langle \operatorname{supp}(\mathbf{e}_i) \rangle| = |\langle i \rangle|$. ■

**Proposition 2.15** *Let $T \in GL_P\left(\mathbb{F}_q^n\right)$. Then, for every $i \in \{1, 2, \ldots, n\}$, we have that $\langle \operatorname{supp}(T(\mathbf{e}_i)) \rangle$ is a principal ideal.*

*Proof* We will first show that there is an element $j \in \operatorname{supp}(T(\mathbf{e}_i))$ satisfying $|\langle j \rangle| = |\langle i \rangle|$. Let $\operatorname{supp}(T(\mathbf{e}_i)) = \{i_1, i_2, \ldots, i_s\}$. Assume the contrary, namely that $|\langle i_u \rangle| < |\langle i \rangle|$ for every $u \in \{1, 2, \ldots, s\}$. It follows from the linearity of $T^{-1}$ that we have $\{i\} = \operatorname{supp}(\mathbf{e}_i) \subseteq \cup_{k=1}^{s} \operatorname{supp}\left(T^{-1}\left(\mathbf{e}_{i_k}\right)\right)$, which implies $i \in \operatorname{supp}\left(T^{-1}\left(\mathbf{e}_{i_k}\right)\right)$ for some $k \in \{1, 2, \ldots, s\}$. Using this and Lemma 2.14, we obtain $|\langle i \rangle| \leq |\langle i_k \rangle| < |\langle i \rangle|$. This is a contradiction, so there is $j \in \operatorname{supp}(T(\mathbf{e}_i))$ satisfying $|\langle j \rangle| = |\langle i \rangle|$. Since $T$ is assumed to preserve the weight, this $j$ is unique and we have that $\langle \operatorname{supp}(T(\mathbf{e}_i)) \rangle = \langle j \rangle$ ■

Since each $T \in GL_P\left(\mathbb{F}_q^n\right)$ maps a principal ideal into a principal ideal, given $i \in [n]$ there is a unique $j \in [n]$ such that $\langle \text{supp}(T(\mathbf{e}_i))\rangle = \left\langle \text{supp}(\mathbf{e}_j)\right\rangle$ and we define the map $\sigma_T : [n] \to [n]$ by setting $\sigma_T(i) = j$.

**Theorem 2.16** *Given $T \in GL_P(\mathbb{F}_q^n)$, the map $\sigma_T$ is a well defined automorphism of $P$.*

*Proof* For the (many) details of the proof, we refer the reader to [87, Theorem 1.1]. ■

Let $\Gamma^{(m)}(P)$ be the set of elements of $P$ that generates prime ideals with cardinality $m$:

$$\Gamma^{(m)}(P) = \{i \in P; |\langle i\rangle| = m\} = \{i \in P; w_P(\mathbf{e}_i) = m\}.$$

We now describe the main result of this section:

**Theorem 2.17** *Let $P = \{1, 2, \ldots, n\}$ be a poset and $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be the canonical base of $\mathbb{F}_q^n$. Then $T \in GL_P(\mathbb{F}_q)$ if and only if*

$$T(\mathbf{e}_j) = \sum_{i \in \langle j\rangle} x_{ij}\mathbf{e}_{\phi(i)}$$

*where $\phi : P \to P$ is an automorphism of $P$ and $x_{jj} \neq 0$, for any $j \in \{1, 2, \ldots, n\}$. Moreover, there is a pair of ordered bases $\beta$ and $\beta'$ of $\mathbb{F}_q^n$ relative to which every linear isometry $T \in GL_P\left(\mathbb{F}_q\right)$ is represented by an $n \times n$ upper triangular matrix $\left(a_{ij}\right)_{1 \leq i, j \leq n}$ with $a_{ii} \neq 0$ for every $i \in \{1, 2, \ldots, n\}$.*

*Proof* Suppose that $T \in GL_P\left(\mathbb{F}_q^n\right)$. Lemma 2.15 ensures that $\left\langle \text{supp}\left(T\left(\mathbf{e}_j\right)\right)\right\rangle$ is a prime ideal for every $j \in \{1, 2, \ldots, n\}$. Given $j \in \{1, 2, \ldots, n\}$, let $j' = \phi(j)$ be the unique maximal element of $\left\langle \text{supp}\left(T\left(\mathbf{e}_j\right)\right)\right\rangle$, where $\phi := \phi_T : P \to P$ is the automorphism induced by the isometry $T$ (see Theorem 2.16). Then

$$\left\langle \text{supp}\left(T\left(\mathbf{e}_j\right)\right)\right\rangle = \left\langle \text{supp}\left(\mathbf{e}_{j'}\right)\right\rangle = \left\langle \text{supp}\left(\mathbf{e}_{\phi(j)}\right)\right\rangle,$$

and since $\phi$ is an automorphism,

$$\left\langle \text{supp}\left(\mathbf{e}_{\phi(j)}\right)\right\rangle = \{\phi(i); i \in \langle j\rangle\}.$$

Therefore, $\left\langle \text{supp}\left(T\left(\mathbf{e}_j\right)\right)\right\rangle = \{\phi(i); i \in \langle j\rangle\}$. Since $\phi(j) = \max\{\phi(i); i \in \langle j\rangle\}$, we conclude that

$$T(\mathbf{e}_j) = \sum_{i \in \langle j\rangle} x_{ij}\mathbf{e}_{\phi(i)} \tag{2.1}$$

with $x_{jj} \neq 0$. It is straightforward to verify that for a given order automorphism $\phi : P \to P$, any linear map defined as in (2.1) is a $P$-isometry.

For the second part of the Theorem, we shall explicitely build two basis $\beta$ and $\beta'$. Let $\beta_m = \left\{\mathbf{e}_i ; i \in \Gamma^{(m)}(P)\right\}$ and

$$\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$$

be a decomposition of the canonical base of $\mathbb{F}_q^n$ as a disjoint union, where $k = \max\{w_P(\mathbf{e}_i) ; i = 1, 2, \ldots, n\}$. We order this basis $\beta = \left\{\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_n}\right\}$ in the following way (and denote this partial order by $\leq_\beta$): if $\mathbf{e}_{i_r} \in \beta_{j_r}$ and $\mathbf{e}_{i_s} \in \beta_{j_s}$ with $r \neq s$ then, $\mathbf{e}_{i_r} \leq_\beta \mathbf{e}_{i_s}$ if and only $j_r \leq j_s$. In other words, we begin enumerating the vectors of $\beta_1$ and after exhausting them, we enumerate the vectors of $\beta_2$ and so on.

We define another ordered basis $\beta'$ as the basis induced by the automorphism $\phi$, that is, $\beta' := \left\{\mathbf{e}_{\phi(i_1)}, \mathbf{e}_{\phi(i_2)}, \ldots, \mathbf{e}_{\phi(i_n)}\right\}$, and let $A$ be the matrix of $T$ relative to the basis $\beta$ and $\beta'$:

$$[T]_{\beta,\beta'} = A = (a_{kl})_{1 \leq k,l \leq n}.$$

We find by the construction of the basis $\beta$ and $\beta'$ that $a_{kl} \neq 0$ implies $i_l \in \langle \phi(i_k) \rangle$. But $i_l \in \langle \phi(i_k) \rangle$ and $\langle i_l \rangle \neq \langle \phi(i_k) \rangle$ implies that $l < k$ so, $A$ is upper triangular. Since $A$ is invertible and upper triangular, we must have $\det(A) = \prod_{i=1}^n a_{ii} \neq 0$ hence, $a_{ii} \neq 0$ for every $i \in \{1, 2, \ldots, n\}$. ■

The upper triangular matrix obtained in the previous theorem is called a *canonical form of* $T$ and they are canonical in the sense they are unique, up to re-ordination within the linearly independent sets $\beta_i, i = 1, 2, \ldots, k$.

A *monomial matrix* is a matrix with exactly one nonzero entry in each row and column. Thus a monomial matrix over $\mathbb{F}_2$ is a permutation matrix, and a monomial matrix over an arbitrary finite field is a permutation matrix times an invertible diagonal matrix.

**Corollary 2.18** *Given $T \in GL_P\left(\mathbb{F}_q^n\right)$, there is an ordering $\beta = \left\{\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_n}\right\}$ of the canonical base such that $[T]_{\beta,\beta}$ is given by the product $A \cdot U$ where $A$ is an invertible upper triangular matrix and $U$ is a monomial matrix obtained from the identity matrix by permutation of the columns, corresponding to the automorphism induced by $T$.*

*Proof* Let $\phi_T$ be the automorphism of order induced by $T$. Let $T_{\phi_T^{-1}}$ be the linear isometry defined as $T_{\phi_T^{-1}}(\mathbf{e}_j) = \mathbf{e}_{\phi_T^{-1}(j)}$, for $j \in \{1, 2, \ldots, n\}$. As we saw in Theorem 2.17, $T(\mathbf{e}_j) = \sum_{i \in \langle j \rangle} x_{ij} \mathbf{e}_{\phi_T(i)}$, so that

$$\begin{aligned} T \circ T_{\phi_T^{-1}}\left(\mathbf{e}_j\right) &= T\left(\mathbf{e}_{\phi_T^{-1}(j)}\right) \\ &= \sum_{i \in \left\langle \phi_T^{-1}(j) \right\rangle} x_{i\phi_T^{-1}(j)} \mathbf{e}_{\phi_T(i)} \\ &= x_{i\phi_T^{-1}(j)} \mathbf{e}_j + \sum_{i \in \left\langle \phi_T^{-1}(j) \right\rangle, i \neq \phi_T^{-1}(j)} x_{i\phi_T^{-1}(j)} \mathbf{e}_{\phi_T(i)}. \end{aligned}$$

It follows that the automorphism induced by $T \circ T_{\phi_T^{-1}}$ is the identity, so, when taking the base $\beta'$ as in Theorem 2.17, we find that $\beta' = \beta$ and the matrix of $T \circ T_{\phi_T^{-1}}$ relative to this base is an upper triangular matrix $A = \left[T \circ T_{\phi_T^{-1}}\right]_\beta$. But $T_{\phi_T^{-1}}$ acts on $\mathbb{F}_q^n$ as a permutation of the vectors in $\beta$, so that in any ordered base containing those vectors, $U^{-1} = \left[T_{\phi_T^{-1}}\right]$ is obtained from the identity matrix by permutation of the columns. We note that $T_{\phi_T} = \left(T_{\phi_T^{-1}}\right)^{-1}$ and it follows that

$$[T]_\beta = \left[T \circ T_{\phi_T^{-1}} \circ T_{\phi_T}\right]_\beta = \left[T \circ T_{\phi_T^{-1}}\right]_\beta \left[T_{\phi_T}\right]_\beta = A \cdot U.$$

■

We already knew that $G_P \cdot \mathrm{Aut}(P) \subseteq GL_P(\mathbb{F}_q^n)$ and the previous corollary may be restated in a simpler way: $G_P \cdot \mathrm{Aut}(P) = GL_P(\mathbb{F}_q^n)$. We shall prove that this is actually a semi-direct product of groups.

**Theorem 2.19** *With the definitions above, the group of isometries of* $\left(\mathbb{F}_q^n, d_P\right)$ *is the semi-direct product* $GL_P\left(\mathbb{F}_q^n\right) \simeq G_P \rtimes \mathrm{Aut}(P)$.

*Proof* First of all we should prove that $G_P$ is a group. Indeed, let $A = \left(a_{ij}\right)$ and $B = \left(b_{ij}\right)$ be elements in $G_P$. Since

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{i \leq_P k \leq_P j} a_{ik} b_{kj},$$

$AB \in G_P$. We note that every element in $G_P$ is an upper triangular matrix with nonzero diagonal entries. Hence, such elements are invertible. Since the inverse of an element in $G_P$ is a polynomial in that element, such an element is in $G_P$. So, we see that $G_P$ is a subgroup of $GL_P\left(\mathbb{F}_q\right)$. Since we already proved that $GL_P\left(\mathbb{F}_q\right) = G_P \cdot \mathrm{Aut}(P)$, all that is left to show is that $G_P$ is a normal subgroup of $GL_P\left(\mathbb{F}_q\right)$. Given $\phi \in S_n$, it acts on $n \times n$ matrices by permuting columns or rows. We denote by $A^\phi$ and ${}^\phi A$ respectively the column and row permutation of the matrix $A$. It is straightforward to show that $\left({}^\phi Id\right)^{-1} = Id^\phi$[17]. It follows that

$$\left({}^{\phi}Id\right) A \left({}^{\phi}Id\right)^{-1} = {}^{\phi}A^{\phi}$$

for every $n \times n$ matrix $A$. If $A = \left(a_{ij}\right) \in G_P$, for each $i = 1, 2, \ldots, n$,

$$\left({}^{\phi}Id\right) A \left({}^{\phi}Id\right)^{-1} (\mathbf{e}_i) = {}^{\phi} A^{\phi} (\mathbf{e}_i) = \sum_{k=1}^{n} a_{\phi(k)\phi(i)} \mathbf{e}_k = \sum_{\phi(k) \leq_P \phi(i)} a_{\phi(k)\phi(i)} \mathbf{e}_k$$
$$= \sum_{k \leq_P i} a_{\phi(k)\phi(i)} \mathbf{e}_k$$

and $a_{\phi(i)\phi(i)} \neq 0$ for every $i$. Thus, we find that $G_P$ is normal in $GL_P\left(\mathbb{F}_q\right)$ and the theorem follows. ■

### 2.3.1 Examples

In this section, we illustrate the results of this chapter with our main classes of poset-metrics: the hierarchical posets and the disjoint union of chains.

*Example 2.20* Consider the hierarchical poset $\mathcal{H} = (H_1, H_2, \ldots, H_l)$ and let $\beta = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, with $\text{supp}(\mathbf{e}_i) \in H_j$ for $h_1 + \cdots + h_{j-1} < i \leq h_1 + \cdots + h_{j-1} + h_j$. As seen in Example 2.4, $\text{Aut}(\mathcal{H})$ is isomorphic to $S_{h_1} \times S_{h_2} \times \cdots \times S_{h_l}$, with $\sigma = (\sigma_1, \ldots, \sigma_l) \in \text{Aut}(\mathcal{R})$ acting on the levels: for

$$\mathbf{x} = \sum_{i=1}^{h_j} \sum_{j=1}^{l} x_{i+L_i} \mathbf{e}_{i+L_i}$$

with $L_1 = 0$, $L_i = h_1 + \cdots + h_{i-1}$ (for $i > 1$), then

$$\sigma(\mathbf{x}) = \sum_{i=1}^{h_j} \sum_{j=1}^{l} x_{\sigma_j(i)+L_i} \mathbf{e}_{i+L_i}.$$

Considering the basis $\beta$, we have that $G_{\mathcal{H}}$ is represented as the group of the matrices of the form

$$A = \begin{pmatrix} A_1 & B_{12} & B_{13} & \ldots & B_{1l} \\ 0 & A_2 & B_{23} & \ldots & B_{2l} \\ 0 & 0 & A_3 & \ldots & B_{3l} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & A_k \end{pmatrix},$$

where each $A_i$ is an invertible $h_i \times h_i$ diagonal matrix and $B_{ij}$ is any $h_i \times h_j$ matrix.

We remark that

$$\left|GL_{\mathcal{H}}\left(\mathbb{F}_q^n\right)\right| = \prod_{i=1}^{l}\left(h_i!\,(q-1)^{h_i}\,q^{h_i L_i}\right)$$

$$= (q-1)^n\, q^{\sum_{i=1}^{l} h_i L_i} \prod_{i=1}^{l} h_i!.$$

*Example 2.21* Let $\mathcal{R} = (R_1, R_2, \ldots, R_r)$ be the NRT poset consisting of a disjoint union of $r$ chains, all of lenght $s$ and let $n = rs$. Considering an ordering

$$\mathbf{e}_1, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_{r(s-1)+1}, \mathbf{e}_2, \mathbf{e}_{r+2}, \ldots, \mathbf{e}_{r(s-1)+2}, \ldots, \mathbf{e}_r, \mathbf{e}_{r+r}, \ldots, \mathbf{e}_{r(s-1)+r}$$

of the usual basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of $\mathbb{F}_q^n$ ($\mathrm{supp}\,(\mathbf{e}_i) = \{i\}$), we have that the support of the first $s$ vectors is contained in the first chain, the support of the following $s$ vectors is contained in the second chain, and so on, so that

$$\mathrm{supp}\left(\left\{\mathbf{e}_i, \mathbf{e}_{r+i}, \ldots, \mathbf{e}_{r(s-1)+i}\right\}\right) = R_i.$$

Considering this ordered basis, every linear isometry $T \in GL_P\left(\mathbb{F}_q\right)$ is represented by the product $A \cdot U$ of $n \times n$ matrices, where $U$ is a monomial matrix obtained by exchanging the $k$ diferent blocks of $m$ collumns of

$$Id = \left(\begin{array}{c|c|c|c} Id_{s\times s} & 0_{s\times s} & \ldots & 0_{s\times s} \\ \hline 0_{s\times s} & Id_{s\times s} & \ldots & 0_{s\times s} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0_{s\times s} & 0_{s\times s} & \ldots & Id_{s\times s} \end{array}\right)$$

and

$$A = \begin{pmatrix} A_1 & 0 & 0 & \ldots & 0 \\ 0 & A_2 & 0 & \ldots & 0 \\ 0 & 0 & A_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & A_r \end{pmatrix},$$

where each $A_i$ is an $s \times s$ upper triangular matrix with non zero diagonal entries. In other words, Aut $(\mathcal{R})$is isomorphic to $S_r$, acting by permuting blocks of coordinates, so that $\sigma\left(\mathbf{e}_{jr+i}\right) = \mathbf{e}_{jr+\sigma(i)}$, where $i = 0, 1, \ldots, r-1$ and $G_{\mathcal{R}}$ is isomorphic to $N^r$, where $N$ is the group of invertible upper triangular $n \times n$ matrices.

If there is a unique chain of length $s = n$, we find that $U$ is necessarily the identity matrix and $A$ may be any $n \times n$ upper triangular matrix. In the Hamming case ($s = 1$ and $r = n$) we obtain $U \in S_n$ and $A$ is an invertible diagonal matrix.

We remark that $\left|GL_{\mathcal{R}}\left(\mathbb{F}_q^n\right)\right| = r!\,(q-1)^s\, q^{\frac{s(s-1)}{2}}$.

## 2.4 Chapter Notes

The description of the group of linear isometries of a poset space as a semi-direct product, valid for a general poset, was preceded by the description of some special cases: NRT [68], crown poset metric [17] and hierarchical poset [58]. The clear understanding of this structure is a key point for the proof of many results, including the canonical decomposition of a hierarchical poset code (to be presented in the next chapter) and it was studied, also as a technical tool, for all the generalizations of a poset metric, to be presented in Chap. 7. There is still work to be done in this direction, including the description of the group of all isometries, not necessarily linear. On this specific case, there are some initial results concerning NRT metric [88], and the construction developed in [88] was used later in [49] to obtain a subgroup of the isometry group and also for classifying the poset metrics for which this subgroup is the full group.

Poset metric is a variation of the Hamming metric concerned exclusively with the relations between the coordinates, and not within the alphabets. The Lee metric is probably the simplest (and most important) case to consider an additional structure on the alphabet. A simultaneous generalization of poset metrics and the Lee metric was developed recently using partially ordered *multisets* and it will be briefly presented in Chap. 7. There is a whole universe of research when considering different algebraic structures on the alphabet (generally a ring $\mathcal{A}$) and when considering a metric on the alphabet adapted to this structure. Considering such structures, the metric on a ring is extended to an $\mathcal{A}$-module additively. A very interesting introduction on the subject can be found in [111]. The main concern is pretty similar to our case (of poset-metrics), but the taste is more algebraic than combinatorial.

## 2.5 Exercises

1. Let $P = (X, \preceq_P)$ and $Q = \left(Y, \preceq_Q\right)$ be posets. Let $f : X \to Y$ be a bijective poset homomorphism. Prove that the inverse $f^{-1}$ is a poset homomorphism if, and only if, $x \preceq_P x' \iff f(x) \preceq_Q f(x')$.
2. Let $P = (X, \preceq_P)$ and $Q = \left(Y, \preceq_Q\right)$ be posets. Suppose there are injective poset homomorphisms $f : X \to Y$ and $g : Y \to X$. Can we conclude that $P$ and $Q$ are isomorphic?
3. Given a poset $P = (X, \preceq_P)$, prove that $\mathrm{Aut}(P)$ is a group.

4. Let $P = ([n], \preceq_P)$ be a finite poset. Prove that $\mathrm{Aut}(P)$ is (group) isomorphic to the permutation group $S_n$ *iff* $P$ is an antichain.
5. Let $S_n$ be the permutation group of $[n]$ and let $\sigma \in S_n$ be a cycle of order $n$. Let $G = \langle \sigma \rangle \subseteq S_n$ be the subgroup generated by $\sigma$. Prove that there is no poset $P = (X, \preceq_P)$ such that $\mathrm{Aut}(P)$ is group-isomorphic to $G$.
6. Show that, up to isomorphism, there are exactly five different posets $P$ over $[3]$. Show that $\mathrm{Aut}(P)$ is isomorphic to either $S_3$, $S_2$ or $S_1$.
7. A rank function on a poset $P$ is a function $\rho : X \to \mathbb{N}$ that is compatible with the order, in the sense that $x \preceq y$ implies $\rho(x) < \rho(y)$ and if $y$ covers $x$, then $\rho(y) = \rho(x) + 1$. Find sufficient and necessary conditions to ensure that the height function is a rank.
8. Considering the order $\leq_n$ over the set $\mathcal{P}_n$ of all posets over $[n]$, prove that an anti-chain is the unique minimal poset and the chains are all the $n!$ maximal posets.
9. Prove that two multi-chain posets with chain arrays $\mathcal{S} = (s_1, \ldots, s_l)$ and $\mathcal{S}' = (s'_1, \ldots, s'_{l'})$ are isomorphic *iff* $l = l'$ and there is a permutation $\sigma \in S_l$ such that $s_{\sigma(i)} = s'_i$ for all $i \in [l]$.
10. For each of the incidences matrices in Example 2.3:

    (a) Draw a Hasse diagram of the poset;
    (b) Determine each of those posets are hierarchical or multi-chain.

11. Prove Theorem 2.8: Given a poset $P$ over $[n]$, the $P$-weight $\varpi_P$ over $\mathbb{F}_q^n$ is a weight that respects the support of vectors, that is, for every $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$, we have that:

    (a) $\varpi_P(\mathbf{x}) \geq 0$;
    (b) $\varpi_P(\mathbf{x}) = 0$ *iff* $\mathbf{x} = \mathbf{0}$;
    (c) $\varpi_P(\mathbf{x}) = \varpi_P(\lambda \mathbf{x})$ for all $0 \neq \lambda \in \mathbb{F}_q$;
    (d) $\varpi_P(\mathbf{x} + \mathbf{y}) \leq \varpi_P(\mathbf{x}) + \varpi_P(\mathbf{y})$;
    (e) If $\mathrm{supp}(\mathbf{x}) \subseteq \mathrm{supp}(\mathbf{y})$, then $\varpi_P(\mathbf{x}) \leq \varpi_P(\mathbf{y})$.

12. Considering the simple fact that an $(n, M)_q$ code $C$ (where $M = |C|$), with packing radius $R_d(C)$ is perfect if, and only if, $|C| \cdot |B_d(\mathbf{x}, R_d(C))| = q^n$, and considering the codes $C_1$ and $C_2$ defined on Example 2.11 and the five posets over $[3]$ presented on that example, show that:

    (a) $C_1$ is perfect only for the metrics $d_Q$ and $d_S$;
    (b) $C_2$ is perfect only for the metrics $d_R$ and $d_S$.

# Chapter 3
# Hierarchical Posets


The hierarchical orders have a peculiar role when considering the poset-metrics, in the sense that, many metric properties of the Hamming metric are actually characterization of hierarchical posets. The most basic goal of coding theory, to maximize the minimum distance, rests on the fact that it determines the error correction capability: considering the Hamming distance $d_H$, given a code $C$ with minimum distance $\delta_{d_H}$, if no more then $\lfloor(\delta_{d_H} - 1)/2\rfloor$ errors occur, a minimum distance decoding algorithm will correctly decode the received messages. As it is well known, in the Hamming case, $\lfloor(\delta_{d_H} - 1)/2\rfloor$ is the packing radius of the code. We saw in Example 2.11 that this is not the case for general posets. We shall see that the minimum distance determines (and is determined by) the packing radius of a poset code, if, and only if, the poset is hierarchical. In this sense, we may say that the primary goal of maximizing the minimum distance remains valid only for the case of hierarchical posets.

We remark that the family of all hierarchical posets is not small, in the sense that it grows exponentially in $n$. However, we should add that it is not known how large this family is, since the family of all posets behaves exponentially as $2^{n^2/4}$ [66].

Metrics determined by hierarchical posets is the only family that is well understood; it is not an exaggeration to say that those metrics are as well understood as the Hamming metric. Our goal in this chapter is to explain them.

In this chapter we present the canonical-systematic form of a generator matrix (Sect. 3.1) and determine the main parameters of a code (minimum distance, packing and covering radius) in Sect. 3.2. Considering the relation between these parameters, we establish necessary and sufficient conditions for a code to be perfect (Sect. 3.3). After that we present some features that are characteristic of hierarchical poset codes (Sect. 3.4). We finish the chapter explaining how the canonical decomposition of a code allows to simplify syndrome decoding (Sect. 3.5).



M. Firer et al., *Poset Codes: Partial Orders, Metrics and Coding Theory*, SpringerBriefs in Mathematics, https://doi.org/10.1007/978-3-319-93821-9_3

## 3.1 Canonical-Systematic Form for the Generator Matrix

In this section, considering a hierarchical poset metric, we aim to present the canonical-systematic form of a generator matrix (Theorem 3.2) and the canonical decomposition of a code (Corollary 3.3). We do not dig into all the details of the proofs, but explain the main ideas. Details of the proofs may be found in [36].

We recall (from Sect. 2.3) that, given a poset $P = ([n], \preceq_P)$ and a vector $\mathbf{x} \in \mathbb{F}_q^n$, $M(\mathbf{x}) := M_P(\mathbf{x})$ is the set of maximal elements in $\mathrm{supp}(\mathbf{x})$ and $\widehat{\mathbf{x}}$ is the cleared out form of $\mathbf{x}$, as defined in Sect. 2.3. We also remind that $\langle \mathrm{supp}(\mathbf{x}) \rangle = \langle \mathrm{supp}(\widehat{\mathbf{x}}) \rangle$ and, for any given $\mathbf{x} \in \mathbb{F}_q^n$, there is $T \in G_P$ such that $T(\mathbf{x}) = \widehat{\mathbf{x}}$.

Let $P_{\mathcal{H}} = ([n], \preceq_{\mathcal{H}})$ be a hierarchical poset with hierarchy spectrum $\mathcal{H} = (H_1, \ldots, H_l)$ and hierarchy array $h = (h_1, \ldots, h_l)$ and let $C \subset \mathbb{F}_q^n$ be an $[n,k]_q$ linear code. Consider the metric $d_{P_{\mathcal{H}}}$ and define $\widehat{C}_0 := \{\mathbf{0}\}$ (the null vector) and $\widehat{C}_i := \{\mathbf{c} \in C;\, M(\mathbf{c}) \subset H_i\}$ for every $i \in \{1, \ldots, l\}$. The sets $\widehat{C}_i$ are not in general vector subspaces (for $i \geq 1$) since $\mathbf{x}, \mathbf{y} \in \widehat{C}_i$ implies $\mathbf{x} + \mathbf{y} \in C_j$ for some $j \leq i$ and, of course, $\mathbf{0} = \mathbf{x} - \mathbf{x} \in \widehat{C}_0 \nsubseteq \widehat{C}_i$ for $i > 0$. However, this same reasoning actually shows that $C_i = \bigcup_{j=0}^{i} \widehat{C}_j$ is a vector subspace, so we have a sequence of nested subspaces

$$\{\mathbf{0}\} = C_0 \subseteq C_1 \subseteq C_2 \subseteq \cdots \subseteq C_l = C.$$

Let $\Lambda(C) := \{t_1, \ldots, t_s\}$ be the set of levels of $P_{\mathcal{H}}$ for which $\widehat{C}_{t_j} \neq \emptyset$, or equivalently, the set of levels for which $C_{t_j} \subsetneq C_{t_{j+1}}$, so that all the non-redundant subspaces are $C_0 \subsetneq C_{t_1} \subsetneq C_{t_2} \subsetneq \cdots \subsetneq C_{t_s} = C$. If we construct a basis for $C$ starting from a basis $\widehat{\beta}_1$ of $C_{t_1}$ and then choosing a maximal set $\widehat{\beta}_2$ of linearly independent vectors in $\widehat{C}_{t_2}$, we get a basis $\widehat{\beta}_1 \cup \widehat{\beta}_2$ of $C_{t_2}$. Continuing recursively, if $\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \cdots \cup \widehat{\beta}_i$ is a basis of $C_{t_i}$ and $\widehat{\beta}_{i+1}$ is a maximal set of linearly independent vectors in $\widehat{C}_{t_{i+1}}$ we find that $\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \cdots \cup \widehat{\beta}_i \cup \widehat{\beta}_{i+1}$ is a basis of $C_{t_{i+1}}$. In doing so we get a basis $\widehat{\beta} = \widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \cdots \cup \widehat{\beta}_s$ of $C$.

We denote $d_{t_j} := |\widehat{\beta}_j|$ and remark that the cardinality $d_{t_j}$ equals $\dim(C_{t_j}) - \dim(C_{t_{j-1}})$, so it is canonical in the sense that it does not depend on the choice of the basis, only on the code $C$.

A matrix $G$ obtained by ordering, bottom-up, the elements of $\widehat{\beta}$ starting from $\widehat{\beta}_1$, is a row reduced generator matrix of $C$. We just made a constructive proof of the following result:

**Theorem 3.1** *Let $P_{\mathcal{H}}$ be a hierarchical poset and let $h = (h_1, \ldots, h_l)$ be its hierarchy array. Then, an $[n,k]_q$ linear code $C$ has a generator matrix $G = \left(G_{k,j}\right)$ where $1 \leq k \leq s$ and $1 \leq j \leq l$, consisting of blocks $G_{k,j}$ of size $d_{t_k} \times h_j$ where $G_{k,j}$ is the null matrix for every $j > t_k$, that is, $C$ has a generator matrix of the form*

$$G = \begin{pmatrix} G_{s,1} & \cdots & G_{s,t_1} & G_{s,t_1+1} & \cdots & G_{s,t_s} & 0 \cdots 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \ddots \vdots \\ G_{1,1} & \cdots & G_{1,t_1} & 0 & \cdots & 0 & 0 \cdots 0 \end{pmatrix}$$

*where the entries of $G_{i,t_i}$ correspond to the $h_{t_i}$ columns of the level $H_{t_i}$.*

The matrix $G$ described in the previous theorem is just a row-echelon reduced form of a generator matrix, and can be obtained by usual elementary operations on the rows. The only fact that concerns hierarchical posets is the following: considering each line as a vector $\mathbf{x} \in \mathbb{F}_q^n$ and writing $\mathbf{x} = (\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^{t_i}, \mathbf{0}, \ldots, \mathbf{0})$, where $\mathbf{x}^j$ is a line of $G_{j,t_i}$, then $M(\mathbf{x}) \subseteq H_{t_i}$.

Since every row of $G_{i,t_i}$ is non-null, we can use them as a pivot (operating on columns) to obtain the cleared out form of each row. This is done one row in a time, considering each $G_{i,t_i}$ to be in a row echelon form, and using always the rightmost position, hence clearing a row does not "spoil" the previously cleared rows. After doing so, we get a matrix of the form

$$\widehat{G} = \begin{pmatrix} 0 \cdots 0 & 0 & 0 \cdots 0 & G_{s,t_s} & 0 \cdots 0 \\ \vdots \ddots \vdots & \vdots & \vdots \ddots \vdots & \ddots & \vdots \ddots \vdots \\ 0 \cdots 0 & G_{1,t_1} & 0 \cdots 0 & \cdots & 0 \cdots 0 \end{pmatrix}.$$

The key point that we should note is that we are substituting the line $\mathbf{x} = (\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^{t_i}, \mathbf{0}, \ldots, \mathbf{0})$ by the line $\mathbf{x}' = (\mathbf{0}, \ldots, \mathbf{0}, \mathbf{x}^{t_i}, \mathbf{0}, \ldots, \mathbf{0})$ of $\widehat{G}$ and since $M(\mathbf{x}) \subseteq H_{t_i}$ we have that $\mathbf{x}' = \widehat{\mathbf{x}}$, the cleared out form of $\mathbf{x}$. To put on a different way, we obtained $\widehat{G}$ out of $G$ by a succession of isometries (of the type operating on ideals, as seen in Sect. 2.3). It follows that $C$ and the code $\widehat{C}$ generated by $\widehat{G}$ are $P_{\mathcal{H}}$-equivalent.

We can move further: considering isometries induced by poset automorphisms, we may assume that each $G_{k,t_k}$ (considering the rows as vectors) has an isometric image $G'_{k,t_k}$ which is in a systematic form $G'_{k,t_k} = [Id_{t_k} | A_{t_k}]$. All that we said is resumed into the following proposition:

**Theorem 3.2 (Canonical-Systematic Form)** *Let $P_{\mathcal{H}}$ be a hierarchical poset with hierarchy array $h = (h_1, \ldots, h_l)$ and let $C$ be an $[n,k]_q$ code. Then $C$ is $P_{\mathcal{H}}$-equivalent to a code $C'$ that has a generator matrix $G' = (G'_{k,j})$ consisting of blocks $G'_{k,j}$ of size $d_{t_k} \times h_j$ such that $G'_{k,j}$ is the null matrix for every $j \neq t_k$ and, for $j = t_k$ it has the form $G'_{k,t_k} = [Id_{t_k} | A_{t_k}]$ where $Id_{t_k}$ is the identity matrix of size $d_{t_k} \times d_{t_k}$ and $A_{t_k}$ is a matrix of size $d_{t_k} \times \left(h_{t_k} - d_{t_k}\right)$. In other words, $G'$ has the following form:*

$$G' = \begin{pmatrix} 0 \cdots 0 & 0 & 0 \cdots 0 & \left[Id_{t_s} | A_{t_s}\right] & 0 \\ \vdots \ddots \vdots & \vdots & \vdots \ddots \vdots & \ddots & \vdots \\ 0 \cdots 0 & \left[Id_{t_1} | A_{t_1}\right] & 0 \cdots 0 & \cdots & 0 \end{pmatrix}.$$

*Proof* Despite the fact that the main idea of the proof is (hopefully) quite clear, the detailed proof is rather technical and lengthy. We refer the interested reader to the original work at [36]. ■

The importance of the canonical-systematic form $G'$ of the generator matrix can be better grasped when we consider the linear code $C'$ generated by $G'$: it is described in the following proposition, an immediate consequence of Theorem 3.2.

**Corollary 3.3 (Canonical Decomposition of Codes)** *Let $P_{\mathcal{H}}$ be a hierarchical poset with spectrum $\mathcal{H} = (H_1, \ldots, H_l)$. Then, a linear code $C$ may be considered, up to $d_{P_{\mathcal{H}}}$-equivalence, to be of the form*

$$C = C_1 \oplus C_2 \oplus \cdots \oplus C_l \tag{3.1}$$

*with* $\operatorname{supp}(C_i) \subseteq H_i$.

*Proof* It follows straightforwardly from Theorem 3.2. ■

The form of a code described in Corollary 3.3 is called the *canonical form*.

*Remark 3.4* We call the generator matrix described in Theorem 3.2 a *canonical-systematic* form, meaning that it is canonical in the levels, in the sense that $\dim(C_i)$ is uniquely determined by the generalized weight hierarchy of $C$ as will be seen in Sect. 3.2. In particular, the levels corresponding to the codes $C_i$ in the decomposition (3.1) with $\dim(C_i) = 0$ correspond to the levels not contained in $\Lambda(C)$, that is, $\dim(C_i) = 0$ if $i \in [n] \setminus \Lambda(C)$.

The restriction of the poset metric $d_{P_{\mathcal{H}}}$ to the $i$-th level $H_i$ is essentially equivalent to the Hamming metric $d_H$ on that subspace of $\mathbb{F}_q^n$, in the sense that, given $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$ with $\operatorname{supp}(\mathbf{x}), \operatorname{supp}(\mathbf{y}) \subseteq H_i$, then $d_P(\mathbf{x}, \mathbf{y}) = d_H(\mathbf{x}, \mathbf{y}) + s_{i-1}$, where $s_i = h_1 + h_2 + \cdots + h_i$.

The block $G'_{i,t_i} = \left[Id_{t_i} | A_{t_i}\right]$ is just the usual systematic generator matrix of $C_{t_i}$ for a Hamming space.

## 3.2 Minimal Distance, Packing and Covering Radius

Let us assume that $C$ is an $[n, k]_q$ code in its canonical decomposition as established in Corollary 3.3:

$$C = C_1 \oplus C_2 \oplus \cdots \oplus C_l.$$

We denote by $k_i$ the dimension of $C_i$ and let $\Omega(C) = \{i; k_i > 0\}$. It is clear that $\Omega(C) = \Lambda(C) = \{t_1, \ldots, t_s\}$, as defined in the beginning of Sect. 3.1, so we may write

$$C = \bigoplus_{i=1,\ldots,s} C_{t_i}.$$

We recall that the restriction $d^l$ of $d_{P_{\mathcal{H}}}$ to $C_l$ is equivalent to the usual Hamming metric $d_H$, differing from it by an additive constant which depends solely on $l$.

**Proposition 3.5** *Under the conditions previously stated, we have that:*

1. *The $i$-th $P_{\mathcal{H}}$-weight $\delta_{i,d_{P_{\mathcal{H}}}}$ of a code $C$ is*

$$\delta_{i,d_{P_{\mathcal{H}}}}(C) = s_{t_j-1} + \delta_{(i-r_{j-1})}\left(C_{t_j}\right),$$

*where $\delta_{(i-r_{j-1})}\left(C_{t_j}\right)$ denotes the $\left(i - r_{j-1}\right)$-generalized Hamming weight of $C_{t_j}$, $r_j = \dim(C_1) + \dim(C_2) + \cdots + \dim(C_j)$ and $r_{j-1} < i \leq r_j$. In particular, the minimal weight is given by*

$$\delta_{d_{P_{\mathcal{H}}}}(C) = s_{t_1-1} + \delta_{d^{t_1}}(C_{t_1});$$

2. *The packing radius is given by*

$$R_{d_{P_{\mathcal{H}}}}(C) = s_{t_1-1} + \left\lfloor \frac{\delta_{d^{t_1}}(C_{t_1}) - 1}{2} \right\rfloor,$$

*where $\lfloor x \rfloor$ is the integer part of $x$;*

3. *The covering radius is*

$$R^{cov}_{d_{P_{\mathcal{H}}}}(C) = R^{cov}_{d_H}(C_r) + \sum_{i=1}^{r} h_i$$

*where $r = \min\{i \in [l-1];\, C_j = \mathbb{F}_q^{h_j}, \forall j > i\}$ if $C_l = \mathbb{F}_q^{h_l}$ or $r = l$ if $C_l \neq \mathbb{F}_q^{h_l}$ and $R^{cov}_{d_H}(C_r)$ is the covering radius of $C_r$ considered as a code in the Hamming space $\mathbb{F}_q^{h_r}$.*

*Proof*

1. Let $\widehat{D}_{t_j} \subset C_{t_j}$ be a code with dimension $(i - r_{j-1})$ such that $|\mathrm{supp}(\widehat{D}_{t_j})| = \delta_{(i-r_{j-1})}(C_{t_j})$. Such code exists because $r_{j-1} < 1$. Hence, $|\langle \mathrm{supp}(\widehat{D}_{t_j})\rangle| = s_{t_j-1} + \delta_{(i-r_{j-1})}(C_{t_j})$. Consider the code

$$\widehat{D} = C_1 \oplus \cdots \oplus C_{t_j-1} \oplus \widehat{D}_{t_j}.$$

It is clear that $\dim(\widehat{D}) = r_{j-1} + (i - r_{j-1}) = i$. Furthermore, noticing that $|\langle \mathrm{supp}(\widehat{D})\rangle| = |\langle \mathrm{supp}(\widehat{D}_{t_j})\rangle| = s_{t_j-1} + \delta_{(i-r_{j-1})}(C_{t_j})$, it follows that $\delta_{P_{\mathcal{H}},i}(C) \leq s_{t_j-1} + \delta_{(i-r_{j-1})}(C_{t_j})$.

We consider now a subcode of $C$ that realizes the $i$-th generalized weight: $D \subset C$ is such that $|\langle \mathrm{supp}(D)\rangle| = \delta_{i,d_{P_{\mathcal{H}}}}$. First of all, note that $D \subset C_1 \oplus \cdots \oplus C_{t_j}$, since otherwise we would have

$$|\langle \mathrm{supp}(D)\rangle| > s_{t_j} > s_{t_j-1} + \delta_{(i-r_{j-1})}\left(C_{t_j}\right) = |\langle \mathrm{supp}(\widehat{D})\rangle|.$$

Hence, we can consider the decomposition

$$D = D_1 \oplus \cdots \oplus D_{t_j},$$

where $D_l \subset C_l$ for $l \in \{1, \ldots, t_j\}$. It is obvious that

$$|\langle \mathrm{supp}(D)\rangle| = |\langle \mathrm{supp}(D_{t_j})\rangle| = |\mathrm{supp}(D_{t_j})| + s_{t_j-1}.$$

By the minimality of $\delta_{(i-r_{j-1})}(C_{t_j})$, we find that

$$|\mathrm{supp}(D_{t_j})| + s_{t_j-1} \geq s_{t_j-1} + \delta_{(i-r_{j-1})}\left(C_{t_j}\right).$$

Since $\delta_{i,d_{P_{\mathcal{H}}}}(C) = |\langle \mathrm{supp}(D)\rangle|$, we have that $\delta_{i,d_{P_{\mathcal{H}}}}(C) \geq s_{t_j-1} + \delta_{(i-r_{j-1})}(C_{t_j})$.

It follows that $\delta_{i,d_{P_{\mathcal{H}}}}(C) = s_{t_j-1} + \delta_{(i-r_{j-1})}(C_{t_j})$.

2. Considering

$$R = s_{t_1-1} + \left\lfloor \frac{\delta_{d^{t_1}} - 1}{2} \right\rfloor,$$

since $C$ is a linear code, it is enough to prove that $B_{P_{\mathcal{H}}}(\mathbf{0}, R) \cap B_{P_{\mathcal{H}}}(\mathbf{c}, R) = \emptyset$ for every non-null $\mathbf{c} \in C$ and that $B_{P_{\mathcal{H}}}(\mathbf{0}, R+1) \cap B_{P_{\mathcal{H}}}(\mathbf{c}, R+1) \neq \emptyset$ for some $\mathbf{c} \in C$.

We recall that $t_1 = \min \Lambda(C)$. Let $\mathbf{c} \in C$ and suppose there is $\mathbf{x} \in B_{P_{\mathcal{H}}}(\mathbf{0}, R) \cap B_{P_{\mathcal{H}}}(\mathbf{c}, R)$. Since $P_{\mathcal{H}}$ is hierarchical, there is a level $H_j$ such that $M(\mathbf{x}) \subseteq H_j$. We claim that $j = t_1$. Indeed, if $j < t_1$, we would have

$$d_{P_{\mathcal{H}}}(\mathbf{x}, \mathbf{c}) = \varpi_{P_{\mathcal{H}}}(\mathbf{c}) \geq \delta_{d_{P_{\mathcal{H}}}}(C) = s_{t_1-1} + \delta_{d^{t_1}}(C_{t_1}) > s_{t_1-1} + \left\lfloor \frac{\delta_{d^{t_1}}(C_{t_1}) - 1}{2} \right\rfloor.$$

On the other hand, if we had $j > t_1$, we would have

$$d_{P_{\mathcal{H}}}(\mathbf{x}, \mathbf{0}) = \varpi_{P_{\mathcal{H}}}(\mathbf{x}) \geq s_{j-1} + 1 > s_{t_1} > s_{t_1-1} + \left\lfloor \frac{\delta_{d^{t_1}}(C_{t_1}) - 1}{2} \right\rfloor,$$

so that $j = t_1$. It follows that

$$\varpi_{P_{\mathcal{H}}}(\mathbf{x} - \mathbf{c}) = s_{t_1-1} + |M(\mathbf{x} - \mathbf{c})|.$$

Since $\varpi_{P_{\mathcal{H}}}(\mathbf{x}-\mathbf{c}) \leq R$,

$$|M(\mathbf{x}-\mathbf{c})| \leq \left\lfloor \frac{\delta_{d^{t_1}}(C_{t_1})-1}{2} \right\rfloor < \frac{\delta_{d^{t_1}}(C_{t_1})}{2}.$$

We remark that $\mathbf{x} \in B_{P_{\mathcal{H}}}(\mathbf{0}, R+1)$ implies that $|M(\mathbf{x}) \cap M(\mathbf{c})| \leq \lfloor \frac{\delta_{d^{t_1}}(C_{t_1})-1}{2} \rfloor$. Assuming $\mathbf{c} \in C$ to be a minimal weight codeword, we have that $\mathbf{c} = \mathbf{c}_{t_1} + \mathbf{c}_{t_2} + \cdots + \mathbf{c}_{t_s}$ with $|\mathrm{supp}(\mathbf{c}_{t_1})| = \delta_{d^{t_1}}(C_{t_1})$. But

$$\begin{aligned} \left|\mathrm{supp}(\mathbf{c}_{t_1})\right| &\leq |M(\mathbf{x}-\mathbf{c})| + |M(\mathbf{x})| \\ &\leq 2\left\lfloor \frac{\delta_{d^{t_1}}(C_{t_1})-1}{2} \right\rfloor < \delta_{d^{t_1}}(C_{t_1}), \end{aligned}$$

a contradiction. Hence, $B_{P_{\mathcal{H}}}(\mathbf{0}, R) \cap B_{P_{\mathcal{H}}}(\mathbf{c}, R) = \emptyset$, the assumption that $\mathbf{x} \in B_{P_{\mathcal{H}}}(\mathbf{0}, R)$.

Now we need to prove that $B_{P_{\mathcal{H}}}(\mathbf{0}, R+1) \cap B_{P_{\mathcal{H}}}(\mathbf{c}, R+1) \neq \emptyset$ for some $\mathbf{c} \in C$. Let $\mathbf{c} \in C$ be a word of minimal weight. We may write $\mathbf{c} = \mathbf{c}_1 + \mathbf{c}_2$ where $\mathrm{supp}(\mathbf{c}_1) \subset H_1 \cup \cdots \cup H_{t_1-1}$ and $\mathrm{supp}(\mathbf{c}_2) \subset H_{t_1}$. So we have that $\varpi_{P_{\mathcal{H}}}(\mathbf{c}) = s_{t_1-1} + |\varpi_H(\mathbf{c}_2)|$ and $|\varpi_H(\mathbf{c}_2)| = \delta_{d^{t_1}}(C_{t_1})$. We consider $A \subset \mathrm{supp}(\mathbf{c}_2)$ with $|A| = \lfloor (\delta_{d^{t_1}}(C_{t_1})-1)/2 \rfloor + 1$. Then, a vector $\mathbf{x}$ with $\mathrm{supp}(\mathbf{x}) = A$ satisfies $\mathbf{x} \in B_{P_{\mathcal{H}}}(\mathbf{0}, R+1) \cap B_{P_{\mathcal{H}}}(\mathbf{c}, R+1)$, as required.

3. Given $\mathbf{x} \in \mathbb{F}_q^n$, we consider the unique decomposition $\mathbf{x} = \mathbf{x}^1 + \cdots + \mathbf{x}^l$ where $\mathrm{supp}(\mathbf{x}^i) \subseteq H_i$. Suppose there is $\mathbf{x} \in \mathbb{F}_q^n$ such that

$$d_{P_{\mathcal{H}}}(\mathbf{x}, \mathbf{c}) > R^{cov}_{d_H}(C_r) + \sum_{i=1}^{r} h_i$$

for every $\mathbf{c} \in \mathcal{C}$. Since $d_{P_{\mathcal{H}}}(\mathbf{x}, \mathbf{c}) = d_H(\mathbf{x}_r, \mathbf{c}_r) + \sum_{i=1}^{r-1} h_i$ for every $\mathbf{c} \in \mathcal{C}$ satisfying $\emptyset \neq M(\mathbf{x}-\mathbf{c}) \subseteq H_r$, then $d_H(\mathbf{x}_r, \mathbf{c}_r) > R^{cov}_{d_H}(C_r)$, which is a contradiction. Therefore,

$$R^{cov}_{d_{P_{\mathcal{H}}}}(C) \leq R^{cov}_{d_H}(C_r) + \sum_{i=1}^{r} h_i.$$

If the equality does not hold in the previous equation, then, for each $\mathbf{x} \in \mathbb{F}_q^n$, there is $\mathbf{c} \in C$ such that $d_H(\mathbf{x}_r, \mathbf{c}_r) < R^{cov}_{d_H}(C_r)$, but this contradicts the minimality of $R^{cov}_{d_H}(C_r)$.

■

As we saw in Example 2.11, for general posets, we may have codes with same minimum distance but different packing radius. This is not the case for hierarchical posets. In the previous proposition we saw that, considering a hierarchical poset, the

packing radius of a code depends only on $\delta_{d^{t_1}}(C_{t_1})$ (on Item 2) and, from Item 1, $\delta_{d^{t_1}}(C_{t_1}) = \delta_{i,d_{P_{\mathcal{H}}}}(C) - s_{t_1-1}$. It follows that, for every linear code, the packing radius is determined by the minimum distance. As we shall see later, this property is a characteristic of hierarchical posets.[1]

## 3.3 Perfect Codes

A code $C \subseteq \mathbb{F}_q^n$ is said to be *$d_P$-perfect* if the balls of radius $R = R_{d_P}(C)$ centered at the codewords covers $\mathbb{F}_q^n$, or, equivalently, if $R_{d_P}(C) = R_{d_P}^{cov}(C)$. As already mentioned, considering the Hamming metric, perfect linear codes are very rare and were classified independently by many researchers in 1973 (see [104, 106] or [112]). In the case of poset metrics, perfect codes are not as seldom, and the study of perfect codes moved mainly into three different directions: (1) fixing a poset and classifying the perfect codes ([1] for the crown poset); (2) fixing a code and determining the poset structures that turn it into a perfect code (as done in [50], considering the extended binary Hamming code) or; (3) fixing the parameters and determining the posets that admit the existence of perfect codes with the given parameters (as in [61]).

For hierarchical posets, considering the expressions for the packing and covering radii of a code given in Proposition 3.5, we can give necessary and sufficient conditions for a code to be $d_P$-perfect. We start with the following Lemma, which proof is omitted since it is trivial.

**Lemma 3.6** *Let $\mathbf{c} \in \mathbb{F}_q^n$ be a codeword with $M(\mathbf{c}) \subset H_{t_1}$. If $\mathbf{x} \in B_{d_{P_{\mathcal{H}}}}(\mathbf{c}, R_{d_{P_{\mathcal{H}}}}(C))$, then $M(\mathbf{x}) \subset H_{t_1}$.*

From this lemma we conclude that, given a $d_{P_{\mathcal{H}}}$-perfect code $C$, $i \in \Lambda(C)$ for every $t_1 \leq i \leq l$, so that $C = C_{t_1} \oplus C_{t_1+1} \oplus \cdots \oplus C_l$ with $\dim(C_i) > 0$ for $i \geq t_1$. We will prove that actually $\dim(C_i) = h_i$, for every $i > t_1$.

**Theorem 3.7** *A code $C$ is $d_{P_{\mathcal{H}}}$-perfect if and only if $C = C_{t_1} \oplus V_{t_1+1} \oplus \cdots \oplus V_h$ and $C_{t_1}$ is a perfect code in $V_{t_1}$ with the Hamming metric, where $V_i = \{\mathbf{x} \in \mathbb{F}_q^n; \operatorname{supp}(\mathbf{x}) \subseteq H_i\}$.*

*Proof* We recall we are assuming $P = P_{\mathcal{H}}$ to be a hierarchical poset with hierarchy array $(h_1, \ldots, h_l)$. Given a linear code $C$, we denote $R = R_{d_P}(C)$. Let us begin assuming that $C$ is $d_P$-perfect and suppose there is $k \in \{t_1+1, \ldots, l\}$ with $V_k \neq C_k$. Let $\mathbf{x} \in V_l \backslash C_l$. Since $C$ is $d_P$-perfect, there is $\mathbf{c} \in C$ such that $\mathbf{x} \in B_P(\mathbf{c}, R)$ and

[1] It is well known that, with the Hamming distance, determining the minimum distance of a code is an NP-hard problem [107], but once the minimum distance is determined, the packing radius is known. For general posets, the problem is much more difficult: considering codes with only two codewords, let us say, $C = \{\mathbf{0}, \mathbf{x}\}$, determining the minimum distance is a linear problem, but determining the packing radius is, even in this two elements case, an NP-hard problem. More details can be found in [25].

by Lemma 3.6 we find that $M(\mathbf{c}) \subset H_k$, so that $\mathbf{c} = \mathbf{c}_{t_1} + \cdots + \mathbf{c}_k$, with $\mathbf{c}_i \in C_i$. But $|\mathrm{supp}(\mathbf{c}_k - \mathbf{x})| \geq 1$, and this implies that $d_P(\mathbf{x}, \mathbf{c}) \geq s_{k-1} + 1 \geq s_{t_1} > R$, a contradiction, hence we must have $C_k = V_k$ for every $k > t_1$. Considering on $V_{t_1}$ the Hamming metric, it follows from Proposition 3.5 that the Hamming minimal weight of $C_{t_1}$ is $\delta_{d^{t_1}}(C_{t_1})$ and this implies that $C_{t_1} \subseteq V_{t_1}$ must be a (Hamming) perfect code in $V_{t_1}$.

Assuming now that $C_{t_1} \subseteq V_{t_1}$ is a (Hamming) perfect code in $V_{t_1}$, we show that $C_{t_1} \oplus V_{t_1+1} \oplus \cdots \oplus V_l$ is a $d_P$-perfect code in $\mathbb{F}_q^n$. Indeed, consider $\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_{t_1} + \cdots + \mathbf{x}_l \in V$, with $\mathbf{x}_i \in V_i$. Assuming $C_{t_1}$ to be a perfect code in $V_{t_1}$ (with the Hamming metric), there is $\mathbf{c}_{t_1} \in C_{t_1}$ such that $d_H\left(\mathbf{c}_{t_1}, \mathbf{x}_{t_1}\right) \leq \left\lfloor (\delta_{d^{t_1}}(C_{t_1}) - 1)/2 \right\rfloor$. Since $\mathbf{c}_{t_1} \in C_{t_1}$ we have that $\mathbf{c} = \mathbf{c}_{t_1} + \mathbf{x}_{t_1+1} + \cdots + \mathbf{x}_l \in C$. But

$$d_P(\mathbf{x}, \mathbf{c}) = \varpi_P(\mathbf{x} - \mathbf{c}) = \varpi_P(\mathbf{x}_{t_1} - \mathbf{c}_{t_1}) = \left|\langle \mathrm{supp}(\mathbf{x}_{t_1} - \mathbf{c}_{t_1}) \rangle\right| \leq s_{t_1-1} + \left\lfloor \frac{\delta_{d^{t_1}}(C_{t_1}) - 1}{2} \right\rfloor = R,$$

so that $\mathbf{x} \in B_P(\mathbf{c}, R)$ and hence $C$ is $d_P$-perfect. ■

We remark that Theorem 3.7 together with the classification of perfect codes under the Hamming metric provide a classification of the $d_{P_{\mathcal{H}}}$-perfect codes.

## 3.4 Characterizations of Hierarchical Poset Metrics

As we stated before, from the point of view of coding theory, hierarchical posets are very peculiar. In this section we shall give some metric properties which characterize a hierarchical poset, in the sense that the property is satisfied for every code if, and only if, the poset which defines the metric is hierarchical.

These properties appear dispersed throughout the literature and were proved using many different and hard-working combinatorial and algebraic tools: characters, association schemes, matroids, etc. For example, the MacWilliams Identity was established using characters in [64] and using matroid theory in [12] and [13]; Wei duality was established in [81] using multisets, and so on. Recently, Machado et al. [74] gave a new proof of these results (and some others), using the canonical decomposition presented in Corollary 3.3, which permits to translate the problems to the much studied Hamming environment. This is a significant shortcut compared to the previous proofs. We follow the line of this last work and present some of the most relevant characterizations presented there. We remark that in [74] there are other characterizations of hierarchical poset metrics, part of which are left to the reader as a list of guided exercises.

**Theorem 3.8** *Let $P = ([n], \preceq)$ be a poset with $l$ levels. Then, $P$ is hierarchical if, and only if, any of the (equivalent) properties below holds:*

***Canonical decomposition:*** *Every linear code admits a $P$-canonical decomposition;*
***MacWilliams' Identity:*** *$P$ admits a MacWilliams Identity*[2]*;*
***Extension Property:*** *$P$ satisfies the MacWilliams Extension Property, that is: given two linear codes $C_1$ and $C_2$, if $t : C_1 \to C_2$ is a linear map preserving the $P$-weight, then $t$ can be extended to a linear isometry $T \in GL_P(\mathbb{F}_q^n)$;*
***Transitivity on spheres:*** *The group of linear isometries acts transitively on spheres of a fixed radius, i.e., $\varpi_P(\mathbf{x}) = \varpi_P(\mathbf{y})$ if, and only if, there is $T \in GL_P(\mathbb{F}_q^n)$ such that $T(\mathbf{x}) = \mathbf{y}$;*
***Minimal distance determines the packing radius:*** *The packing radius $R_{d_P}$ of a linear code $C$ is a function of its minimum distance $\delta_{d_P}$.*

*Proof* **Canonical decomposition:** Corollary 3.3 ensures that a hierarchical poset has a canonical decomposition. For the "*if*" part, let us suppose that $P$ is not hierarchical. Let $H_\alpha$ be the first level where the poset "fails" to be hierarchical, that is, there are $a \in H_{\alpha-1}$ and $b \in H_\alpha$ such that $a$ and $b$ are not comparable and $\alpha$ is minimal with this property. Since $b \in H_\alpha$, there must be $c \in H_{\alpha-1}$ distinct from $a$ such that $c \preceq b$. As a subposet, $\{a, b, c\}$ is isomorphic to the second poset of Example 2.11. The code generated by $\mathbf{e}_a + \mathbf{e}_b$ does not admit a canonical decomposition. Indeed, given $T \in GL_P(\mathbb{F}_q^n)$, Proposition 2.15 ensures that both $\langle T(\mathbf{e}_a)\rangle$ and $\langle T(\mathbf{e}_b)\rangle$ are principal ideals. The isometry $T$ induces an order automorphism $\sigma_T$ and we denote $t_a = \sigma_T(a)$ and $t_b = \sigma_T(b)$. Since $\sigma_T$ is an automorphism, it preserves the structure of levels, so we have that $t_a \in H_{\alpha-1}$ and $t_b \in H_\alpha$. Moreover, since $a$ and $b$ are not comparable, the same applies to $t_a$ and $t_b$, so the ideal $\langle\{t_a, t_b\}\rangle_P$ is not a principal ideal. Since $T$ is linear, we have that $T(\mathbf{e}_a + \mathbf{e}_b) = T(\mathbf{e}_a) + T(\mathbf{e}_b)$ and it follows that $\langle T(\mathbf{e}_a + \mathbf{e}_b)\rangle_P$ is contained in $H_\alpha \cup H_{\alpha-1}$, but not in $H_\alpha$ neither in $H_{\alpha-1}$.

**MacWilliams' Identity:** The proof that a hierarchical poset satisfies the MacWilliams Identity will follow from a more general situation which will be explored in Chap. 5. The guidelines for a proof using the canonical decomposition is given in Exercise 4.

For the "*if*" part, let us consider the smallest non-hierarchical poset $P = (X, \preceq)$: $X = \{1, 2, 3\}$ with a single non-trivial relation $1 \preceq 3$. We consider the codes $C = \{000, 001\}$ and $C' = \{000, 110\}$ over $\mathbb{F}_2$. We have that $W_C^P(z) = W_{C'}^P(z) = 1 + z^2$. On the other hand, we have that

$$C^\perp = \{000, 100, 010, 110\} \text{ and } (C')^\perp = \{000, 110, 001, 111\}.$$

[2]We denote by $W_C^P(z) = \sum_{i=1}^n A_i^P(C) z^i$ the *$P$-weight enumerator* of the code $C$, where $A_i^P(C) = |\{\mathbf{c} \in C; \varpi_P(\mathbf{c}) = i\}|$. Given a poset $P$, the *opposite poset* $P^*$ is defined by $i \preceq_{P^*} j \iff j \preceq_P i$. We say that $P$ admits a MacWilliams identity if the $P$-weight enumerator of a code determines the $P^*$-weight enumerator of its dual, that is, given codes $C_1$ and $C_2$, $W_{C_1}^P(z) = W_{C_2}^P(z)$ if, and only if, $W_{(C_1)^\perp}^{P^*}(z) = W_{(C_2)^\perp}^{P^*}(z)$.

Direct computations show that

$$W_{C^\perp}^{P^*}(z) = 1 + z + z^2 + z^3 \neq 1 + z + 2z^3 = W_{(C')^\perp}^{P^*}(z).$$

For the general case, we consider $a, b$, and $c$ as in the proof of the canonical decomposition. Let $\mathbf{x} = \sum_{i \in \langle\{a,b\}\rangle_P \cap H_{\alpha-1}} \mathbf{e}_i$ and define $C$ and $C'$ as the codes spanned by $\mathbf{e}_b$ and $\mathbf{x}$, respectively. Direct computations show that

$$W_C^P(z) = W_{C'}^P(z) = 1 + (q-1)^{|\langle b \rangle_P|}.$$

It is possible to prove that $W_{C^\perp}^{P^*}(z) \neq W_{(C')^\perp}^{P^*}(z)$. Details can be found in [74, Property $\mathfrak{P}_1$, Theorem 3].

**Extension property:** We consider the canonical decompositions $C = C_1 \oplus \cdots \oplus C_l$ and $C' = C'_1 \oplus \cdots \oplus C'_l$ of two linear codes $C_1$ and $C_2$ and let $t : C \to C'$ be a linear map that preserves the $P$-weight. Since $t$ is assumed to preserve the $P$-weight, given $\mathbf{c}_i \in C_i$, we have that $t(\mathbf{c}_i) = t_i(\mathbf{c}_i) + f_i(\mathbf{c}_i)$ where $f_i : C_i \to \bigoplus_{j<i} C'_j$ and $t_i : C_i \to C'_i$ is a $P$-isometry. Moreover, since $t$ is assumed to be a linear map, we have that both $t_i$ and $f_i$ are linear. The linearity of $t$ ensures that, given $\mathbf{c} = \mathbf{c}_1 + \cdots + \mathbf{c}_l$ with $\mathbf{c}_i \in C_i$, we have that

$$t(\mathbf{c}) = \sum_{i=1}^{l} t_i(\mathbf{c}_i) + f_i(\mathbf{c}_i).$$

Since $P$ is hierarchical and both $\mathrm{supp}(C_i)$ and $\mathrm{supp}(C'_i)$ are subsets of $H_i$, it follows that $t_i$ is also a linear isometry according to the Hamming metric.

The classical MacWilliams Extension ensures the existence of a linear isometry $T_i : \mathbb{F}_q^{h_i} \to \mathbb{F}_q^{h_i}$ such that $T_i|_{C_i} = t_i$. Because $P$ is hierarchical and $\mathrm{supp}(\mathbb{F}_q^{h_i}) = H_i$, the map $T_i$ is a linear isometry according to the poset metric $P$. For each $i \in [l]$, let us consider $\mathbb{F}_q^{h_i}$ as the direct sum $\mathbb{F}_q^{h_i} = C_i \oplus W_i$ and define $F_i : \mathbb{F}_q^{h_i} \to \mathbb{F}_q^n$ as the linear map determined by $F_i(\mathbf{c} + \mathbf{y}) = f_i(\mathbf{c})$ for $\mathbf{c} \in C_i$ and $\mathbf{y} \in W_i$. Since each $\mathbf{x} \in \mathbb{F}_q^n$ may be uniquely decomposed as $\mathbf{x} = \mathbf{x}_i + \cdots + \mathbf{x}_l$ with $\mathrm{supp}(\mathbf{x}_i) \subseteq H_i$ for every $i \in [l]$, the map

$$\begin{array}{rcl} T : \mathbb{F}_q^n & \to & \mathbb{F}_q^n \\ \sum_{i=1}^l \mathbf{x}_i & \mapsto & \sum_{i=1}^l T_i(\mathbf{x}_i) + F_i(\mathbf{x}_i) \end{array}$$

is well defined. Furthermore, by construction, $T$ is a linear $P$-isometry satisfying $T|_C = t$.

For the "*if*" part, let $C_1$ and $C_2$ be the codes spanned by $\mathbf{e}_b$ and $\mathbf{x}$, respectively, where $a, b$, and $c$ are defined as in the proof of the canonical decomposition and $\mathbf{x} = \sum_{i \in \langle\{a,b\}\rangle \cap H_{\alpha-1}} \mathbf{e}_i$. The map $t : C_1 \to C_2$ defined by $t(\lambda \cdot \mathbf{e}_b) = \lambda \cdot \mathbf{x}$ is a linear $P$-isometry between $C_1$ and $C_2$. We claim that this map cannot be extended to a linear $P$-isometry of $\mathbb{F}_q^n$. Indeed, Proposition 2.15 ensures that if $T \in GL_P(\mathbb{F}_q)$,

then $\langle \mathrm{supp}(T(\mathbf{e}_b))\rangle$ is a principal ideal, but $T(\mathbf{e}_b) = t(\mathbf{e}_b) = \mathbf{x}$ and $\langle \mathrm{supp}(\mathbf{x})\rangle$ is clearly not principal.

**Transitivity on spheres:** Given $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$ with $\varpi_P(\mathbf{x}) = \varpi_P(\mathbf{y})$, the map defined by $t(\lambda\mathbf{x}) = \lambda\mathbf{y}$ for every $\lambda \in \mathbb{F}_q$ is a linear map between the spaces spanned by $\mathbf{x}$ and $\mathbf{y}$ preserving the $P$-weight. The extension property ensures that $t$ may be extended to a map $T \in GL_P(\mathbb{F}_q^n)$. Since $T(\mathbf{x}) = \mathbf{y}$, the group $GL_P(\mathbb{F}_q^n)$ acts transitively on the sphere of radius $r = \varpi_P(\mathbf{x}) = \varpi_P(\mathbf{y})$.

For the "*if*" part, we consider $\mathbf{e}_b$ and $\mathbf{x} = \sum_{i \in \langle\{a,b\}\rangle_P \cap H_{\alpha-1}} \mathbf{e}_i$, where $a, b$, and $c$ are defined as in the proof of the canonical decomposition. Since $\varpi_P(\mathbf{e}_b) = \varpi_P(\mathbf{x})$ and $\mathrm{supp}(\mathbf{e}_b)$ generates a principal ideal while $\mathrm{supp}(\mathbf{x})$ does not, it follows, from Proposition 2.15, that $\mathbf{e}_b$ cannot be mapped into $\mathbf{x}$ by any $P$-isometry $T \in GL_P(\mathbb{F}_q^n)$.

**Minimal distance determines the packing radius:** The expression for the packing radius in Proposition 3.5 ensures that, for a hierarchical poset, the packing radius is determined by the minimum distance.

For the "*if*" part, let us consider the smallest non-hierarchical poset: $X = \{1, 2, 3\}$ with a single non-trivial relation $1 \preceq 3$. Consider the binary codes $C_1 = \{000, 001\}$ and $C_2 = \{000, 110\}$. Both codes have minimum distance 2. However, direct computations show that

$$R_{d_P}(C_1) = 1 \text{ and } R_{d_P}(C_2) = 0.$$

As happens with the MacWilliams property, the general case uses the same kind of construction, considering the first level where the poset fails to be hierarchical. Exercise 5 suggests a guidance for the proof, which appears in [74, Property $\mathfrak{P}_5$, Theorem 3]. ■

## 3.5 Syndrome Decoding

Let $C$ be an $[n, k]_q$ linear code with parity check matrix $H$. To perform syndrome decoding, once a message $\mathbf{y}$ is received, we compute $s(\mathbf{y}) = H\mathbf{y}^T$ and then we need to make a search in the table of all syndromes (of coset leaders). The decoding table has $q^{n-k}$ elements.

Considering a hierarchical poset $P$ with $l$ levels, we may consider the canonical decomposition

$$T(C) = C' = C'_1 \oplus C'_2 \oplus \cdots \oplus C'_l.$$

Given $\mathbf{y} \in \mathbb{F}_q^n$ let us consider the decomposition $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2 + \cdots + \mathbf{y}_l$ where $\mathrm{supp}(\mathbf{y}_i) \subset H_i$, the $i$-th level of the poset. Considering each $C'_i \subseteq \mathbb{F}_q^{h_i}$ as a code of a subspace endowed with the Hamming metric $d_H$, we have that, if $\mathbf{c}_i \in \underset{\mathbf{c}\in C'_i}{\mathrm{argmin}}\, d_H(\mathbf{y}_i, \mathbf{c})$, then

$$\mathbf{c}_1 + \mathbf{c}_2 + \cdots + \mathbf{c}_l \in \operatorname*{argmin}_{\mathbf{c} \in C'} d_P(\mathbf{y}, \mathbf{c}).$$

It follows that we just need to perform usual syndrome decoding for each $\mathbf{y}_i$ relatively to the $i$-th component $C'_i$.

To do so, we have two different situations. If $C'_i = \{\mathbf{0}\}$, we know that, independently of other components, $\mathbf{y}_i$ should be $\mathbf{0}$. In case $C'_i \neq \{\mathbf{0}\}$, we need to look at the table of syndromes of $C'_i$, and this is a table with $q^{h_i - k_i}$ elements, where $k_i = \dim(C'_i)$. It follows that we are making at most $\sum_{t_i \in \Lambda(C')} q^{h_i - k_i}$ searches and this, in general, is much smaller than the $q^{n-k}$ entries of the Hamming syndrome decoding table.

## 3.6 Chapter Notes

The canonical decomposition of a code is a feature available only in the case of hierarchical posets. This decomposition allows us to translate many coding problems with hierarchical posets into coding problems with the usual Hamming metric in a smaller code. In this sense we may say that working with a hierarchical poset metric is at least as easy (or at most as difficult) as working with the Hamming metric. It enables us to use many of the results known to be valid in the Hamming case, and this is what permitted us to prove, for example, a MacWilliams identity without resorting to powerful and complex combinatorial tools such as characters, matroids, and association schemes, as the reader can see, for example, in the next two chapters. We stress that, since the existence of a canonical decomposition can be ensured only in the hierarchical case, there is no hope to produce such "clean" proofs for general posets.

Considering the properties stated in Theorem 3.8, the authors suspected that the extension property and the MacWilliams identity would be equivalent properties in a broader range of metrics. This is not the case, as it will be seen in Chap. 7, for both the graph metrics and the combinatorial metrics.

Finally, the extension property of codes can be translated to an extension property of ideals. Indeed, assuming this extension property we have that for any two words with same weight, there is a linear isometry mapping one word to another. It follows that the automorphism of posets induced by such isometry is an extension of an ordering preserving map between the ideals generated by the support of those words. Since any ideal can be realized as the support of some word, we say that hierarchical posets satisfy the *extension property of ideals*. A natural question rises in this context: are hierarchical posets the only ones satisfying such extension property? The answer is easy and is negative (see, for example, Exercise 7 or [102]). However, this leads us to a classification problem, which is still unsolved: Characterize the posets satisfying the extension property of ideals.

## 3.7 Exercises

1. Let $P = ([n], \preceq)$ be a poset. Given a code $C$, we define $\widehat{C}_0 := \{\mathbf{0}\}$ (the null vector) and $\widehat{C}_i := \{\mathbf{x} \in C; M(\mathbf{x}) \subset H_i\}$, the set of codewords whose maximal elements in their support belong to level $i$. Prove that $\widehat{C}_i$ is well defined for every code $C$ if, and only if, $P$ is hierarchical.
2. [74, Theorem 3] Let $P = ([n], \preceq)$ be a poset. Prove that any two ideals $I, J \subseteq [n]$ with $|I| = |J|$ are order isomorphic if, and only if, $P$ is hierarchical.
3. Let $P = ([n], \preceq)$ be a poset. Prove that, given $\mathbf{x} \in \mathbb{F}_q^n$ there is a single level $H_i$ such that $M_P(\mathbf{x}) \subseteq H_i$ if, and only if, $P$ is hierarchical.
4. [74, Theorem 3] (*MacWilliams' Identity*) Let $P = ([n], \preceq)$ be a hierarchical poset and let $C = C_1 \oplus \cdots \oplus C_l$ be the canonical decomposition of a linear code $C$. Let $C_i^{\perp}$ be the dual code of $C_i$ and define $D_i = \{\mathbf{x} \in C_i^{\perp}; \text{supp}(\mathbf{x}) \subseteq H_i(P)\}$.

   (a) Prove that $\mathcal{C}^{\perp} = D_1 \oplus D_2 \oplus \cdots \oplus D_l$.

   Given a code $C \subset \mathbb{F}_q^n$ and $T \subset [n]$, we denote by $C^T$ the code obtained by puncturing (removing) $C$ at the coordinates corresponding to $T$.[3]

   (b) Prove that $W^P_{C_i^{[n]\setminus H_i}}(z) = W^P_{C_i}(z)$ and $W^{P^*}_{D_i^{[n]\setminus H_i}}(z) = W^{P^*}_{D_i}(z)$.

   (c) Consider the codes $\texttt{punct}(C) := C_1^{[n]\setminus H_1} \oplus \cdots \oplus C_l^{[n]\setminus H_l}$ and $\texttt{punct}(D) := D_1^{[n]\setminus H_1} \oplus \cdots \oplus D_l^{[n]\setminus H_l}$. Prove that $\texttt{punct}(C)$ and $\texttt{punct}(D)$ are dual to each other.

   (d) Use the MacWilliams identity for the Hamming metric (actually the existence of such identity) to prove the MacWilliams identity for hierarchical poset codes: if $W^P_C(z) = W^P_{C'}(z)$, then $W^{P^*}_{C^{\perp}}(z) = W^{P^*}_{(C')^{\perp}}(z)$.

5. [74, Proposition 1 and Theorem 3] (*Minimal distance determines the packing radius*) Let $H_\alpha$ be the minimal level where $P$ fails to be hierarchical and consider $a, b \in H_{\alpha-1}, b \in H_\alpha$ with $c \preceq b$ and $a, b$ not comparable. Let $\mathbf{x} = \sum_{i \in \langle\{a,b\}\rangle_P \cap H_{\alpha-1}} \mathbf{e}_i$ and define $C_1$ and $C_2$ as the codes spanned by $\mathbf{e}_b$ and $\mathbf{x}$, respectively.

   (a) Let $C$ be a linear code and let $I = \langle \text{supp}(C) \rangle$ be the ideal generated by the support of $C$ and let $\widehat{C}$ be the code obtained by puncturing $C$ on $[n] \setminus I$. Then, the packing radius of $\widehat{C}$ according to $I$ coincides with the packing radius of $C$ according to $P$.

   (b) Use the previous item to prove that the packing radius of $C_1$ and $C_2$ equals the packing radius of the codes punctured in the coordinates corresponding to the ideals $I = \langle b \rangle_P$ and $J = \langle a, b \rangle_P \setminus \{b\}$, respectively.

[3] Puncturing a code in a set of coordinates means to remove these coordinates from every codeword. By puncturing in a single coordinate an $[n, k]_q$ code we get a shorter code with possibly the same dimension, that is, the punctured code is either an $[n-1, k]_q$ or $[n-1, k-2]_q$ code.

(c) Consider the fact that $P$ is hierarchical up to the level $\alpha - 1$ and use the expression for the packing radius of a hierarchical poset in Proposition 3.5 to show that

$$R_{d_P}(C_1) = \sum_{i=1}^{\alpha-2} n_i + |H_{\alpha-1} \cap \langle b \rangle| \text{ and } R_{d_P}(C_2) = \sum_{i=1}^{\alpha-2} n_i + \left\lfloor \frac{|H_{\alpha-1} \cap \langle b \rangle|}{2} \right\rfloor.$$

(d) Conclude that, if $P$ is not hierarchical, there are codes with same minimum distance but different packing radius.

6. Classify all the posets over $[n]$ for which the packing radius $R_{d_P}(C)$ and the minimum distance $\delta_{d_P}(C)$ are related by the equation $R_{d_P}(C) = \delta_{d_P}(C) - 1$, for every linear code $C \subset \mathbb{F}_q^n$.
7. Let $P$ be a poset consisting of disjoint chains. Prove that $P$ satisfies the extension property of ideals if, and only if, all chains have the same length.

# Chapter 4
# Disjoint Chains with Equal Length: The Niederreiter-Rosenbloom-Tsfasman Metric


## 4.1 Introduction

The study of codes for the Niederreiter-Rosenbloom-Tsfasman metric has a long history of parallel developments.

A first source comes from a series of papers by Niederreiter [82–84]. The three essential parameters of a linear code in Hamming space are its length, dimension and minimum distance and, as we have seen in the Chap. 1, these parameters influence each other (recall the Hamming and Singleton bounds). The most fundamental problems in coding theory consist of fixing a pair of them and maximizing the third one. For instance, one may search for a code with largest minimum distance $\delta$ of given length $n$ and dimension $k$. Using parity-check matrices, this problem may be formulated in matrix form or, equivalently, in terms of families of vectors in $\mathbb{F}_q^{(n-k)}$; in this latter form it was generalized by Niederreiter in [82]. When Brualdi, Graves and Lawrence introduced poset metrics in 1995, one of their main motivations was to present Niederreiter's problem on families of vectors in $\mathbb{F}_q^k$ in a (broader) coding theoretic setting. They indeed showed that this problem corresponds to the coding-theoretic problem of finding a code with largest minimum distance $\delta_P$, of given length $n$ and dimension $k$, with respect to a $P$-metric where $P$ is a disjoint union of chains.

In 1997 Rosenbloom and Tsfasman introduced a new metric in the $\mathbb{F}_q$-vector space of $r \times s$ matrices with an information-theoretic motivation [96]. Since then coding-theoretic questions with respect to this metric have been investigated, questions such as MDS codes [31, 33, 101], MacWilliams Duality [32], structure and decoding of linear codes [86, 89], coverings [15]. The study of those metric spaces was also motivated by the study of distributions of points in the unit cube (see for instance [6, 7, 31, 101]). It was eventually realized that this matrix metric corresponds to a $P$-metric where $P$ is the union of $r$ chains of length $s$.



M. Firer et al., *Poset Codes: Partial Orders, Metrics and Coding Theory*, SpringerBriefs in Mathematics, https://doi.org/10.1007/978-3-319-93821-9_4

A poset which is the disjoint union of chains is hierarchical if and only if it consists of a single chain. The characterization theorem of hierarchical posets of last chapter tells us then that several coding-theoretic properties which are "to be expected" do not hold in the multichains case: the packing radius of a code is not determined by the minimum distance, the group of linear isometries does not act transitively on spheres and MacWilliams identities do not hold. These problems have led naturally to new questions: if the group of linear isometries does not act transitively, can we describe its orbits? If the weight enumerator of a code does not determine the weight enumerator of its dual, is there some kind of metric invariant of the code that does determine the corresponding metric invariant of the dual? We present some of these questions and their answers in this chapter, while other questions will be adressed in the following chapters.

In this chapter we present constructions of perfect codes and MDS codes. Examples of codes with the same minimum distance and different packing radii are also presented. We describe the orbits of the group of linear isometries and show that these orbits are parametrized by *shapes*, which are invariants of the action of the group of linear isometries which describe the equivalence classes of ideals modulo poset automorphisms. As an application of the description of this group we obtain a standard form for generator matrices. We close this chapter with a memoryless channel which is weakly matched to the NRT metric, in the sense that the probability that the vector $\mathbf{y}$ is received given that the vector $\mathbf{x}$ is sent depends not only on the distance between these vectors but also on the shape of the support ideal of the difference vector $\mathbf{y} - \mathbf{x}$.

## 4.2 Geometry of NRT Spaces

Let $\mathcal{R}(r, s)$ be the poset which is the union of $r$ chains $R_1, R_2, \ldots, R_r$ of equal length $s$. Contrary to what has been done in the previous chapters, we will not take indices in the set $[rs]$ but rather use explicitly the poset $\mathcal{R}(r, s)$ as index set for the coordinates in $\mathbb{F}_q^{rs}$. Accordingly, each vector $\mathbf{x} \in \mathbb{F}_q^{rs}$ will be written with double indexed coordinates (as in Example 2.13):

$$\mathbf{x} = \left(x_1^1, x_2^1, \ldots, x_s^1; x_1^2, x_2^2, \ldots, x_s^2; \cdots; x_1^r, x_2^r, \ldots, x_s^r\right) \in \mathbb{F}_q^{rs}$$

where $x_1^i, x_2^i, \ldots, x_s^i$ are the coordinates corresponding to the $i$-th chain. It is useful to think of $\mathbb{F}_q^{rs}$ as the direct sum of its "chain subspaces", and we will write

$$\mathbf{x} = (\mathbf{x}^{(1)}; \mathbf{x}^{(2)}; \ldots; \mathbf{x}^{(r)})$$

where $\mathbf{x}^{(i)}$ is the vector $\mathbf{x}^{(i)} = \left(x_1^i, x_2^i, \ldots, x_s^i\right)$. We use this decomposition in order to express the NRT-weight of a nonzero $\mathbf{x} \in \mathbb{F}_q^{rs}$:

$$\varpi_{\mathcal{R}(r,s)}(\mathbf{x}) = \sum_{i=1}^{r} m\left(\mathbf{x}^{(i)}\right)$$

where $m\left(\mathbf{x}^{(i)}\right) := \max\left\{j; x^i_j \neq 0\right\}$ for $\mathbf{x}^{(i)} \neq 0$, and $m(\mathbf{0}) = 0$. In fact, it follows from the definition of $P$-weight that the NRT-weight of a nonzero vector $\mathbf{x}$ is the sum of the heights of the maximal elements of $\langle \text{supp}(\mathbf{x})\rangle$. Note that $\mathcal{R}(r, s)$ is a hierarchical poset *iff* it consists of a single chain of height $s$ (when $r = 1$) or of multiple chains of unit height (when $s = 1$).

In order to emphasize the poset structure of $\mathcal{R}(r, s)$ and the poset metric being used we will adopt the notation $\mathbb{F}_q^{[r\times s]}$ as a shorthand for the metric space $(\mathbb{F}_q^{rs}, d_{\mathcal{R}(r,s)})$. On the other hand, we will usually write only $\mathcal{R}$ instead of $\mathcal{R}(r, s)$.

### 4.2.1 Orbits of the Group of Linear Isometries and Shapes of Vectors

In NRT spaces the distance between vectors may not tell the whole story; codes with same minimum distance may have distinct packing radii and, as we will see in the following, vectors with the same NRT-weight may not lie in the same $GL_{\mathcal{R}}(\mathbb{F}_q^{[r\times s]})$-orbit. Nevertheless, there is a $GL_{\mathcal{R}}(\mathbb{F}_q^{[r\times s]})$-invariant, the *shape* of a vector, which parametrizes those orbits.

It is clear that if $\mathbf{u}$ and $\mathbf{v}$ lie in the same $GL_{\mathcal{R}}(\mathbb{F}_q^{[r\times s]})$-orbit then these vectors have the same $\mathcal{R}$-weight. The converse will not hold in general, since by Theorem 3.8 we know that $GL_P(\mathbb{F}_q^n)$ acts transitively on spheres if and only if $P$ is a hierarchical poset. Moreover, by Theorem 2.16 each linear NRT symmetry $T : \mathbb{F}_q^{[r\times s]} \to \mathbb{F}_q^{[r\times s]}$ induces an automorphism $\sigma_T : \mathcal{R} \to \mathcal{R}$, and if $T(\mathbf{u}) = \mathbf{v}$, then $\sigma_T(\langle \text{supp}(\mathbf{u})\rangle) = \langle \text{supp}(\mathbf{v})\rangle$. Hence, if two vectors lie in the same $GL_{\mathcal{R}}(\mathbb{F}_q^{[r\times s]})$-orbit, their support ideals lie in the same $\text{Aut}(\mathcal{R})$-orbit.

The automorphism group of $\mathcal{R}$ consists of permutations of chains, and as such it is isomorphic to the symmetric group $S_r$. It is straigthforward to identify which ideals lie in the same orbit if we parametrize ideals by the positions (chains) where their maximal elements appear. Let $\mathcal{I}(P)$ be the set of all ideals of a poset $P$.

Given $I \in \mathcal{I}(\mathcal{R})$, for each $l = 1, 2, \ldots, r$, let $m_l$ be the maximal element of $I$ in the $l$-th chain, if such element exists, and let $m_l$ be equal to zero otherwise; let $\mathbf{m}(I) = (m_1, m_2, \ldots, m_r)$. This is an integer vector such that $0 \leq m_l \leq s$ for all $l$. Conversely, any element of the set

$$X = \{(m_1, m_2, \ldots, m_r) \in \mathbb{Z}^r; 0 \leq m_l \leq s \text{ for all } l\}$$

determines an ideal of $\mathcal{R}$ by following (backwards) the recipe above. Thus $I \mapsto \mathbf{m}(I)$ defines a bijection from the set of ideals $\mathcal{I}(\mathcal{R})$ of $\mathcal{R}$ to the set $X$. Moreover, the $\text{Aut}(\mathcal{R})$-action on the set of ideals of $\mathcal{R}$ corresponds to the $S_r$-action on $X$ by permutations of coordinates, and therefore two ideals $I, J$ are isomorphic via an automorphism *iff* $\mathbf{m}(I)$ and $\mathbf{m}(J)$ differ only by a permutation of coordinates.

For instance, if $s = 2$ and $r = 3$ then the vectors $\mathbf{u} = (1, 0; 1, 0; 0, 0)$ and $\mathbf{v} = (1, 1; 0, 0; 0, 0)$ have the same NRT-weight, but their support ideals are described respectively by the vectors $(1, 1, 0)$ and $(2, 0, 0)$, which are clearly not in the same $S_3$-orbit; therefore there exists no linear symmetry mapping $\mathbf{u}$ to $\mathbf{v}$.

So when are two vectors in the same orbit of the linear symmetry group? This question was adressed by Dougherty and Skriganov in [32]; we will rephrase the answer in terms of *shapes* as introduced (later) by Barg and Purkayastha [7] (see also [9]).

Let $I$ be an (nonempty) ideal of $\mathcal{R}$ and let $\mathbf{m}(I) = (m_1, m_2, \ldots, m_r)$ be its vector of maximal elements. The *shape* of the ideal $I$ is the vector $\text{shape}(I) = (e_1, e_2, \ldots, e_s)$, where

$$e_j := |\{l \in \{1, \ldots, r\}; m_l = j\}| \tag{4.1}$$

for $j = 1, 2, \ldots, s$.

A moment's reflection shows that there exists a permutation of coordinates taking $\mathbf{m}(I)$ to $\mathbf{m}(J)$ *iff* $I$ and $J$ have the same shape. Hence, the orbits of the ideals of $\mathcal{R}$ under the action of $\text{Aut}(\mathcal{R})$ are in bijection with the set of all possible shapes:

$$\text{Shapes}(s, r) := \{(e_1, e_2, \ldots, e_s) \in \mathbb{Z}^s; 0 \le e_j \text{ for all } j, \sum_j e_j \le r\}.$$

Finally, we can talk about the shape of a *vector* $u \in \mathbb{F}_q^{[r\times s]}$ by defining $\text{shape}(\mathbf{u}) = \text{shape}(\langle \text{supp}(\mathbf{u})\rangle)$. It is easy to see that the shape of a vector $\mathbf{u} = (\mathbf{u}^{(1)}; \mathbf{u}^{(2)}; \ldots; \mathbf{u}^{(r)})$ is the vector $\mathbf{e} = (e_1, e_2, \ldots, e_s)$ where

$$e_j := |\{l \in \{1, \ldots, r\}; \varpi_{\mathcal{R}}(\mathbf{u}^{(l)}) = j\}|. \tag{4.2}$$

Note that the NRT-weight can be defined in terms of shapes: if $\mathbf{e}$ is the shape of the vector $\mathbf{u}$ then

$$\varpi_{\mathcal{R}}(\mathbf{u}) = \sum_{i=1}^{s} i e_i. \tag{4.3}$$

In the language of shapes, the description of $GL_{\mathcal{R}}(\mathbb{F}_q^{[r\times s]})$-orbits is as follows:

**Proposition 4.1** *[32, Proposition 2.2 (ii)] The $GL_{\mathcal{R}}(\mathbb{F}_q^{[r\times s]})$-orbit of a nonzero vector $\mathbf{u} \in \mathbb{F}_q^{[r\times s]}$ is the set of all vectors $\mathbf{v} \in \mathbb{F}_q^{[r\times s]}$ which have the same shape as $\mathbf{u}$.*

By the previous arguments, two vectors $\mathbf{u}$ and $\mathbf{v}$ are in the same $GL_{\mathcal{R}}(\mathbb{F}_q^{[r\times s]})$-orbit *iff* their support ideals are in the same $\text{Aut}(\mathcal{R})$-orbit.

### 4.2.2 Minimum Distance and Packing Radius

A property that characterizes hierarchical posets is that the packing radius of a linear code is a function of the minimal distance. In NRT spaces the packing radius depends on the minimum distance and on the *shapes* of vectors of minimum weight.

We remark that if $r = 1$ then we are dealing with a hierarchical poset; in this case it is easy to see that if $\delta_{d_{\mathcal{R}}}(\mathcal{C}) = \delta$ then $R_{d_{\mathcal{R}}}(\mathcal{C}) = \delta - 1$ for any linear code $\mathcal{C}$. By Proposition 1.4 and the fact that $\mathcal{C}$ is linear it is enough to show that $B_{d_R}(\mathbf{0}, \delta - 1) \cap B_{d_R}(\mathbf{c}, \delta - 1) = \emptyset$ for every $\mathbf{c} \in \mathcal{C}$. Suppose, by contradiction, that there exist vectors $\mathbf{v} \in \mathbb{F}_q^s$ and $\mathbf{0} \neq \mathbf{c} \in \mathcal{C}$ such that

$$\mathbf{v} \in B_{d_R}(\mathbf{0}, \delta - 1) \cap B_{d_R}(\mathbf{c}, \delta - 1).$$

Then $\varpi_{\mathcal{R}}(\mathbf{v}) \leq \delta - 1$ and $\varpi_{\mathcal{R}}(\mathbf{c} - \mathbf{v}) \leq \delta - 1$. However, by the expression of the NRT metric in the single chain case, if $\varpi_{\mathcal{R}}(\mathbf{v}) \leq \delta - 1$ then $\varpi_{\mathcal{R}}(\mathbf{c} - \mathbf{v}) = \delta$, a contradiction. Hence $R_{d_{\mathcal{R}}}(\mathcal{C}) = \delta - 1$.

In the case of multiple chains, $r \geq 2$, the packing radius is affected by the shape of the minimum vectors.

Consider, for instance, $s = r = m$. Let $\beta = \{\mathbf{e}_j^i; 1 \leq i, j \leq m\}$ be the canonical basis of $\mathbb{F}_q^{[m \times m]}$; let $\mathcal{C}_0$ be the code generated by the vector

$$\mathbf{v}_0 = \sum_{j=1}^{m} \mathbf{e}_j^1$$

and, for $k = 1, 2, \ldots, m - 1$, let $\mathcal{C}_k$ be the code generated by the vector

$$\mathbf{v}_k = \sum_{j=1}^{m-k} \mathbf{e}_j^1 + (\mathbf{e}_1^2 + \cdots + \mathbf{e}_1^{k+1}).$$

All of those codes possess the same minimum distance $\delta_{d_{\mathcal{R}}}(\mathcal{C}_k) = \varpi_{\mathcal{R}}(\mathbf{v}_k) = m$. However, the shapes of the vectors $\mathbf{v}_k$ are all distinct, and so are the packing radii of the associated codes which are given by

$$R_{d_{\mathcal{R}}}(\mathcal{C}_k) = \begin{cases} m - k - 1 & \text{if } 0 \leq k < \lfloor m/2 \rfloor, \\ \lfloor (m-1)/2 \rfloor & \text{if } \lfloor m/2 \rfloor \leq k \leq m - 1 \end{cases}$$

For instance, for $k = 1$,

$$\mathbf{u} = \sum_{j=1}^{m-1} \mathbf{e}_j^1 \in B_{d_{\mathcal{R}}}(0, m-1) \cap B_{d_{\mathcal{R}}}(\mathbf{v}_k, m-1)$$

and $B_{d_{\mathcal{R}}}(0, m-2) \cap B_{d_{\mathcal{R}}}(\alpha \mathbf{v}_k, m-2) = \varnothing$ for all $\alpha \in \mathbb{F}_q^*$, showing that the packing radius of $\mathcal{C}_1$ is $m-2$.

Those codes show that all the possible values for the packing radius for a code with minimum distance $m$ are attained in $\mathbb{F}_q^{[m\times m]}$ (Proposition 1.4). See Exercise 1 for a further development of this problem. More generally, the simple problem of determining the packing radius of a unidimensional code in $\mathbb{F}_q^{[r\times s]}$ is NP-hard. More details will be given in Sect. 6.1.

## 4.3 Perfect Codes and MDS Codes

### *4.3.1 Perfect Codes*

One of the first results in Brualdi et al. [14] is the classification of perfect poset codes when the poset is a single chain. However, constructing perfect codes in NRT spaces is a hard task: in the same paper it is proved that all perfect codes in the case of 2 chains are trivial. In this section we will present those results together with an elementary "lifting" construction in the cases of 3 or more chains. A perfect code $\mathcal{C}$ with packing radius $m$ will be called an "$m$-perfect code".

**Single Chain**

In the single chain case we can describe all perfect codes, even the nonlinear ones [14]. Note that here we deal with a hierarchical poset and, as such, the linear perfect codes were already described in Chap. 3.

Suppose that $\mathcal{C}$ is a (nonlinear) perfect code in $\mathbb{F}_q^{[1\times s]}$ of minimum distance $\delta$. In the case of a single chain we have seen that the packing radius is $R = \delta - 1$. Since $\mathcal{C}$ is perfect,

$$\mathbb{F}_q^{[1\times s]} = \bigcup_{\mathbf{c}\in\mathcal{C}} B_{d_{\mathcal{R}}}(\mathbf{c}, R)$$

and $B_{d_{\mathcal{R}}}(\mathbf{c}, R) \cap B_{d_{\mathcal{R}}}(\mathbf{c}', R) = \emptyset$ if $\mathbf{c}$ and $\mathbf{c}'$ are distinct points of $\mathcal{C}$.

Given $\mathbf{v} \in \mathbb{F}_q^{[1\times s]}$, there exists a unique $\mathbf{c} \in \mathcal{C}$ such that $\mathbf{v} \in B_{d_{\mathcal{R}}}(\mathbf{c}, R)$. By the expression of the NRT metric, this means that $\mathcal{C}$ and $\mathbf{v}$ have the same last $s-R$ coordinates, i.e., $v_j = c_j$ for $j = R+1, R+2, \ldots, s$. Something interesting happens if we consider another vector $\mathbf{u} \in \mathbb{F}_q^{[1\times s]}$ which has the same last $s-R$ coordinates as those of $\mathbf{v}$: there is a unique codeword $\mathbf{c}' \in \mathcal{C}$ such that $d_{\mathcal{R}}(\mathbf{u}, \mathbf{c}') \leq R$, but once again this happens if and only if $c'_j = u_j$ for every $j \geq R+1$. Since $u_j = v_j = c_j$ for every such coordinate, it follows that $d_{\mathcal{R}}(\mathbf{c}, \mathbf{c}') \leq R$, which implies that $\mathbf{c}$ and $\mathbf{c}'$ are equal!

Hence, if we consider the direct sum decomposition $\mathbb{F}_q^s = \mathbb{F}_q^R \oplus \mathbb{F}_q^{(s-R)}$ induced by the vector decomposition

$$(v_1, \ldots, v_s) = (v_1, \ldots, v_R, 0, \ldots, 0) + (0, \ldots, 0, v_{R+1}, \ldots, v_s)$$

then every codeword $\mathbf{c}$ is of the form

$$\mathbf{c} = (\mathbf{a}', \mathbf{a})$$

with $\mathbf{a} \in \mathbb{F}_q^{s-R}$ and $\mathbf{a}' \in \mathbb{F}_q^R$.

Every point of $\mathbb{F}_q^{(s-R)}$ appears as the last $s - R$ coordinates of a codeword: If $\mathbf{b}$ is in $\mathbb{F}_q^{(s-R)}$ then there must be an element $\mathbf{c} \in \mathcal{C}$ such that $d_{\mathcal{R}}(\mathbf{c}, (\mathbf{0}, \mathbf{b})) < R$, i.e., $\mathbf{c}$ is of the form $\mathbf{c} = (\mathbf{b}', \mathbf{b})$. The perfectness condition ensures that $\mathbf{b}'$ is unique. Thus we have a mapping $f : \mathbb{F}_q^{(s-R)} \to \mathbb{F}_q^R$ defined by the rule $f(\mathbf{a}) = \mathbf{a}'$ *iff* $(\mathbf{a}', \mathbf{a})$ is an element of $\mathcal{C}$.

Conversely, it is easy to see that any function $f : \mathbb{F}_q^{(s-R)} \to \mathbb{F}_q^R$ defines a code

$$\mathcal{C} = \{(f(\mathbf{a}), \mathbf{a}) \in \mathbb{F}_q^{[1\times s]};\ \mathbf{a} \in \mathbb{F}_q^{(s-R)}\}$$

which is $R$-perfect with respect to the NRT metric. This code is linear if and only if $f$ is linear.

**Two Chains**

In [14] the authors show, by a counting argument, that there exist no nontrivial perfect codes in the poset which is a union of two chains of equal length. The only perfect codes are the whole space $\mathbb{F}_q^{[2\times s]}$ itself and its unitary subsets, and in particular the linear ones are the zero subspace and the whole space.

**Three or More Chains**

As far as we know, little has been done for three chains or more and no classification is available. We present here an elementary construction for perfect codes based on the description of the perfect codes in one-chain case above [14] and on the construction of 1-perfect codes in $\mathbb{F}_2^{[3\times s]}$ presented by Gastoldi and Carmelo in [15]. Given $1 \leq t \leq s$, let $\pi_t : \mathbb{F}_q^{[r\times s]} \to \mathbb{F}_q^{[r\times t]}$ be the projection that erases the last $s - t$ coordinates in each chain, i.e., $\pi_t(\mathbf{v}) = (v_1^1, \ldots, v_t^1; \ldots; v_1^r, \ldots, v_t^r)$. This projection is a particular case of the puncturing construction which will be used in the following for MDS codes, but in this part we are interested not in codes obtained by projecting; we are interested instead in preimage by $\pi_t$ of perfect codes in $\mathbb{F}_q^{[r\times t]}$.

Let $m$ be an integer such that $1 \leq m \leq t$. If $C$ is an $m$-perfect code in $\mathbb{F}_q^{[r\times t]}$, let $\overline{C} \subseteq \mathbb{F}_q^{[r\times s]}$ be the preimage of $C$ by $\pi_t$. We can describe this code explicitly if we consider the direct sum decomposition

$$\mathbb{F}_q^{sr} = \mathbb{F}_q^{tr} \oplus \mathbb{F}_q^{(s-t)r}$$

in the following sense: identifying $\mathbb{F}_q^{r\times s}$ with $r \times s$ matrices over $\mathbb{F}_q$, this decomposition is obtained as follows:

$$\mathbf{v} = \begin{bmatrix} v_1^1 & \dots & v_s^1 \\ v_1^2 & \dots & v_s^2 \\ \vdots & \ddots & \vdots \\ v_1^r & \dots & v_s^r \end{bmatrix} = \begin{bmatrix} v_1^1 & \dots & v_t^1 & 0 & \cdots & 0 \\ v_1^2 & \dots & v_t^2 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ v_1^r & \dots & v_t^r & 0 & \cdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 & v_{t+1}^1 & \dots & v_s^1 \\ 0 & \cdots & 0 & v_{t+1}^2 & \dots & v_s^2 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & v_{t+1}^r & \dots & v_s^r \end{bmatrix}$$

Note that the inclusion of $\mathbb{F}_q^{[r\times t]}$ in $\mathbb{F}_q^{[r\times s]}$ is an isometry onto its image. With regard to this decomposition, $\overline{C}$ is the code

$$\overline{C} = C \oplus \mathbb{F}_q^{(s-t)r} = \{(\mathbf{c}_1, \mathbf{c}_2)\,;\, \mathbf{c}_1 \in C, \mathbf{c}_2 \in \mathbb{F}_q^{(s-t)r}\}.$$

It is clear that $\overline{C}$ is an $m$-covering of $\mathbb{F}_q^{[r\times s]}$: given $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2) \in \mathbb{F}_q^{[r\times s]}$, since $C$ is $m$-perfect in $\mathbb{F}_q^{[r\times t]}$ there exists $\mathbf{c}_1 \in C$ such that $d_{\mathcal{R}(r,t)}(\mathbf{v}_1, \mathbf{c}_1) \leq m$. By the definition of $\overline{C}$ the vector $\mathbf{c} = (\mathbf{c}_1, \mathbf{v}_2)$ lies in $\overline{C}$ and

$$d_{\mathcal{R}(r,s)}(\mathbf{v}, \mathbf{c}) = d_{\mathcal{R}(r,t)}(\mathbf{v}_1, \mathbf{c}_1) \leq m.$$

On the other hand, $m$ is the packing radius of $\overline{C}$: If $\mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2), \mathbf{c}' = (\mathbf{c}_1', \mathbf{c}_2')$ are two codewords of $\overline{C}$ and $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2)$ lies in the intersection of $B_{d_{\mathcal{R}(r,s)}}(\mathbf{c}, m)$ and $B_{d_{\mathcal{R}(r,s)}}(\mathbf{c}', m)$ then, since $m \leq t$, it follows that $\mathbf{c}_2 = \mathbf{v}_2 = \mathbf{c}_2'$, but then it also follows that $d_{\mathcal{R}(r,t)}(\mathbf{c}_1, \mathbf{v}_1) = d_{\mathcal{R}(r,s)}(\mathbf{c}, \mathbf{v}) \leq m$ and $d_{\mathcal{R}(r,t)}(\mathbf{c}_1', \mathbf{v}_1) = d_{\mathcal{R}(r,s)}(\mathbf{c}', \mathbf{v}) \leq m$, a contradiction, since $\mathbf{c}_1$ and $\mathbf{c}_1'$ are distinct codewords of $C$ which is a $m$-perfect code in $\mathbb{F}_q^{[r\times t]}$.

In particular, if $C \subset \mathbb{F}_q^r$ is a 1-perfect code with respect to the Hamming metric, i.e., 1-perfect in $\mathbb{F}_q^{[r\times 1]}$, then $C$ can be "lifted" to a 1-perfect code $\overline{C}$ in any NRT space $\mathbb{F}_q^{[r\times s]}$ for every $s \geq 2$. Applying this to the Hamming $q$-ary codes, this construction yields 1-perfect codes in $\mathbb{F}_q^{[r\times s]}$ whenever $r = \dfrac{q^k - 1}{q - 1}$, $k \geq 1$ and $s \geq 2$. We remark that every 1-perfect code in Hamming space, linear or not, has the same parameters as a Hamming code [48].

### 4.3.2 MDS Codes

**Singleton Bound**

The Singleton bound for the NRT metric appears in [96]; later, it was proved for every $P$-metric [51]. This is the proof that we present below.

Let us consider briefly an arbitrary poset $P$. Let $I$ be a nontrivial ideal of $P$ (i.e. $P$ is neither empty nor the whole poset $P$). It can be shown that if $|I| = r$ then:

1. If $r < r_1 \leq |P|$ then there exists an ideal $J$ of $P$ such that $I \subset J$ and $|J| = r_1$;
2. If $0 \leq r_0 < r$ then there exists an ideal $J$ of $P$ such that $J \subset I$ and $|J| = r_0$.

**Proposition 4.2 ([51])** *Let $P$ be a poset with $n$ elements and let $C \subseteq \mathbb{F}_q^n$ be a (possibly nonlinear) $(n, M, \delta)_q$ code. Then*

$$M \leq q^{n-\delta+1}.$$

*If $C$ is a linear $[n, k, \delta]_q$ code, then*

$$k + \delta \leq n + 1.$$

*Proof* Let $\mathbf{u}$ and $\mathbf{v}$ be elements of $C$ such that $d_P(\mathbf{u}, \mathbf{v}) = \delta(C)$. Then the ideal $\langle \text{supp}(\mathbf{u} - \mathbf{v}) \rangle_P$ has $\delta(C)$ elements and, by the remarks above, there exists an ideal $J \subset \langle \text{supp}(\mathbf{u} - \mathbf{v}) \rangle_P$ with $\delta(C) - 1$ elements. No two codewords of $C$ coincide on the complement of $J$, since otherwise their distance would be less than $\delta(C)$, and therefore the restriction of a codeword to its vector of "$J$-coordinates" defines an injective mapping from $C$ to $\mathbb{F}_q^{n-\delta(C)+1}$. ■

We will say that a $P$-code $C$ is *maximum distance separable* (MDS) if equality is attained in the Singleton bound. In this section we present some constructions of MDS codes in NRT spaces.

**Hamming Weight and NRT Weight**

First of all, every code $C$ in $\mathbb{F}_q^{[r\times s]}$ which is an MDS code with respect to the Hamming weight is also an MDS code for the NRT-weight. Since $\varpi_H(v) \leq \varpi_{\mathcal{R}}(\mathbf{v})$ for every $\mathbf{v} \in \mathbb{F}_q^{[r\times s]}$ it follows that $\delta_{d_H}(C) \leq \delta_{d_{\mathcal{R}}}(C)$ for every linear code $C \subseteq \mathbb{F}_q^{[r\times s]}$, and if $C$ is Hamming-MDS with dimension $k$ then

$$rs - k + 1 = \delta_{d_H}(C) \leq \delta_{d_{\mathcal{R}}}(C) \leq rs - k + 1$$

and hence $rs - k + 1 = \delta_{d_{\mathcal{R}}}(C)$.

**Puncturing MDS Codes**

A second construction appears in [33] and provides a kind of "punctured code construction" for the NRT metric. If $C$ is a linear code in $\mathbb{F}_q^n$, a punctured code is obtained by deleting one or more coordinates of $C$. Deleting $m$ coordinates of a Hamming-MDS $[n, k]_q$ code yields a $[n-m, k]_q$ Hamming-MDS code if $m < n-k$ (see [38], for instance). The following construction does the same for the NRT metric but in this context one has to chose more carefully the coordinates to be deleted. In what follows we will identify $\mathbb{F}_q^{[r\times s]}$ with the $\mathbb{F}_q$-vector space $\text{Mat}_{r,s}(\mathbb{F}_q)$ of $r \times s$ matrices in the usual way.

If $r' \leq r$ and $s' < s$, consider the projection of $\text{Mat}_{r,s}(\mathbb{F}_q)$ onto $\text{Mat}_{r',s'}(\mathbb{F}_q)$ that selects the upper right corner, i.e., the projection

$$\begin{bmatrix} A_{(r')\times(s-s')} & B_{r'\times s'} \\ C_{(r-r')\times(s-s')} & D_{(r-r')\times s'} \end{bmatrix} \in \text{Mat}_{r,s}(\mathbb{F}_q) \mapsto B_{r'\times s'} \in \text{Mat}_{r',s'}(\mathbb{F}_q).$$

Let $C$ be an MDS code in $\mathrm{Mat}_{r,s}(\mathbb{F}_q)$. If its dimension is not "too big" then the projection takes $C$ onto another MDS code in $\mathrm{Mat}_{r',s'}(\mathbb{F}_q)$.

**Theorem 4.3 ([33])** *Let $C$ be an MDS $[rs, k]_q$ code in $\mathrm{Mat}_{r,s}(\mathbb{F}_q)$. If $r' \leq r$, $s' < s$ and $k \leq r's'$ then the restriction of $\pi : \mathrm{Mat}_{r,s}(\mathbb{F}_q) \to \mathrm{Mat}_{r',s'}(\mathbb{F}_q)$ to $C$ is injective and $\pi(C)$ is a MDS $[rs - r's', k]_q$-code in $\mathrm{Mat}_{r',s'}(\mathbb{F}_q)$.*

The injectivity goes as follows: writing $\varpi_{r,s}$ (resp. $\varpi_{r',s'}$) for the NRT-weight in $\mathrm{Mat}_{r,s}(\mathbb{F}_q)$ (resp. $\mathrm{Mat}_{r',s'}(\mathbb{F}_q)$), the weight of $\mathbf{v} \in \mathrm{Mat}_{r,s}(\mathbb{F}_q)$ satisfies

$$\varpi_{r,s}(\mathbf{v}) \leq \varpi_{r',s'}(\pi(\mathbf{v})) + r'(s - s') + s(r - r') = \varpi_{r',s'}(\pi(\mathbf{v})) + rs - r's'.$$

If $\mathbf{v} \in C$ and $\pi(\mathbf{v}) = 0$ then $\varpi_{r,s}(\mathbf{v}) \leq rs - r's' \leq rs - k$. On the other hand, since $C$ is an MDS code, if $\mathbf{v} \neq \mathbf{0}$ then $\varpi_{r,s}(\mathbf{v}) \geq rs - k + 1$, and hence $\pi(\mathbf{v}) = \mathbf{0}$ implies $\mathbf{v} = \mathbf{0}$.

The minimum distance can be calculated by considering two cases: *(i)* $r < r'$ and $s = s'$, *(ii)* $r = r'$ and $s < s'$ (the third case where both $s < s'$ and $r < r'$ follows by composing projections of the first two cases). In the first case, for instance, if $\mathbf{v}$ is a nonzero vector of $C$ then

$$\begin{aligned}\varpi_{r',s}(\pi(\mathbf{v})) &= \varpi_{r,s}(\mathbf{v}) - (m(v^{r'+1}) + \cdots + m(v^{r})) \\ &\geq rs - k + 1 - s(r - r') = r's - k + 1\end{aligned}$$

and it follows that $\pi(C)$ is MDS. An analogous reasoning takes care of the second case.

**Evaluation Codes**

A third explicit construction was presented in [101] (see also [96]) and it is an extension of the Reed-Solomon code presented in Chap. 1, which we briefly recall.

Let $\mathbb{P}(t)$ be the vector space of all polynomials in $z$ with coefficients in $\mathbb{F}_q$ that have degree at most $t$ (including the zero polynomial). Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be distinct elements of $\mathbb{F}_q$ and take $1 \leq t \leq n$. The associated Reed-Solomon code is the image of the linear map

$$\begin{aligned}\phi : \mathbb{P}(t) &\to \mathbb{F}_q^n \\ f(z) &\to (f(\alpha_1), \ldots, f(\alpha_n)).\end{aligned}$$

Since a nonzero $f \in \mathbb{P}(t)$ has at most $t$ roots amongst the $\alpha_i$'s, the vector $\phi(f)$ has at least $n - t$ nonzero coordinates. This remark proves both that $\phi$ is injective and that its image is an MDS code.

The idea now is to build MDS codes in NRT spaces by evaluating polynomials and their (hyper) derivatives at the points $\alpha_1, \ldots, \alpha_n$, taking into account that the degree of a polynomial is still an upper bound for the number of roots even when we count multiplicities. It will be useful to write polynomials in the form $f(z) = \sum_{i=0}^{t-1} f_i z^i$, where $t = \deg(f) + 1$. For such a polynomial $f(z) = \sum_{i=0}^{t-1} f_i z^i$ with

coefficients in $\mathbb{F}_q$, its $j$-th *hyperderivative* is

$$\partial^j f(z) = \sum_{i=0}^{t-1} \binom{i}{j} f_i z^{i-j},$$

where $\binom{i}{j} = 0$ if $j > i$. This alternative to the usual formal derivative of a polynomial was introduced by Hasse [47]; a crucial advantage over the formal derivative of a polynomial is that the hyperderivatives provide the coefficients of the Taylor expansion: given $\alpha \in \mathbb{F}_q$,

$$f(z) = \sum_{j=0}^{t-1} \partial^j f(\alpha)(z-\alpha)^j. \tag{4.4}$$

From this expression we conclude that $\alpha$ is a root of $f(z)$ of multiplicity $m$ *iff* $\partial^j f(\alpha) = 0$ for $j = 0, 1, \ldots, m-1$ and $\partial^m f(\alpha) \neq 0$ (see [72, Lemma 6.51]).

**Theorem 4.4 ([101])** *Suppose that $1 \leq r \leq q$ and fix $\alpha_1, \ldots, \alpha_r$ in $\mathbb{F}_q$. Assume that $t \leq rs$ and consider the linear mapping*

$$\begin{aligned} \phi : \mathbb{P}(t-1) &\to \mathbb{F}_q^{[r\times s]} \\ f(z) &\mapsto (\partial^{(s-j)} f(\alpha_i)). \end{aligned}$$

*Then the image of the map $\phi$ above is an MDS code.*

*Proof* In the course of this proof, $\mu_f(\alpha)$ is the multiplicity of $\alpha$ as a root of $f$, if $\alpha$ is indeed a root of this polynomial, and $\mu_f(\alpha) = 0$ if $f(\alpha) \neq 0$.

It is clear that the map $\phi$ is linear. If $f$ is a nonzero element of $\mathbb{P}(t-1)$ then $\phi(f) \neq 0$, since $\phi(f) = 0$ iff

$$\partial^j f(\alpha_i) = 0, \quad j = 0, \ldots, s-1, \quad 1 \leq i \leq r$$

and therefore each $\alpha_i$ is a root of $f$ with multiplicity equal or greater than $s$ . On the other hand,

$$rs > t-1 \geq \deg(f) \geq \sum_i \mu_f(\alpha_i) = rs,$$

a contradiction. Hence $\phi$ is injective and its image has dimension $t$.

The code $C$ which is the image of $\phi$ is an MDS code *iff*

$$\varpi_{\mathcal{R}}(\phi(f)) \geq rs - t + 1$$

for every nonzero $f \in \mathbb{P}(t-1)$. Given $f$, let

$$\mathbf{v}^{(i)} = (\partial^{s-1} f(\alpha_i), \partial^{s-2} f(\alpha_i), \ldots, \partial^1 f(\alpha_i), f(\alpha_i))$$

be the $i$-th row of $\phi(f)$. It follows that if $\mu_f(\alpha_i) \leq s$ then

$$\mathbf{v}^{(i)} = (\partial^{s-1} f(\alpha_i), \partial^{s-2} f(\alpha_i), \ldots, \partial^{\mu_f(\alpha_i)} f(\alpha_i), 0, \ldots, 0)$$

and therefore $m(\mathbf{v}^{(i)}) = s - \mu_f(\alpha_i)$. If we define $\mu_i = \mu_f(\alpha_i)$ when $\mu_f(\alpha_i) \leq s$ and $\mu_i = s$ when $\mu_f(\alpha_i) > s$, then we can express the weight of $\phi(f)$ as

$$\varpi_{\mathcal{R}}(\phi(f)) = rs - \sum_i \mu_i. \tag{4.5}$$

Hence, if $f \neq 0$ and $\varpi_{\mathcal{R}}(\phi(f)) \leq rs - t$ then

$$\deg(f) < t \leq \sum_i \mu_i \leq \deg(f),$$

contradiction. It follows that $\varpi_{\mathcal{R}}(\phi(f)) \geq rs-t+1$ for every nonzero $f \in \mathbb{P}(t-1)$. ■

## 4.4 NRT Triangular Form

In this section we present a standard form for generator matrices of NRT codes, the "NRT triangular form", which is an interesting application of the description of the group of linear isometries of $\mathbb{F}_q^{[r\times s]}$. However, this standard form for generator matrices in the NRT case does not bear the same importance of the canonical form for generator matrices (and codes) in the hierarchical case, since the former only provides sparse generator matrices whereas the latter essentially yields the whole coding theory over hierarchical posets.

The group of linear isometries of $\mathbb{F}_q^{[r\times s]}$ was presented in Example 2.21 in matrix form; we will look at it from another angle, closer to Theorem 2.19.

In order to describe the group $G_{\mathcal{R}}$ and also to write down generator matrices for NRT codes we must consider a total ordering of the indices. For instance, we may order the coordinates lexicographically via the bijection $(i, j) \in [r] \times [s] \mapsto r(j-1)+i \in [rs]$. If $\mathbf{v} \in \mathbb{F}_q^{[r\times s]}$ then $\mathbf{v} = (\mathbf{v}^{(1)}; \mathbf{v}^{(2)}; \ldots; \mathbf{v}^{(r)})$ where, in these new coordinates,

$$(\mathbf{v}_i, \mathbf{v}_{i+r}, \cdots, \mathbf{v}_{i+r(s-1)}).$$

It follows from the definition of $G_{\mathcal{R}}$ that its elements are block-diagonal matrices

$$A = \operatorname{diag}(A_1, A_2, \ldots, A_r)$$

where each matrix $A_i$ is invertible and upper triangular. Letting $\mathbf{T}_s(\mathbb{F}_q)$ denote the group of the $s \times s$ invertible upper triangular matrices with coefficients in $\mathbb{F}_q$, it is easy to see that the map that sends $A = \text{diag}(A_1, A_2, \ldots, A_r) \in G_{\mathcal{R}}$ to $(A_1, A_2, \ldots, A_r) \in (\mathbf{T}_s(\mathbb{F}_q))^r$ is a group isomorphism between $G_{\mathcal{R}}$ and $(\mathbf{T}_s(\mathbb{F}_q))^r$.

We have already seen that the group of automorphisms of the poset $\mathcal{R} = \mathcal{R}(r, s)$ is isomorphic to $S_r$. Using Theorem 2.19, we conclude that $GL_{\mathcal{R}}(\mathbb{F}_q^n)$ is isomorphic to the semidirect product $(\mathbf{T}_s(\mathbb{F}_q))^r \rtimes S_r$, where $S_r$ acts by permuting the components of $(\mathbf{T}_s(\mathbb{F}_q))^r$.

The action of a permutation $\sigma \in S_r$ on a vector $\mathbf{v} = (\mathbf{v}^{(1)}; \mathbf{v}^{(2)}; \ldots; \mathbf{v}^{(r)})$ is given by

$$\sigma \cdot (\mathbf{v}^{(1)}; \mathbf{v}^{(2)}; \ldots; \mathbf{v}^{(r)}) = (\mathbf{v}^{(\sigma(1))}; \ldots; \mathbf{v}^{(\sigma(r))})$$

and the action of an element $(T_1, \ldots, T_r) \in (\mathbf{T}_s(\mathbb{F}_q))^r$ on $\mathbf{v}$ is

$$(T_1, \ldots, T_r) \cdot (\mathbf{v}^{(1)}; \mathbf{v}^{(2)}; \ldots; \mathbf{v}^{(r)}) = (T_1\mathbf{v}^{(1)}; \ldots; T_r\mathbf{v}^{(r)}).$$

In particular, the group of linear isometries of the NRT space over one chain $\mathbb{F}_q^{[1\times s]}$ is the group $\mathbf{T}_s(\mathbb{F}_q)$. The chain poset is a hierarchical poset and therefore $\mathbf{T}_s(\mathbb{F}_q)$ acts transitively on NRT spheres (see Theorem 3.8; it is also easy to check it directly); hence, if $\varpi_{\mathcal{R}}(\mathbf{u}) = k > 0$ then there exists $T \in \mathbf{T}_s(\mathbb{F}_q)$ such that $T(\mathbf{u}) = \mathbf{e}_k$, where $\{\mathbf{e}_1, \ldots, \mathbf{e}_s\}$ is the canonical basis of $\mathbb{F}_q^s$.

**Definition 4.5** We will say that a $k \times s$ matrix $M$ is $\mathbf{T}_s(\mathbb{F}_q)$-reduced if, modulo permutations of rows,

$$M = [\mathbf{c}_1|\mathbf{c}_2|\cdots|\mathbf{c}_s]$$

where each *column* $\mathbf{c}_j$ is either the zero vector or a vector of the form

$$\mathbf{c}_j = (c_{(1,j)}, \ldots, c_{(\omega_j-1,j)}, 1, 0, \ldots, 0)^T$$

with $\omega_j < \omega_{j'}$ whenever $j < j'$ and $\mathbf{c}_j$ and $\mathbf{c}_{j'}$ are both nonzero.

**Lemma 4.6** *Every $\mathbf{T}_s(\mathbb{F}_q)$-orbit of a $k\times s$ matrix contains a $\mathbf{T}_s(\mathbb{F}_q)$-reduced matrix modulo row permutation.*

The idea is the following: Let $M = (v_{i,j})$ be a nonzero $k \times s$ matrix and assume, for simplicity, that $M$ has a row with last entry $v_{i,s} \neq 0$. Permuting rows we may assume that $i = k$ (last row), and there is a symmetry $T_1 \in \mathbf{T}_s(\mathbb{F}_q)$ such that

$$T(v_{k,1}, \ldots, v_{k,s}) = \mathbf{e}_s.$$

If there is at least another nonzero row of $M$, we obtain

$$T_1(M) = \left[\begin{array}{ccc|ccc} \vdots & & \vdots & \vdots & \vdots & \vdots \\ v_{k-1,1} & \cdots & v_{k-1,j} & 0 & \cdots & 0 \quad v_{k-1,m} \\ \hline 0 & \cdots & 0 & 0 & \cdots & 0 \quad 1 \end{array}\right]$$

and modulo a permutation of rows we may assume that $v_{k-1,j} \neq 0$. Once more, there exists $T_2 \in \mathbf{T}_s(\mathbb{F}_q)$ that takes the $(k-1)$-th row to the vector $\mathbf{e}_j$. Therefore

$$T_2(T_1(M)) = \left[\begin{array}{ccc|ccc|ccc} \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ v_{k-2,1} & \cdots & v_{k-2,j'} & 0 & \cdots & v_{k-2,j} & 0 & \cdots 0 & v_{k-2,m} \\ \hline 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots 0 & v_{k-1,m} \\ \hline 0 & \cdots & 0 & 0 & \cdots & 0 & \cdot 0 & \cdot \quad 0 & 1 \end{array}\right]$$

and so on.

Let $G$ be a generator matrix of a linear code $C$ in $\mathbb{F}_q^{[r\times s]}$. Let us write

$$G = [G_1 \mid G_2 \mid \ldots \mid G_r]$$

where each $G_i$ is a $k \times s$-block and corresponds to the positions which lie in the $i$-th chain.

We describe in the following an algorithm for obtaining a new generator matrix, for a code equivalent to $C$, which is in a "block echelon form" called *NRT triangular* form in [2].

**First Step**

Modulo permutations (action of $S_r$) we may assume that $G_1$ has the least rank among the blocks $G_i$. An appropriate $T_1 \in \mathbf{T}_s(\mathbb{F}_q)$ transforms $G_1$ to a $\mathbf{T}_s(\mathbb{F}_q)$-reduced block $T_1(G_1)$, and

$$(T_1, I, \ldots, I)[G_1 \mid G_2 \mid \cdots \mid G_r] = [T_1(G_1) \mid G_2 \mid \cdots \mid G_r].$$

Combinations of rows of $G$ erase any nonzero column entry in $T_1(G_1)$ which is not the last nonzero one, and we end with a matrix

$$G' = [G'_1 \mid G'_2 \mid \cdots \mid G'_r]$$

which generates a code equivalent to $C$, whose first block $G'_1$ consists of rows which are either zero or have only one nonzero entry, equal to 1.

**Second Step**

Modulo row permutations, $G'_2$ is the block of least rank amongst $G'_2, \ldots, G'_r$. There exists $T_2 \in \mathbf{T}_s(\mathbb{F}_q)$ taking $G'_2$ to a $\mathbf{T}_s(\mathbb{F}_q)$-reduced block $T_2(G'_2)$, and

$$(I, T_2, I, \ldots, I)[G'_1 \mid G'_2 \mid \cdots \mid G'_r] = [G'_1 \mid T_2(G'_2) \mid \cdots \mid G'_r].$$

Now there are two possibilities: If $\text{rank}(G'_2) = k_2 > k_1 = \text{rank}(G'_1)$ then we may erase any nonzero column entry in $T_2(G'_2)$ which is not the last nonzero one via linear combinations, without disturbing the first block, and then we move on to the third block; if $\text{rank}(G'_2) = k_2 = k_1 = \text{rank}(G'_1)$ we move on to the third block. Hence we end with a matrix

$$G'' = [G''_1 \mid G''_2 \mid \cdots \mid G''_r]$$

which generates a code equivalent to $C$ and whose first two blocks are already simplified.

**Next $r-2$ Steps**
If $r = 2$ we are done; if not, once again, modulo row permutations, we may assume that $G''_3$ is the block of least rank amongst $G''_3, \ldots, G''_r$. Now we proceed as before, applying the same considerations of the second step to the pair $(G''_2, G''_3)$ in place of the pair $(G'_1, G'_2)$.

**Theorem 4.7 ([2])** *Let $\mathbb{F}_q^{[r\times s]}$ be the NRT space over the union of $r$ chains of length $s$, and let $C$ be a linear $[rs, k]$-code. There exists a generator matrix $G$ for a code equivalent to $C$ such that $G = [G_1|G_2|\cdots|G_r]$, where each $G_i$ is a $k \times s$ matrix, satisfying:*

1. *$G$ is in "block echelon form", i.e., if the last $t$ rows of $G_i$ are zero, then the last $t$ rows of $G_1, \ldots, G_{i-1}$ are also zero;*
2. *The nonzero rows of $G_1$ are distinct canonical vectors, arranged in order of increasing NRT-weight;*
3. *For each $i = 2, \ldots, r$,*
   (a) *$G_i$ is $\mathbf{T}_s(\mathbb{F}_q)$-reduced, or*
   (b) *$G_i = \begin{bmatrix} A^i & B^i \\ J^i & 0 \end{bmatrix}$, where $A^i$ and $B^i$ are $\mathbf{T}_s(\mathbb{F}_q)$-reduced, $J^i$ is a matrix whose nonzero rows are distinct canonical vectors (also arranged in order of increasing NRT-weight) and whose last column is nonzero, and all entries of $A^i$ above each nonzero entry of $J^i$ are zero.*

## 4.5 Ordered Symmetric Channel and the NRT Metric

Already in [96] Rosenbloom and Tsfasman presented a qualitative information-theoretic interpretation as a motivation for the NRT metric. Following [5], in this section we present a family of channels associated to this metric. If there is more than one chain, those channels are not matched to the NRT-weight as defined before (see Chap. 1); the error probability will depend not only on the weight but also on the $\text{Aut}(\mathcal{R})$-orbit of the support ideal and therefore, as we have just seen, it will depend on *shapes*.

Given an integer $r \geq 1$ consider a vector $\varepsilon = (\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_s)$ where $0 \leq \varepsilon_i \leq 1$ for all $i = 0, 1, \ldots, s$ and $\sum_{i=0}^{s} \varepsilon_i = 1$. Our *alphabet* will be the vector space $\mathbb{F}_q^s$. Let $P_s : \mathbb{F}_q^s \to \mathbb{F}_q^s$ define the transition matrix

$$P_s(\mathbf{y} \mid \mathbf{x}) = \begin{cases} \varepsilon_0 & \text{if } \mathbf{y} = \mathbf{x}, \\ \dfrac{\varepsilon_i}{q^{i-1}(q-1)} & \text{if } d_{\mathcal{R}}(\mathbf{x}, \mathbf{y}) = i. \end{cases} \tag{4.6}$$

Note that $q^{i-1}(q-1)$ is the number of elements of $\mathbf{y}$ such that $d_{\mathcal{R}}(\mathbf{x}, \mathbf{y}) = i$ in the NRT space when it has a unique chain of length $s$.

Assume also that $\varepsilon$ is chosen in such a manner that

$$\varepsilon_0 > \frac{\varepsilon_1}{q-1} > \frac{\varepsilon_2}{(q-1)q} > \ldots > \frac{\varepsilon_s}{q^{s-1}(q-1)}.$$

In this fashion the error probability decreases with the NRT distance between symbols of the alphabet $\mathbb{F}_q^s$: If $d_{\mathcal{R}}(\mathbf{y}, \mathbf{x}) \geq d_{\mathcal{R}}(\mathbf{z}, \mathbf{x})$ then $P_s(\mathbf{y} \mid \mathbf{x}) \leq P_s(\mathbf{z} \mid \mathbf{x})$.

Following the same steps of Chap. 1 for the symmetric channel, consider a memoryless channel on $\mathbb{F}_q^{[r\times s]} = (\mathbb{F}_q^s)^r$ by defining

$$P_s(\mathbf{y} \mid \mathbf{x}) = \prod_{i=1}^{r} P_s(\mathbf{x}^{(i)} \mid \mathbf{y}^{(i)})$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^{[r\times s]}$.

Since the NRT metric is translation invariant, so is the transition matrix, in the sense that $P_s(\mathbf{y} \mid \mathbf{x}) = P_s(\mathbf{y} + \mathbf{z} \mid \mathbf{x} + \mathbf{z})$ for any $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in $\mathbb{F}_q^{[r\times s]}$. Therefore we may assume, without loss of generality, that $\mathbf{x} = \mathbf{0}$ was sent. Let $\mathbf{y} \in \mathbb{F}_q^{[r\times s]}$ be a nonzero vector of shape $\mathbf{e} = (e_1, e_2, \ldots, e_s)$. Taking into account the shape formula for weights (4.3) we obtain the following formula for $P_s(\mathbf{y} \mid \mathbf{0})$ (see Exercise 5):

$$P_s(\mathbf{y} \mid \mathbf{0}) = \frac{\varepsilon_0^{e_0}}{q^{\varpi_{\mathcal{R}}(\mathbf{y})}} \prod_{i=1}^{s} \left(\frac{q\varepsilon_i}{q-1}\right)^{e_i}. \tag{4.7}$$

Therefore $P_s$ depends not only on the NRT-weight of $\mathbf{y}$ but also on its shape vector, and it is not apparent that a increase in weight corresponds to a decrease in the error probability. On the other hand, since the weight of a vector is a function of its shape, we should focus on the dependency on *shape* instead of *weight*. Another clue comes from the previous section: There it was shown that the $GL_{\mathcal{R}}(\mathbb{F}_q^{[r\times s]})$-orbits are parametrized by shapes, and that spheres are unions of orbits (and not a single orbit) if there is more than one chain. Therefore, one should look for a ordering of shapes.

Such an order of shapes is defined in [5]: given shapes $\mathbf{e} = (e_1, e_2, \ldots, e_s)$ and $\mathbf{f} = (f_1, f_2, \ldots, f_s)$ we say that $e \leq f$ if

$$e_l + e_{l+1} + \cdots + e_s \leq f_l + f_{l+1} + \cdots + f_s$$

for $l = 1, 2, \ldots, s$. If $r \geq 2$ this is not a total order. The meaning of this order is the following: Let $\overline{\mathcal{I}}(\mathcal{R})$ be the set of $\mathrm{Aut}(\mathcal{R})$-orbits of ideals of $\mathcal{R}$, and let $[I]$ denote the orbit of $I$. Let us define on $\overline{\mathcal{I}}(\mathcal{R})$ the partial order

$$[I] \leq [J] \textit{ iff } \text{there exists } I' \in [I] \text{ such that } I' \subseteq J. \tag{4.8}$$

Then it is proved in [5, Lemma 2.4] that

$$[I] \leq [J] \textit{ iff } \mathrm{shape}(I) \leq \mathrm{shape}(J). \tag{4.9}$$

Finally, let $\mathbf{y}$ and $\mathbf{z}$ be nonzero vectors of $\mathbb{F}_q^{[r\times s]}$ of shapes $\mathbf{e}$ and $\mathbf{f}$ respectively. According to [5] it follows from Eqs. (4.7) and (4.9) that if $\mathbf{e} \geq \mathbf{f}$ then $P_s(\mathbf{y} \mid \mathbf{0}) \leq P_s(\mathbf{z} \mid \mathbf{0})$, i.e., the probability of error is decreasing with respect to the order of shapes. Hence the NRT metric is matched to the ordered channel in the single chain case but it is only weakly matched in the multiple chains case.

## 4.6 Chapter Notes

Working with several chains presents serious technical difficulties, such as the "decoupling" of packing radius and minimum distance. On the other hand, although computations are much more complicated than in the hierarchical case (considering only codes in canonical form), one can still carry then through: since prime ideals either coincide or have empty intersection, and since the number of elements in a prime ideal depends only on the height of the maximal element, the double indices (where the first one localizes the chain, and the second one the height in this chain) provide a simple and effective system of coordinates to work with.

The important topics of uniform distributions in unit cubes, ordered orthogonal arrays and $(t, m, s)$-nets, which overlap with the research on NRT codes and have provided much of the impetus for the development of this research, have not been adressed in this book.

## 4.7 Exercises

1. Let $\mathcal{R}$ be the poset consisting of $r$ chains of lenght $s$. Let $\mathbf{c}$ be a nonzero vector and let $\mathcal{C} = \{\mathbf{0}, \mathbf{c}\}$ be a code in $\mathbb{F}_q^{[r\times s]}$ consisting of two codewords, so that $\delta_{d_\mathcal{R}}(C) = \varpi_\mathcal{R}(\mathbf{c})$.

(a) Assume that $\varpi_{\mathcal{R}}(\mathbf{c}) = s$, with $s \leq r$. Show that, depending on $\mathbf{c}$, the packing radius $R_{d_{\mathcal{R}}}(\mathcal{C})$ may assume any integer value between $\lfloor \frac{\varpi_{\mathcal{R}}(\mathbf{c})-1}{2} \rfloor$ and $\varpi_{\mathcal{R}}(\mathbf{c}) - 1$.
(b) Assume that $\varpi_{\mathcal{R}}(\mathbf{c}) = s+1$, with $s+1 \leq r$. Show that, depending on $\mathbf{c}$, the packing radius $R_{d_{\mathcal{R}}}(\mathcal{C})$ may assume any integer value between $\lfloor \frac{\varpi_{\mathcal{R}}(\mathbf{c})-1}{2} \rfloor$ and $\varpi_{\mathcal{R}}(\mathbf{c}) - 2$.
(c) Assume that $\varpi_{\mathcal{R}}(\mathbf{c}) = s+2$, with $s+2 \leq r$. Show that, depending on $\mathbf{c}$, the packing radius $R_{d_{\mathcal{R}}}(\mathcal{C})$ may assume any integer value between $\lfloor \frac{\varpi_{\mathcal{R}}(\mathbf{c})-1}{2} \rfloor$ and $\varpi_{\mathcal{R}}(\mathbf{c}) - 3$.
(d) Try to generalize it considering integers $a, b$ such that $as < \varpi_{\mathcal{R}}(\mathbf{c}) \leq (a+1)s$ and $br < \varpi_{\mathcal{R}}(\mathbf{c}) \leq (b+1)r$.

2. ([51, Proposition 1.1]) Let $P$ be a poset and let $I$ be a nontrivial ideal of $P$ (i.e. $P$ is neither empty nor the whole poset $P$). Prove that if $|I| = r$ then
   (a) If $r < r_1 \leq |P|$ then there exists an ideal $J$ of $P$ such that $I \subset J$ and $|J| = r_1$.
   (b) If $0 \leq r_0 < r$ then there exists an ideal $J$ of $P$ such that $J \subset I$ and $|J| = r_0$.
3. Show that every automorphism of $\mathcal{R}(r, s)$ is a permutation of the $r$ disjoint chains. **Hint:** *automorphisms preserve chains and preserve length, hence the image of an $s$-chain must be another $s$-chain.*
4. (Ordered channel) Considering the ordered channel defined in Section 4.5:
   (a) Show that $P_s(\mathbf{y} \mid \mathbf{x}) = P_s(\mathbf{y} + \mathbf{z} \mid \mathbf{x} + \mathbf{z})$ for any $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in $\mathbb{F}_q^{[r\times s]}$. Conclude that $P_s(\mathbf{y} \mid \mathbf{x}) = P_s(\mathbf{y} - \mathbf{x} \mid \mathbf{0})$.
   (b) Show that if $T : \mathbb{F}_q^{[r\times s]} \to \mathbb{F}_q^{[r\times s]}$ is a linear isometry and $\mathbf{y} \in \mathbb{F}_q^{[r\times s]}$ then $P_s(\mathbf{y} \mid \mathbf{0}) = P_s(T(\mathbf{y}) \mid \mathbf{0})$.
5. (Ordered channel) Consider the channel defined on $\mathbb{F}_q^{rs}$ by the transition matrix $P_s$ as in (4.6) and let $\mathbf{y} \in \mathbb{F}_q^{rs}$ with shape vector $e = e(\mathbf{y}) = (e_1, \ldots, e_s)$; remember that $|e| = e_1 + \cdots + e_s$ and that $e_0 = r - |e|$.
   (a) Show that $P_s(\mathbf{y} \mid \mathbf{0}) = (\varepsilon_0^{e_0} \ldots \varepsilon_s^{e_s}) \left(\frac{q-1}{q}\right)^{|e|} q^{\sum_{i=1}^{s} i e_i}$
   (b) Show that $P_s(\mathbf{y} \mid \mathbf{0}) = \dfrac{\varepsilon_0^{e_0}}{q^{\omega_{NRT}(y)}} \prod_{i=1}^{r} \left(\dfrac{q\varepsilon_i}{q-1}\right)^{e_i}$.
6. (Puncturing codes) Show that an $[rs, k]_q$ code $C$ is MDS with respect to the NRT metric if and only if the restriction of $\pi : \mathbb{F}_q^{[r\times s]} \to \mathbb{F}_q^{[r'\times s']}$ to $C$ is injective whenever $k < r's'$. **Remark**: *Other projections can be considered if you permute rows. And it seems that it is as far as it goes.*
7. Find a code in some NRT space which is not equivalent to a code which has a canonical decomposition in the sense of Chap. 3.
8. [51, Lemma 4.1] Let $P$ and $Q$ be posets with same underlying set $[n]$. We recall that $Q$ *refines* $P$ if for all $x, y \in [n]$, $x \preceq_P y$ implies $x \preceq_Q y$. Show that

if $C$ is a $P$-MDS code then $C$ is also a $Q$-MDS code. **Hint**: *this follows from the Singleton bound and from definition of MDS.*

9. MDS codes in poset spaces.

   (a) Let $P$ be a poset with $n$ elements. Use Exercise 8 and Reed-Solomon codes to show that if $q > n$ then there exists an MDS $[n, k]_q$ code in $(\mathbb{F}_q^n, d_P)$ for every positive integer $k$ such that $n \geq k$.

   (b) Let $P$ be the poset which is the disjoint union of $r$ chains of (possibly) distinct lengths $t_1, t_2, \ldots, t_r$ and let $n = t_1 + \cdots + t_r$ (which is the number of elements of $P$). Make an adaptation of the evaluation code in Sect. 4.3.2 in order to show that if $q > r$ then there exists an MDS $[n, k]_q$ code in $(\mathbb{F}_q^n, d_P)$ for every positive integer $k$ such that $n \geq k$.

   (c) It is possible to partition a poset $Q$ in a disjoint union of chains (disjoint union as a *set*, not as a poset). Dilworth's Theorem states that the minimum number of chains in such a partition of $Q$ equals the largest size of an antichain contained in $Q$.

   Let $Q$ be a poset of size $n$ and let $Q = Q_1 \cup Q_2 \cup \cdots \cup Q_r$ be a "Dilworth partition" of the set $Q$, each $Q_i$ a chain of length $t_i$. Use the Dilworth partition, item (a) and Exercise 8 to prove the following result:

   [52, Thm II.3] If $q > r$ then there exists an MDS $[n, k]_q$-code in $(\mathbb{F}_q^n, d_Q)$ for every positive integer $k$ such that $n \geq k$.

   (d) Compare the results of (a) and (c): what have we gained using the more intricate method of item (c)?

10. (Open problem) Find a perfect code in the NRT space $\mathbb{F}_q^{[r \times s]}$, where $r \geq 3$ and $s \geq 2$, which is not a preimage of a perfect code in the Hamming space $(\mathbb{F}_q^r, d_H)$.

# Chapter 5
# Duality


In a general setting, duality theorems relate invariants of a mathematical structure to the invariants of a dual structure. In lattice theory, they are called *transference theorems*, since they allow to transfer problems in the lattice to problems in the dual lattice, where they may be simpler to tackle. In coding theory, the two main results in this direction are the MacWilliams identities, which establishes relations between the weight distribution of a code and of its dual, and the duality theorem of generalized weights, initially proposed by Wei in [108], as a sort of weakened MacWilliams' identity for the subspace hierarchy of codes.

Several works are devoted to MacWilliams-type identities and Wei's duality theorem for poset codes, see [4, 13, 18, 32, 63, 64, 81, 85]. Furthermore, many combinatorial tools have been used in the literature in order to obtain these identities, such as association schemes [85], group characters [64], matroids [13], and multisets [81]. In this chapter, we are going to explore both the results, using two different techniques: the MacWilliams-type identities (using group characters) and the Wei duality theorem (using matroids).

## 5.1 Equivalence Relations Over Ideals

As seen in Chap. 3, for hierarchical poset metrics the MacWilliams identity can be obtained by using the canonical decomposition and the MacWilliams identity for the Hamming metric. Furthermore, Theorem 3.8 ensures that among the poset metrics, the hierarchical metrics are the only ones admitting the MacWilliams identity in the classical meaning, that is, an identity relating the weight distributions of a code and its dual. Many attempts have been made to derive a MacWilliams-type identity for the non-hierarchical case, either by changing same algebraic aspect of the dual code or by extending the concept of weight enumerator, as we can see in [32] and in [18]. In this chapter we aim to explore the approach given by Choi et al. in [18]



M. Firer et al., *Poset Codes: Partial Orders, Metrics and Coding Theory*, SpringerBriefs in Mathematics, https://doi.org/10.1007/978-3-319-93821-9_5

since it provides a unifying way to present the majority of the results obtained in this field by using the same tools that F. J. MacWilliams employed in [76] to obtain the MacWilliams identities for the Hamming metric.

In order to obtain MacWilliams' identities for non-hierarchical posets, as proposed in [18], we shall explore the concept of equivalence relations on the set $\mathcal{I}(P)$ of all ideals of $P$.

We recall that, given a poset $P$, the *dual (or opposite) poset* $P^*$ is defined by $i \preceq_{P^*} j$ if, and only if, $j \preceq_P i$, for every $i, j \in P$. Due to the fact that every ideal in the dual poset is the complementary of an ideal in the original poset, when considering $P^*$, we will use the notation $I^c$ to denote ideals of $\mathcal{I}(P^*)$. When $I, J \in \mathcal{I}(P)$ are isomorphic as posets, that is, when there exists an order preserving bijection $\sigma : I \to J$, we will denote by $I \sim_\sigma J$, or just by $I \sim J$ when the bijection is not relevant in the context. With this notation we state our first relation on $\mathcal{I}(P)$.

**($\mathbf{E_S}$) Isomorphism Relation** $E_S$ is the equivalence relation on $\mathcal{I}(P)$ given by:

$$(I, J) \in E_S \text{ if, and only if, } I \sim J.$$

It is an easy exercise to show that $E_S$ is indeed an equivalence relation.

If $E$ is an equivalence relation on $\mathcal{I}(P)$, then a natural attempt to induce an equivalence relation $E'$ over $\mathcal{I}(P^*)$ would be putting $(I^c, J^c) \in E'$ if, and only if, $(I, J) \in E$. However, in this case, the characteristics of $E$ are not transferred to $E'_S$. Indeed, considering the isomorphism relation previously defined, it is clear that for a hierarchical poset, $E_S$ admits a dual relation $E'$ constructed in this way. However, in the general case, it does not work well, since $I \sim J$ not necessarily implies $I^c \sim J^c$. Hence, we need to modify the definition of dual relation so that the properties of the relation are transfered to the dual structure.

**Definition 5.1** Let $P$ be a poset over $[n]$ and $E$ an equivalence relation on $\mathcal{I}(P)$. An equivalence relation $E^*$ on $\mathcal{I}(P^*)$ is the *dual relation of* $E$ if it satisfies the following property: if $(I, J) \in E$ is defined by property $(A)$ on $\mathcal{I}(P)$, then $(I^c, J^c) \in E^*$ is defined by property $(A)$ on $\mathcal{I}(P^*)$.

It is clear that the dual relation is unique and well-defined since $E^{**} = E$. The *price we are paying* to use $E^*$ instead of the natural one $E'$ is that some equivalence relations over $\mathcal{I}(P)$ may not have a dual relation. Indeed, the isomorphism relation $E_S$ may or may not admit a dual relation $E_S^*$, depending on the poset structure. The family of posets in which $E_S$ always has a dual relation is larger than the hierarchical family and, based on the previous discussion, can be easily characterized.

**Definition 5.2** A poset $P$ is called a *complement isomorphism poset* if for any $I, J \in \mathcal{I}(P)$ we have that

$$I \sim J \text{ if, and only if, } I^c \sim J^c.$$

We will let the reader to prove as an exercise (Exercise 1), that $E_S$ admits a dual relation only in the case of complement isomorphism posets. We shall now present two more equivalence relations that will play an important role in the sequence.

**($\mathbf{E_C}$) Cardinality Relation** $E_C$ is the relation on $\mathcal{I}(P)$ defined by

$$(I, J) \in E_C \text{ if, and only if, } |I| = |J|,$$

where $I, J \in \mathcal{I}(P)$. $E_C$ is an equivalence relation with dual relation $E_C^*$ on $\mathcal{I}(P^*)$ defined by

$$(I^c, J^c) \in E_C^* \text{ if, and only if, } |I^c| = |J^c|.$$

**($\mathbf{E_H}$) Automorphism Relation** Let $H$ be a subgroup of $Aut(P)$. The relation $E_H$ on $\mathcal{I}(P)$ is defined by

$$(I, J) \in E_H \text{ if, and only if, } \sigma(I) = J \text{ for some } \sigma \in H$$

is an equivalence relation whose dual relation is naturally defined by the rule $\sigma(I^c) = J^c$.

Regarding the three relations previously defined, we have that $E_H \subset E_S \subset E_C$ (see Exercise 3). In the subsequent sections, we will show that the amount of relations is inversely proportional to the number of posets in which we can find a MacWilliams-type identity.

## 5.2 $I$-Spheres and MacWilliams' Equivalences

It is well-known that, in general, the weight enumerator of a code and of its dual cannot be related: in the proof of Theorem 3.8 (classification of hierarchical posets by the MacWilliams identity), we provide an example of two codes with same weight distribution whose dual codes have different weight distributions. Hence, in order to obtain MacWilliams-type identities, distributions of invariants of codes which in some sense generalize the weight distribution should be explored. To do so, we will use the equivalence relations defined in the previous section as the main tool to work with ideals and to determine the MacWilliams-type identities.

When considering the classical MacWilliams' identity, we partition a code into subsets whose elements are codewords with same weight. It is worth to note that in addition to determine the weight distribution of a codeword, the ideal generated by its support may provide more information about the codeword: considering a binary one-dimensional code, the weight of its unique non-null codeword does not determine its packing radius (see Exercise 1 of Chap. 4 and Sects. 4.2.2 and 6.1), but considering two binary one-dimensional codes $C_1 = \{\mathbf{0}, \mathbf{c}_1\}$ and $C_2 = \{\mathbf{0}, \mathbf{c}_2\}$ for which the ideals $\langle \text{supp}(\mathbf{c}_1) \rangle$ and $\langle \text{supp}(\mathbf{c}_2) \rangle$ are isomorphic, they do have the same

packing radius. Hence, a natural extension of the partition of a code determined by the weight is a partition of the code into subsets determined by an equivalence relation on ideals, for instance, being isomorphic is an example of such relations. To construct such a partition, we consider an equivalence relation on $\mathcal{I}(P)$ and denote by $\overline{I} \in \mathcal{I}(P)/E$ the equivalence class of an ideal $I$. We shall define the concept of $\overline{I}$-spheres, which is derived from the definition of $I$-spheres, introduced in [51].

Let $P$ be a poset and $I \in \mathcal{I}(P)$. For every $\mathbf{u} \in \mathbb{F}_q^n$, define the $I$-sphere centered at $\mathbf{u}$ as

$$S_I^P(\mathbf{u}) := \{\mathbf{v} \in \mathbb{F}_q^n \ ; \ \langle supp(\mathbf{u}-\mathbf{v})\rangle_P = I\}.$$

Given an equivalence relation $E$ over $\mathcal{I}(P)$, for each coset $\overline{I} \in \mathcal{I}(P)/E$, the $\overline{I}$-sphere $S_{\overline{I},E}^P(\mathbf{u})$ centered at $\mathbf{u}$ with respect to $E$ is defined as

$$S_{\overline{I},E}^P(\mathbf{u}) := \{\mathbf{v} \in \mathbb{F}_q^n \ : \ (\langle supp(\mathbf{u}-\mathbf{v})\rangle_P, I) \in E\}.$$

The definition of $\overline{I}$-sphere is the natural extension of poset metric spheres in the following sense: considering the equivalence relation $E_C$, if $P$ is a hierarchical poset then

$$S_{\overline{I},E_C}^P(\mathbf{u}) = S_{d_P}(\mathbf{u}, |I|), \tag{5.1}$$

where $S_{d_P}(\mathbf{u}, |I|)$ is the *poset metric sphere* centered in $\mathbf{u}$ with radius $|I|$, that is, $S_{d_P}(\mathbf{u}, |I|) := \{\mathbf{v} \in \mathbb{F}_q^n : d_P(\mathbf{u}, \mathbf{v}) = |I|\}$. In other words, when considering hierarchical posets, the partition of the codewords provided by $E_C$ coincides with the partition provided by the weight distribution and, roughly speaking, the ideal generated by the support of a codeword does not provide more information on the codeword excepts for its weight: we saw in Theorem 3.8 that, in the hierarchical case, the minimum distance of a code determines its packing radius.

Let $C$ be a linear code in $\mathbb{F}_q^n$. For $\overline{I} \in \mathcal{I}(P)/E$ and $J \in \mathcal{I}(P)$, we denote $S_{\overline{I},E}^P(\mathbf{0}) = S_{\overline{I},E}^P$ and $S_J^P(\mathbf{0}) = S_J^P$ and define

$$A_{\overline{I},E}(C) := |S_{\overline{I},E}^P \cap C|. \tag{5.2}$$

**Definition 5.3** Let $P$ be a poset, $C$ be a linear code over $\mathbb{F}_q^n$ and $E$ be an equivalence relation on $\mathcal{I}(P)$. The vector

$$W(C, P, E) := [A_{\overline{I},E}(C)]_{\overline{I} \in \mathcal{I}(P)/E}$$

is the *spectrum* of $C$ with respect to $E$.

In the setting of hierarchical posets with cardinality relation $E_C$, we get that the spectrum of $C$ with respect to $E_C$ coincides with the weight distribution of $C$, i.e.,

$$A_{\overline{I},E_C}(C) = |\{\mathbf{c} \in C; \varpi_{d_P}(\mathbf{c}) = |I|\}|. \tag{5.3}$$

From here on, our goal in this section is to characterize the posets for which the spectrum of any code is uniquely determined by the spectrum of its dual, that is, the posets (or poset metrics) admitting a MacWilliams-type identity according to a equivalence relation over $\mathcal{I}(P)$.

**Definition 5.4** Let $P$ be a poset and $E$ an equivalence relation on $\mathcal{I}(P)$ admitting a dual relation. The relation $E$ is a *MacWilliams equivalence* if, given linear codes $C_1$ and $C_2$, $W(C_1, P, E) = W(C_2, P, E)$ implies $W(C_1^{\perp}, P^*, E^*) = W(C_2^{\perp}, P^*, E^*)$.

By the previous discussion and the characterization of hierarchical poset according to the existence of a MacWilliams-type identity presented in Chap. 3 (whose proof is outlined in Exercise 4 of that chapter), we conclude that when considering hierarchical posets, the equivalence relation $E_C$ is a MacWilliams equivalence. However, the proof presented in Chap. 3 uses the canonical decomposition and the existence of the MacWilliams identity for the Hamming metric. In the subsequent sections of this chapter we will provide a direct proof for this fact, so that the reader can grasp something of the power (and difficulties) of using characters. Furthermore, we will characterize the posets admitting a MacWilliams-type identity for $E_S$ and $E_H$ where $H$ is a subgroup of $Aut(P)$.

Before we continue, we shall describe some general results concerning equivalence relations over $\mathcal{I}(P)$. It is worth to note that the results are rather technical and we will try to avoid part of this technicality, either by referring to the original work [18] as a supplementary reading or by using exercises as a guideline to verify some statements.

From here on assume we only consider equivalence relations which admit a dual relation.

### 5.2.1 MacWilliams-Type Identities: A Summary of Tools

The MacWilliams identity can be proved using either characters or association schemes. We have chosen to use most classical approach, using characters, so that we need mainly two tools, the orthogonality relation and the Möbius Inversion Formula. We briefly introduce our tool kit.

#### (*i*) Additive Characters over $\mathbb{F}_q$

An *additive character* $\chi$ of a finite field $\mathbb{F}_q$ is a homomorphism from the additive group of $\mathbb{F}_q$ into the multiplicative group of unitary complex numbers, that is, $\chi : \mathbb{F}_q \to \mathbb{C}^*$ satisfies $\chi(\alpha + \beta) = \chi(\alpha)\chi(\beta)$ and $|\chi(\alpha)| = 1$ for every $\alpha, \beta \in \mathbb{F}_q$. A character $\chi$ is *trivial* if $\chi(\alpha) = 1$ for every $\alpha \in \mathbb{F}_q$. From here on, we consider only

non-trivial characters. The main relation on characters we will use is the following (a proof can be found in [71]): For any non-trivial character $\chi$,

$$\sum_{\beta\in\mathbb{F}_q}\chi(\alpha\beta)=\begin{cases}q & \text{if } \alpha=0\\ 0 & \text{if } \alpha\neq 0\end{cases}\qquad \text{(orthogonality relation)}$$

**(*ii*) Möbius Inversion Formula**

Let $\mathcal{P}=(\mathcal{P}([n]),\subseteq)$ be the poset of all subsets of $[n]$. Given functions $f:\mathcal{P}([n])\to\mathbb{C}$ and $g:\mathcal{P}([n])\to\mathbb{C}$, then for any $E,F\subset[n]$, it is known (see [103]) that

$$f(E)=\sum_{A\subset E}g(A) \text{ if, and only if, } g(F)=\sum_{B\subset F}(-1)^{|F\setminus B|}f(B).$$

Therefore,

$$f(E)=\sum_{A\subset E}\sum_{B\subset A}(-1)^{|A\setminus B|}f(B). \tag{5.4}$$

**Orthogonality Relation of Characters: Derived Identities**

Since the steps used to obtain some identities are rather technical, we will omit their proofs. We invite the reader to prove Identity (5.5) by following the instructions of Exercise 6. For a complementary reading and complete proof of these identities, see [18].

(a) Given $\overline{I}\in\mathcal{I}(P)/E$ and $C$ a linear code over $\mathbb{F}_q^n$, then

$$A_{\overline{I},E}(C)=\frac{1}{|C^{\perp}|}\sum_{\overline{J^c}\in\mathcal{I}(P^*)/E^*}\ \sum_{\mathbf{u}\in C^{\perp}\cap S^{P^*}_{\overline{J^c},E^*}}\ \sum_{\mathbf{v}\in S^{P}_{\overline{I},E}}\chi(\mathbf{u}\cdot\mathbf{v}) \tag{5.5}$$

where $\mathbf{u}\cdot\mathbf{v}=\sum_{i=1}^{n}u_iv_i$ and $\chi$ is a non-trivial additive character.

(b) ([18], Lemma 3.7) Denote by $M(I)$ and $m(I)$ the *set of maximal* and *non-maximal* elements of $I$, respectively. Let $I\in\mathcal{I}(P)$, $J^c\in\mathcal{I}(P^*)$ and $\mathbf{u}\in S_I^P$. Then

$$\sum_{\mathbf{v}\in S^{P^*}_{J^c}}\chi(\mathbf{u}\cdot\mathbf{v})=\begin{cases}(-1)^{|I\cap J^c|}(q-1)^{|M(J^c)|-|I\cap J^c|}q^{|m(J^c)|} & \text{if } m(I)\cap J^c=\emptyset\\ 0 & \text{if } m(I)\cap J^c\neq\emptyset\end{cases}. \tag{5.6}$$

We stress that, by duality, a similar identity (dual identity of the previous ones) occurs for $A_{\overline{J^c},E^*}(\mathcal{C}^\perp)$ and for $\sum_{\mathbf{v}\in S_I^P}\chi(\mathbf{u}\cdot\mathbf{v})$ when $\mathbf{u}\in S^{P^*}_{J^c}$.

### 5.2.2 *Alternative Formulation of MacWilliams' Equivalence*

The following conditions, which together are equivalent to the existence of a MacWilliams type relation, will be used for the characterization of MacWilliams-type identities to come in the sequence of this section.

**Proposition 5.5 ([18, Theorem 2.3])** *An equivalence relation $E$ on $\mathcal{I}(P)$ admitting a dual relation is a MacWilliams equivalence if, and only if, for every $\overline{I}\in\mathcal{I}(P)/E$ and $\overline{J^c}\in\mathcal{I}(P^*)/E^*$ the two following items are satisfied:*

*(i) If $\mathbf{u}$ and $\mathbf{u}'$ are in $S^P_{\overline{I},E}$, then $\sum_{\mathbf{v}\in S^{P^*}_{\overline{J^c},E^*}}\chi(\mathbf{u}\cdot\mathbf{v})=\sum_{\mathbf{v}\in S^{P^*}_{\overline{J^c},E^*}}\chi(\mathbf{u}'\cdot\mathbf{v})$;*

*(ii) If $\mathbf{v}$ and $\mathbf{v}'$ are in $S^{P^*}_{\overline{J}^c,E^*}$, then $\sum_{\mathbf{u}\in S^P_{\overline{I},E}}\chi(\mathbf{u}\cdot\mathbf{v})=\sum_{\mathbf{u}\in S^P_{\overline{I},E}}\chi(\mathbf{u}\cdot\mathbf{v}')$.*

*Proof* Let $E$ be a MacWilliams equivalence and suppose that $(i)$ is not valid for some $\overline{I_0}\in\mathcal{I}(P)/E$. This means there are $\mathbf{u},\mathbf{u}'\in S^P_{\overline{I_0},E}$ not satisfying condition $(i)$. Let $C_1=span\{\mathbf{u}\}$ and $C_2=span\{\mathbf{u}'\}$ be two 1-dimensional codes. By construction, $W(C_1,P,E)=W(C_2,P,E)$. On the other hand, considering the dual of Identity (5.5) and the fact that $C_1\cap S^P_{\overline{I_0},E}=C_1\setminus\{\mathbf{0}\}$, $C_1\cap S^P_{\overline{\emptyset},E}=\{\mathbf{0}\}$ and $C_1\cap S^P_{\overline{I},E}=\emptyset$ for every coset $\overline{I}\neq\overline{I_0}$ and $\overline{I}\neq\overline{\emptyset}$, we get a contradiction since

$$A_{\overline{J^c},E^*}(C_1^\perp)=\frac{1}{|C_1|}\sum_{\overline{I}\in\mathcal{I}(P)/E}\sum_{\mathbf{w}\in C_1\cap S^P_{\overline{I},E}}\sum_{\mathbf{v}\in S^{P^*}_{\overline{J^c},E^*}}\chi(\mathbf{w}\cdot\mathbf{v}) \tag{5.7}$$

$$=\frac{1}{q}\left(|S^{P^*}_{\overline{J^c},E^*}|+\sum_{\alpha\in\mathbb{F}_q\setminus\{0\}}\chi(\alpha)\sum_{\mathbf{v}\in S^{P^*}_{\overline{J^c},E^*}}\chi(\mathbf{u}\cdot\mathbf{v})\right)\neq A_{\overline{J^c},E^*}(C_2^\perp).$$

Note that we are also using the fact that $\mathbf{w}$ is running over all elements of the form $\alpha\mathbf{u}$ for every $\alpha\in\mathbb{F}_q\setminus\{0\}$. With the same reasoning, using Identity (5.5) instead of its dual identity (5.7), we obtain the proposition if we assume that $(ii)$ is not valid.

For the reciprocal, suppose that the relation $E$ satisfies conditions $(i)$ and $(ii)$ and consider two codes $C_1,C_2\in\mathbb{F}_q^n$ such that $W(C_1,P,E)=W(C_2,P,E)$. Since the last summation of identity (5.5) is constant by hypothesis (item $(i)$), Identity (5.2) ensures that $W(C_1^\perp,P^*,E^*)=W(C_2^\perp,P^*,E^*)$. ■

**Corollary 5.6** *Let $P$ be a poset, $C\subset\mathbb{F}_q^n$ a linear code, and $E$ a MacWilliams equivalence over $\mathcal{I}(P)$. Then, for every $\overline{I}\in\mathcal{I}(P)/E$,*

$$A_{\overline{I},E}(C) = \frac{1}{|C^{\perp}|} \sum_{\overline{J^c} \in \mathcal{I}(P^*)/E^*} A_{\overline{J^c},E^*}(C^{\perp}) t(\overline{I}, \mathbf{v}_{\overline{J^c}})$$

*where* $t(\overline{I}, \mathbf{v}_{\overline{J^c}}) := \sum_{\mathbf{u} \in S^P_{I,E}} \chi(\mathbf{u} \cdot \mathbf{v})$ *and* $\mathbf{v}_{\overline{J^c}} \in S^{P^*}_{\overline{J^c},E^*}$.

Consider the product $[A_{\overline{I}}]_{\overline{I} \in \mathcal{I}(P)/E} \cdot [B_{\overline{I}}]_{\overline{I} \in \mathcal{I}(P)/E} := \sum_{\overline{I} \in \mathcal{I}(P)/E} A_{\overline{I}} B_{\overline{I}}$ where $A_{\overline{I}}, B_{\overline{I}} \in \mathbb{Z}$ for every $\overline{I} \in \mathcal{I}(P)/E$. Therefore, by Corollary 5.6, we get the MacWilliams-type identity for MacWilliams' equivalences over $\mathcal{I}(P)$.

**Theorem 5.7** *Let $P$ be a poset, $C \subset \mathbb{F}_q^n$ a linear code and $E$ a MacWilliams equivalence over $\mathcal{I}(P)$. Then,*

$$W(C, P, E) = \frac{1}{|C^{\perp}|} \left[ W(C^{\perp}, P^*, E^*) \cdot [t(\overline{I}, \mathbf{v}_{\overline{J^c}})]_{\overline{J^c} \in \mathcal{I}(P^*)/E^*} \right]_{\overline{I} \in \mathcal{I}(P)/E}, \tag{5.8}$$

*where* $t(\overline{I}, \mathbf{v}_{\overline{J^c}}) = \sum_{\mathbf{u} \in S^P_{I,E}} \chi(\mathbf{u} \cdot \mathbf{v})$ *and* $\mathbf{v}_{\overline{J^c}} \in S^{P^*}_{\overline{J^c},E^*}$.

We stress that $t(\overline{I}, \mathbf{v}_{\overline{J^c}})$ does not depend on $C$ and, furthermore, it is the dual equation of (5.6), that is,

$$t(\overline{I}, \mathbf{v}_{\overline{J^c}}) = \begin{cases} (-1)^{|J^c \cap I|}(q-1)^{|M(I)|-|J^c \cap I|} q^{|m(I)|} & \text{if } m(J^c) \cap I = \emptyset \\ 0 & \text{if } m(J^c) \cap I \neq \emptyset \end{cases}.$$

### 5.2.3 Classifications of Posets

We end the discussion concerning MacWilliams' equivalences by classifying the posets turning each of $E_S$, $E_C$, and $E_H$ into a MacWilliams equivalence. This classification was presented in [18], but the classification concerning $E_C$ is known since 2005 due to the work [64]. An alternative way to prove this classification was presented in [74], using the canonical decomposition and by assuming the Hamming case. This provides a shorter proof of this case, which is delineated in Exercise 4 of Chap. 3.

In the remaining of this section consider $\mathbf{u}, \mathbf{u}' \in S^P_{\overline{I},E}$ and let $I_1 = \langle supp(\mathbf{u}) \rangle_P$ and $I_2 = \langle supp(\mathbf{u}') \rangle_P$. Note that $\mathbf{u} \in S^P_{I_1} \subset S^P_{\overline{I},E}$. Hence, from Exercise 4 in this chapter and Eq. (5.6), we have that for every $\overline{J^c} \in \mathcal{I}(P^*)/E^*$,

$$\sum_{\mathbf{v} \in S^{P^*}_{\overline{J^c},E^*}} \chi(\mathbf{u} \cdot \mathbf{v}) = \sum_{K^c \in \overline{J^c}} \sum_{\mathbf{v} \in S^{P^*}_{K^c}} \chi(\mathbf{u} \cdot \mathbf{v})$$

$$= q^{|m(J^c)|}(q-1)^{|M(J^c)|} \sum_{\substack{K^c \in \overline{J^c} \\ m(I_1)\cap K^c = \emptyset}} \left(\frac{1}{1-q}\right)^{|I_1 \cap K^c|}, \tag{5.9}$$

where in the last equality we are using the following (easy to prove) facts: $|M(K^c)| = |M(J^c)|$ and $|m(K^c)| = |m(J^c)|$. An analogous equation (dual to the one presented) is obtained when we consider $\mathbf{u}, \mathbf{u}' \in S^{P^*}_{\overline{J^c}, E^*}$.

**Theorem 5.8** *Let $P$ be a poset over $[n]$. $P$ is hierarchical if, and only if, $E_C$ is a MacWilliams equivalence.*

*Proof* Suppose $P$ is hierarchical, hence there exist $\sigma \in Aut(P)$ such that $\sigma(I_2) = I_1$. We shall prove condition $(i)$ of Proposition 5.5. By identity (5.9), for every $\overline{J^c} \in \mathcal{I}(P^*)/E^*$,

$$\sum_{\mathbf{v} \in S^{P^*}_{\overline{J^c}, E^*_C}} \chi(\mathbf{u} \cdot \mathbf{v}) = q^{|m(J^c)|}(q-1)^{|M(J^c)|} \sum_{\substack{\sigma(K^c) \in \overline{J^c} \\ m(\sigma(I_2)) \cap \sigma(K^c) = \emptyset}} \left(\frac{1}{1-q}\right)^{|\sigma(I_2) \cap \sigma(K^c)|}$$

$$= q^{|m(J^c)|}(q-1)^{|M(J^c)|} \sum_{\substack{\sigma(K^c) \in \overline{J^c} \\ \sigma(m(I_2)) \cap K^c = \emptyset}} \left(\frac{1}{1-q}\right)^{|\sigma(I_2 \cap K^c)|},$$

where the last equality can be justified by the following two facts: $\sigma(m(I_2)) = m(\sigma(I_2))$ and $\sigma(A \cap B) = \sigma(A) \cap \sigma(B)$. Furthermore, it is clear that the last equation coincides with $\sum_{\mathbf{v} \in S^{P^*}_{\overline{J^c}, E^*_C}} \chi(\mathbf{u}' \cdot \mathbf{v})$ because $\mathbf{u}' \in S^P_{I_2}$ and hence we can perform the same steps used to obtain identity (5.9). Condition $(ii)$ of Proposition 5.5 can be proved in a similar way. For the opposite direction, we saw in Eq. (5.3) that when considering the cardinality relation $E_C$ and hierarchical posets, the spectrum of a code coincides with the weight distribution. The proof is concluded by using the construction presented in Theorem 3.8 on the MacWilliams Identity item. ■

**Theorem 5.9** *If $P$ is a poset over $[n]$ and $H$ is a subgroup of $Aut(P)$, then $E_H$ is a MacWilliams equivalence on $\mathcal{I}(P)$.*

*Proof* Since here $\mathbf{u}, \mathbf{u}' \in S^P_{I, E_H}$, by the definition of $E_H$, we have that there is an element $\sigma \in H$ such that $\sigma(I_1) = I_2$. The proof follows in a similar way to the proof of Theorem 5.8. ■

Note that the main characteristic used to prove the last two theorems was the existence of an automorphism $\sigma \in Aut(P)$ such that $\sigma(I_1) = I_2$. When considering complement isomorphism posets, the existence (or not) of such automorphism is still an open problem, however we can work around this problem by using the Möbius inversion formula.

**Theorem 5.10** *Let $P$ be a poset over $[n]$. $P$ is a complement isomorphism poset if, and only if, $E_S$ is a MacWilliams equivalence.*

*Proof* First, note that for every $\overline{J^c} \in \mathcal{I}(P^*)/E^*$,

$$\sum_{\substack{K^c \in \overline{J^c} \\ I_1 \cap K^c \subset M(I_1)}} \left(\frac{-1}{q-1}\right)^{|I_1 \cap K^c|} = \sum_{A \subset M(I_1)} \left(\frac{-1}{q-1}\right)^{|A|} \sum_{\substack{K^c \in \overline{J^c} \\ I_1 \cap K^c = A}} 1.$$

Applying the Möbius Inversion Formula (Eq. (5.4)) to the right-hand side of the previous equality, we get

$$\sum_{\substack{K^c \in \overline{J^c} \\ I_1 \cap K^c \subset M(I_1)}} \left(\frac{-1}{q-1}\right)^{|I_1 \cap K^c|} = \sum_{A \subset M(I_1)} \left(\frac{-1}{q-1}\right)^{|A|} \sum_{B \subset A} (-1)^{|A \setminus B|} \sum_{\substack{K^c \in \overline{J^c} \\ I_1 \cap K^c \subset B}} 1.$$

Suppose $I_1 \sim I_2$, so that there is a bijective order preserving map $\sigma : I_1 \to I_2$. Therefore, by Eq. (5.9) and by the fact that $m(I_1) \cap K^c = \emptyset$ if, and only if, $I_1 \cap K^c \subset M(I_1)$, we conclude the proof of item (a) of Proposition 5.5. Item (b) can be proved in a similar way. ■

## 5.3 Matroids and Poset Duality

We will use matroids theory to prove Wei's duality theorem, so we start by introducing some basic matroids' concepts.

### *5.3.1 Matroids*

Proposed by Whitney in ([109], 1935), matroid theory has been attracting increasing attention since the work of Tutte concerning the characterization of matroids arising from graphs (see [105]).

Since a matroid is a combinatorial structure which abstracts the notion of linear independence in vector spaces, some results from coding theory were extended and proved in this broader context. In this section we aim to describe how to obtain the Wei duality theorem for poset metrics using matroids. The approach used in this book to obtain the Wei duality theorem is less general than the one of [13], where a duality theorem for demi-matroids was obtained. Since we are interested only in duality for linear codes, we will avoid the terminology of demi-matroids. We stress that the first work providing a proof for the duality theorem was [81], using a multiset structure.

We can find in the literature several equivalent ways to define matroids, each of them allowing a nice and interesting interpretation when doing a parallel with the notion of linear independence, see [110] and [109].

**Definition 5.11** A *matroid* $M = ([n], \rho)$ over $[n]$ consists of a function $\rho : \mathcal{P}([n]) \to \mathbb{Z}$ (called *rank function*) satisfying the following properties:

*(i)* $0 \leq \rho(A) \leq |A|$ for every $A \in \mathcal{P}([n])$;
*(ii)* if $A \subset B \subset [n]$, then $\rho(A) \leq \rho(B)$;
*(iii)* for any $A, B \in \mathcal{P}([n])$, $\rho(A \cup B) + \rho(A \cap B) \leq \rho(A) + \rho(B)$.

The integer value $\rho([n])$ is called the *rank of* $M$.

**Definition 5.12** Let $M$ be a matroid on $[n]$ with rank function $\rho$. The *dual matroid* of $M$ is a matroid $M^*$ on $[n]$ whose rank function $\rho^*$ is defined by

$$\rho^*(A) = |A| + \rho([n] \setminus A) - \rho([n]) \tag{5.10}$$

for every $A \subset [n]$.

It is easy to verify that $M^{**} = M$ and that $\rho^*$ satisfies the properties of a matroid, hence the concept of dual matroid is well defined. A nice motivation for the definition of the dual matroid relies on graph theory and can be found in [110]. Note that

$$(\rho^*)^*(A) = |A| + \rho^*([n] \setminus A) - \rho^*([n]),$$

for every $A \subset [n]$. Hence, using the fact that $\rho(\emptyset) = 0$ and expanding (using 5.10) the right-hand side of the previous equality, we obtain that $\rho^{**} = \rho$. Therefore,

$$\rho(A) = |A| + \rho^*([n] \setminus A) - \rho^*([n]) \tag{5.11}$$

for every $A \subset [n]$.

Let $P$ be a poset over $[n]$ and $M$ be a matroid with rank $k = \rho([n])$. By Eq. (5.10), $\rho^*([n]) = n - k$. For each $i \in \{0, \ldots, k\}$ and $j \in \{0, \ldots, n-k\}$, we define

$$s_i^P := \max\{|[n] \setminus \langle [n] \setminus A \rangle_P| \,;\ A \subset [n],\ \rho(A) \leq i\}$$

and

$$t_j^{P^*} := \max\{|[n] \setminus \langle [n] \setminus A \rangle_{P^*}| \,;\ A \subset [n],\ \rho^*(A) \leq j\}.$$

We stress that for every such $i$ and $j$, $s_i^P \geq 0$ and $t_j^{P^*} \geq 0$ because $\rho(\emptyset) = \rho^*(\emptyset) = 0$. Furthermore, by a similar argument,

$$\overline{s}_i := \min\{|\langle A \rangle_P| \,;\ A \subset [n],\ \rho([n]) - \rho([n] \setminus A) \geq i\}$$

and

$$\overline{t}_j := \min\{|\langle A\rangle_{P^*}| \,;\ A \subset [n],\ \rho^*([n]) - \rho^*([n] \setminus A) \geq j\}$$

are also well defined for every $i \in \{0, \ldots, k\}$ and $j \in \{0, \ldots, n-k\}$.

**Proposition 5.13** *$n - s^P_{k-i} = \overline{s}_i$ and $n - t^{P^*}_{n-k-j} = \overline{t}_j$*

*Proof* Exercise 7. ■

**Theorem 5.14 ([13, Theorem 5])** *Let $M$ be a matroid on $[n]$ with rank $\rho([n]) = k$ and $P$ be a poset over $[n]$. Then,*

$$\{n - s^P_{k-1}, n - s^P_{k-2}, \ldots, n - s^P_0\} \bigcap \{t^{P^*}_0 + 1, t^{P^*}_1 + 1, \ldots, t^{P^*}_{n-k-1}\} = \emptyset.$$

*Proof* Suppose that $n - s^P_i = t^{P^*}_j + 1$ for some $i \in \{0, \ldots, k\}$ and $j \in \{0, \ldots, n-k\}$. Hence, by Proposition 5.13, $\overline{s}_i = n + 1 - \overline{t}_j$. Consider $A \subset [n]$ such that $|\langle A\rangle_P| = \overline{s}_i$ and $\rho([n]) - \rho([n] \setminus A) \geq i$. If $B = [n] \setminus \langle A\rangle_P$, then $B \in \mathcal{I}(P^*)$ and $|\langle B\rangle_{P^*}| = |B| = n - |\langle A\rangle_P| = n - \overline{s}_i$, so that $|\langle B\rangle_{P^*}| = \overline{t}_j - 1$. By the definition of $\overline{t}_j$, $\rho^*([n]) - \rho^*([n] \setminus B) \leq j - 1$, hence

$$\begin{aligned} n - k - \overline{s_i} + i &\leq \rho^*([n]) - |\langle A\rangle_P| + (\rho([n]) - \rho([n] \setminus A)) \\ &\leq \rho^*([n]) - |[n] \setminus B| + \rho([n]) - \rho([n] \setminus \langle A\rangle_P) \\ &= \rho^*([n]) - (|[n] \setminus B| - \rho([n]) + \rho(B)) \\ &= \rho^*([n]) - \rho^*([n] \setminus B) \\ &\leq j - 1. \end{aligned}$$

In a similar way, $n - (n-k) - \overline{t}_j + j \leq i - 1$. Therefore, $-1 = n - \overline{s}_i - \overline{t}_j \leq -2$, which is a contradiction. ■

### 5.3.2 *Wei's Duality Theorem*

As described in Sect. 1.3.1, Victor K. Wei in ([108], 1991) defined the generalized Hamming weights, which can be seen as an $n$-dimensional generalization of the concept of minimum distance. In that work, it was proved the *Wei Duality Theorem*, which states that the weight hierarchy (according to the Hamming metric) of a code is uniquely determined by the weight hierarchy of its dual. We are going to use the elements of the theory of matroids described in the previous section to obtain a Wei-type duality theorem for poset codes.

The generalized weights and its monotonicity (Theorem 1.11) were introduced, for the Hamming metric, in Sect. 1.3. We start generalizing it for poset metrics.

**Definition 5.15** Let $P$ be a poset over $[n]$ and $C$ an $[n,k]_q$ linear code. The $i$-th *generalized $P$ weight* of $C$ is defined as

$$\delta_{i,d_P}(C) := \min\{|\langle \mathrm{supp}(D)\rangle_P|;\ D \subseteq C \text{ and } \dim(D) = i\}$$

where $\mathrm{supp}(D) = \{i \in [n];\ x_i \neq 0 \text{ for some } \mathbf{x} = (x_1, \ldots, x_n) \in D\}$.

**Proposition 5.16** *If $P$ is a poset over $[n]$ and $C$ is an $[n,k]$ linear code, then*

$$\delta_{1,d_P}(C) < \cdots < \delta_{k,d_P}(C).$$

*Proof* By definition, it is clear that $\delta_{1,d_P}(C) \leq \cdots \leq \delta_{k,d_P}(C)$. Suppose $\delta_{i,d_P}(C) = \delta_{i+1,d_P}(C)$ for some $i \in \{1, \ldots, k-1\}$. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}\}$ be a basis of a code $D_1 \subset C$ such that $\dim(D_1) = i+1$ and $\delta_{i+1,d_P}(C) = |\langle \mathrm{supp}(D_1)\rangle_P|$. Let $j \in \langle \mathrm{supp}(D_1)\rangle_P$ be a maximal element (according to $P$) in $\mathrm{supp}(D_1)$. Suppose, without loss of generality, that $j \in \mathrm{supp}(\mathbf{v}_1)$. It follows that for every $s \in \{2, \ldots, i+1\}$ there exist $\lambda_s \in \mathbb{F}_q$ such that $j \notin \mathrm{supp}(\mathbf{v}_s - \lambda_s \mathbf{v}_1)$. Therefore, $\{\mathbf{v}_1, \mathbf{v}_2 - \lambda_2\mathbf{v}_1, \ldots, \mathbf{v}_{i+1} - \lambda_{i+1}\mathbf{v}_1\}$ is a basis of $D_1$ satisfying $j \in \mathrm{supp}(\mathbf{v}_1)$ and $j \notin \mathrm{supp}(\mathbf{v}_s)$ for every $s \neq 1$. Hence, $\{\mathbf{v}_2 - \lambda_2\mathbf{v}_1, \ldots, \mathbf{v}_{i+1} - \lambda_{i+1}\mathbf{v}_1\}$ is a basis for an $i$-dimensional space $D_2 \subset D_1$. Since $j$ is maximal in $\mathrm{supp}(D_1)$, $|\langle \mathrm{supp}(D_2)\rangle_P| < |\langle \mathrm{supp}(D_1)\rangle_P|$, which is a contradiction. ■

The set

$$\{\delta_{1,d_P}(C), \delta_{2,d_P}(C), \ldots, \delta_{k,d_P}(C)\}$$

is called the *weight hierarchy* of $C$.

Let $C$ be a $k$-dimensional linear code over $\mathbb{F}_q^n$. As described in [48], we may obtain new codes from $C$ by puncturing or by shortening it in the following way: Given $T \subset [n]$, the punctured code $C^T$ is obtained by deleting the coordinates corresponding to $T$ from each codeword, while the shortened code $C_T$ is formed by all the codewords $c \in C$ such that $\mathrm{supp}(c) \subset T$, therefore, $\dim C = \dim C^T + \dim C_T$. For more details concerning punctured and shortened codes, see [48, Section 1.5].

Define the function $\rho_C : \mathcal{P}([n]) \to \mathbb{Z}$ by

$$\rho_C(A) = \dim(C^{[n]\setminus A}).$$

It is straightforward to verify that $\rho_C$ is a rank function. The matroid $M_C = ([n], \rho_C)$ is called the *matroid of $C$*. Note that

$$\begin{aligned}\rho_{C^\perp}(A) &= \dim(C^\perp)^{[n]\setminus A} = \dim(C_{[n]\setminus A})^\perp \\ &= (n - |[n] \setminus A|) - \dim C_{[n]\setminus A} \\ &= |A| - \dim C + \dim C^{[n]\setminus A}\end{aligned}$$

$$= |A| - \rho_C([n]) + \rho_C([n] \setminus A) = {\rho_C}^*(A).$$

Hence, the matroid determined by ${\rho_C}^* = \rho_{C^\perp}$ is the matroid of $C^\perp$. We will see now that the weight hierarchy of a code is essentially a combinatorial invariant.

**Proposition 5.17** *Let $C \in \mathbb{F}_q^n$ be a linear code and $P = ([n] \preceq)$ a poset. Then, $\delta_{i,d_P}(C) = n - s^P_{k-i}$ and $\delta_{j,d_{P^*}}(C^\perp) = n - t^{P^*}_{n-k-j}$*

*Proof* Let $D \subset C$ be an $i$-dimensional subcode such that $|\langle \mathrm{supp}(D)\rangle_P| = \delta_{i,d_P}(C)$ and let $A = \mathrm{supp}(D)$. Hence, $k - \rho_C([n] \setminus A) = \dim(D) = i$ and by Proposition 5.13,

$$\delta_{i,d_P}(C) = |\langle \mathrm{supp}(D)\rangle_P| = |\langle A\rangle_P| \geq \overline{s}_i = n - s^P_{k-i}.$$

On the other hand, if $A$ is any subset of $[n]$ satisfying $k - \rho_C([n] \setminus A) = i$ and $|\langle A\rangle_P| = \overline{s}_i$, since $\dim(C_A) = k - \rho_C([n] \setminus A) = i$, it follows that

$$\delta_{i,d_P}(C) \leq |\langle \mathrm{supp}(C_A)\rangle_P| \leq |\langle A\rangle_P| = \overline{s}_i = n - s^P_{k-i},$$

so that $\delta_{i,d_P}(C) = n - s^P_{k-i}$.

Similarly, we can proof that $\delta_{j,d_{P^*}}(C^\perp) = n - t^{P^*}_{n-k-j}$. ■

We stress that Proposition 5.16 is also a direct consequence of the previous proposition and of a combinatorial property of matroids, see Exercise 9. Furthermore, by the previous proposition and Theorem 5.14, we get the Wei duality theorem for poset codes:

**Theorem 5.18 ([13, Theorem 11])** *Wei's Duality Theorem for poset codes:*

$$\{\delta_{1,d_P}(C), \ldots, \delta_{k,d_P}(C)\} \bigcup \{\delta_{1,d_{P^*}}(C^\perp), \ldots, \delta_{n-k,d_{P^*}}(C^\perp)\} = [n]$$

*and*

$$\{\delta_{1,d_P}(C), \ldots, \delta_{k,d_P}(C)\} \bigcap \{\delta_{1,d_{P^*}}(C^\perp), \ldots, \delta_{n-k,d_{P^*}}(C^\perp)\} = \emptyset.$$

*Proof* Propositions 5.14 and 5.17 ensure that the intersection between the two weight hierarchies is empty. The emptiness of the intersection and Proposition 5.16 complete the proof. ■

Theorem 5.18 ensures that once the weight hierarchy of a code $C$ is known, the weight hierarchy of its dual code can be determined.

## 5.4 Chapter Notes

Since the publication of [46] in 1998, which was, up to our knowledge, the first work about MacWilliams' identities for poset codes, a number of authors have generalized in many coding-theoretical ways such identities (see [9, 32, 37, 57] and [60], for example). With some exceptions (see the work [12] by Britz and Shiromoto), most of these generalizations are obtained by using two main theories: characters over finite fields and association schemes, both of them consist in translating a combinatorial problem into an algebraic problem. We decided to use the character approach, since it demands few new tools (the code language can be used) and the prerequisites are minimum. In order not to relegate the association schemes approach into orphanhood, we should sketch the results obtained by using schemes and also the connections between the two approaches (see [22] for details about association schemes).

Either using characters or association schemes, the first step to obtain the MacWilliams identity consists in defining a partition of codes according to some property concerning poset ideals, for example:

- When dealing with hierarchical posets, the cardinality of ideals can be used. This allows us to partition a code into subsets whose elements have same weight, hence obtaining the well-known weight enumerator (see [64] and [85]);
- For NRT posets, we can use the shape of ideals (defined in Chap. 4) to obtain the desired partition and this allows us to define the shape enumerator (see [9]).

In our approach, equivalence relations over ideals provide such partitions. The group of linear isometries acts transitively on equivalence classes of codewords if we consider the automorphism relation ($E_H$). The transitivity of isometry groups is a key property which allows us to obtain MacWilliams-type identities without using the Möbius inversion formula. This is a core property to construct association schemes from codes (details can be found in [22]). The transitivity allows us to construct an association scheme whose inner distribution is uniquely determined by the inner distribution of the dual code scheme (which have to be, in some way, related to the dual of the original scheme).

In [9], it was proved that for shape enumerators, the dual of the association scheme is isomorphic to the association scheme obtained by the dual code if, and only if, the poset is self-dual. Furthermore, concerning weight enumerators, in [85], it was shown that the dual of the association scheme is isomorphic to the association scheme obtained by the dual code if, and only if, the poset is hierarchical. We conjecture that this is also true for complement isomorphism posets, that is, concerning the partition given by $E_S$, the dual of the association scheme is isomorphic to the association scheme obtained by the dual code if, and only if, the poset is complement isomorphism. Since an NRT poset (which is a self-dual poset) is also a complement isomorphism poset (see Exercise 9) and in this case, the shape enumerator coincides with the distribution given by relation $E_S$, that would be a natural extension of the results of [9]. We also stress that it is still not

known whether $GL_P(\mathbb{F}_q^n)$ acts transitively on the equivalence classes of codewords determined by $E_S$ (see Exercise 11).

It is worth noting that the association scheme approach provides a natural way to extend the results for non-linear codes, including a definition for the dual of a non-linear code which generalizes the usual definition for linear subspaces. This was first observed by Delsarte in [22].

Finally, MacWilliams-type identities and Wei's duality theorem can be generalized when weakening the algebraic toolkit, working on modules over finite rings instead of vector spaces over finite fields. See [10] and [44], were results where obtained for Frobenius rings.

## 5.5 Exercises

1. Show that the relation $E_S$ is an equivalence relation and that it admits a dual relation if, and only if, the poset is a complement isomorphism.
2. Prove that if $P$ is hierarchical, then $P$ is a complement isomorphism. Provide an example of a complement isomorphism poset which is not hierarchical.
3. Prove the following properties of the equivalence relations defined in Sect. 5.1.

   (a) $E_{Aut(P)} \subset E_S \subset E_C$.
   (b) $P$ is hierarchical if, and only if, $E_S = E_C$.

4. Assuming the definitions and notations of Sect. 5.2, show that

$$S^P_{\overline{I},E}(\mathbf{u}) = \overset{\circ}{\bigcup_{J \in \overline{I}}} S^P_J(\mathbf{u})$$

   where the union is disjoint. Conclude that for a linear code $C$ in $\mathbb{F}_q^n$,

$$A_{\overline{I},E}(C) = \sum_{J \in \overline{I}} |S^P_J \cap C|.$$

   Furthermore,

$$C = \overset{\circ}{\bigcup_{\overline{I} \in \mathcal{I}(P)/E}} C \cap S^P_{\overline{I},E}.$$

5. Prove that for any linear code $C \subseteq \mathbb{F}_q^n$,

$$\sum_{\mathbf{v}\in C}\chi(\mathbf{u}\cdot\mathbf{v}) = \begin{cases} |C| & \text{if } \mathbf{u}\in C^{\perp} \\ 0 & \text{if } \mathbf{u}\notin C^{\perp} \end{cases}$$

where $\mathbf{u}\cdot\mathbf{v} = \sum_{i=1}^{n} u_i v_i$. **Hint:** *Use the orthogonality relation of characters.*

6. ([18, Lemma 3.1]) Use Exercise 5 to conclude that for $\overline{I} \in \mathcal{I}(P)/E$,

$$A_{\overline{I},E}(C) = \frac{1}{|C^{\perp}|} \sum_{\overline{J^c}\in\mathcal{I}(P^*)/E^*} \sum_{\mathbf{u}\in C^{\perp}\cap S^{P^*}_{\overline{J^c},E^*}} \sum_{\mathbf{v}\in S^{P}_{\overline{I},E}} \chi(\mathbf{u}\cdot\mathbf{v}).$$

**Hint:** *Prove that* $A_{\overline{J^c},E^*}(C^{\perp}) = \sum_{\mathbf{v}\in C^{\perp}\cap S_{\overline{J^c},E^*}} 1$.

7. ([13, Lemma 4]) Let $M = ([n], \rho)$ be a matroid with $\rho([n]) = k$ and let $P$ be a poset over $[n]$.
   (a) Prove that $s_i^P = \max\{|[n] \setminus \langle [n] \setminus A\rangle_P| \,;\ A \subset [n],\ \rho(A) = i\}$.
   (b) Prove that $\overline{s}_i = \min\{|\langle A\rangle_P| \,;\ A \subset [n],\ \rho([n]) - \rho([n] \setminus A) = i\}$.
   (c) Conclude that $n - s^P_{k-i} = \overline{s}_i$ for every $i \in \{0, \ldots, k\}$.
   (d) In a similar way, prove that $n - t^{P^*}_{n-k-j} = \overline{t}_j$ for every $j \in \{0, \ldots, n-k\}$.

8. Let $M = ([n], \rho)$ be a matroid with $\rho([n]) = k$ and $P$ be a poset over $[n]$. Suppose $n - s_i^P = t_j^{P^*} + 1$ for some $i \in \{0, \ldots, k\}$ and $j \in \{0, \ldots, n-k\}$. Prove that

$$k - \overline{t_j} + j \leq i - 1.$$

**Hint:** *See the proof of Theorem 5.14.*

9. Let $M = ([n], \rho)$ be a matroid with $\rho([n]) = k$.
   (a) ([13], Lemma 1) For every $i \in [n]$ and $A \subset [n]$, show that $\rho(A \setminus \{i\}) = \rho(A) - 1$ or $\rho(A \setminus \{i\}) = \rho(A)$.
   (b) ([13], Lemma 3) Conclude that $s^P_{i-1} < s^P_i$ for every $i \in \{0, \ldots, k\}$.

10. Show that an NRT poset is also a complement isomorphism poset.
11. (Open problem) Considering complement isomorphism posets, try to prove (or disprove) that the group of linear isometries acts transitively on the equivalence classes of vectors determined by the relation $E_S$.

# Chapter 6
# The General Case: Dead Ends and Hidden Passages


In the previous chapters we saw that the hierarchical poset metrics are as tractable as the Hamming metric, while the case of metrics defined by multi-chains is much more complicated. In this chapter we explore some topics that are studied concerning coding with general poset metrics, with no restrictions on the poset. As it should be expected, the problems are difficult.

In Sect. 6.1 we study the packing and covering radii of a code and show that, in the general situation, this is an intractable problem, even in one-dimensional case. In Sect. 6.2 we see that there is a way to tackle somehow this situation: using the knowledge about hierarchical poset metrics to produce bounds for the general case. Finally, in Sect. 6.3, we briefly survey what is known about perfect and MDS codes in the general situation.

## 6.1 Packing and Covering Radii

For those who are used to the most common setting of coding theory—a vector space endowed with the Hamming metric—the packing radius of a code encompasses some surprises. In the Hamming setting, the minimum distance $\delta_{d_H}(C)$ and the packing radius $R_{d_H}(C)$ are related by the well-known formula $R_{d_H}(C) = \lfloor(\delta_{d_H}(C) - 1)/2\rfloor$. As seen in Chap. 3, Proposition 3.5, for a hierarchical poset, the relation between these invariants is more complicated, but still, the packing radius is determined by the minimum distance. However, the existence of a relation between the packing radius and the minimum distance is a peculiar characteristic of hierarchical posets (Theorem 3.8). In Chap. 4 we saw that the packing radius of a code may attain every value between the two essential bounds: $\lfloor(\delta_{d_P}(C) - 1)/2\rfloor \leq R_{d_P}(C) \leq \delta_{d_P}(C) - 1$. In Sect. 4.2.2 we saw an interesting example (considering the NRT poset) in which the packing radius attains these values.



M. Firer et al., *Poset Codes: Partial Orders, Metrics and Coding Theory*, SpringerBriefs in Mathematics, https://doi.org/10.1007/978-3-319-93821-9_6

In this section we shall see that determining the packing radius of a code is a very difficult problem even in the most simple case of a code containing only two elements: $C = \{\mathbf{0}, \mathbf{c}\}$, $\mathbf{c} \neq \mathbf{0}$. Since $C$ is determined by the element $\mathbf{c}$, we shall denote $R_{d_P}(C) = R_{d_P}(\mathbf{c})$.

We will explain only the main ideas, but detailed proofs may be found in [26]. Considering only the fact that a poset metric $d_P$ is invariant by translations, it is immediate to prove that $\mathbf{x} \in B_{d_P}(\mathbf{0}, r)$ if, and only if, $\mathbf{x} - \mathbf{c} \in B_{d_P}(\mathbf{c}, r)$ and hence

$$R_{d_P}(\mathbf{c}) = \min_{\mathbf{x} \in \mathbb{F}_q^n} \{\max\{\varpi_P(\mathbf{x}), \varpi_P(\mathbf{x} - \mathbf{c})\}\} - 1. \tag{6.1}$$

Any vector $\mathbf{x}$ attaining this minimum is called a *packing vector* of $\mathbf{c}$.

Let $\mathbf{z}$ be a packing vector of $\mathbf{c}$, and let $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{F}_q^n$ be defined by $x_i = c_i$ if $z_i \neq 0$ and $x_i = 0$ otherwise. We have that $\mathbf{x}$ is also a packing vector of $\mathbf{c}$ and $\operatorname{supp}(\mathbf{x}) \subset \operatorname{supp}(\mathbf{c})$. This means that the packing radius of a vector $\mathbf{c}$ depends only on what happens with the ideal $\langle \operatorname{supp}(\mathbf{c}) \rangle_P$ so we can transform the determination of the packing radius of a vector into a problem concerning solely ideals. Given $A \subseteq \operatorname{supp}(\mathbf{c})$, we define $\varpi_P(A) = |\langle A \rangle_P|$ and $A_{\mathbf{c}} = \operatorname{supp}(\mathbf{c}) \setminus A$. With these definitions we can translate the expression of the packing radius in (6.1) into

$$R_{d_P}(\mathbf{c}) = \min_{A \subseteq \operatorname{supp}(\mathbf{c})} \{\max\{|\langle A \rangle_P|, |\langle A_{\mathbf{c}} \rangle_P|\}\} - 1. \tag{6.2}$$

Considering the set $M_P(A)$ of all $P$-maximal entries of $A$, we have that $\langle A \rangle_P = \langle M_P(A) \rangle_P$, so that $|\langle A \rangle_P| = |\langle M(A) \rangle_P|$, hence

$$\min_{A \subseteq \operatorname{supp}(\mathbf{c})} \{\max\{|\langle A \rangle_P|, |\langle A_{\mathbf{c}} \rangle_P|\}\} = \min_{X \subseteq M_P(\operatorname{supp}(\mathbf{c}))} \{\max\{|\langle X \rangle_P|, |\langle X_{\mathbf{c}} \rangle_P|\}\}.$$

We remark that for any subset $X \subseteq M_P(\operatorname{supp}(\mathbf{c}))$, the set $M_P(\operatorname{supp}(\mathbf{c}))$ is the disjoint union of $X = M_P(X)$ and $M_P(X_{\mathbf{c}})$. To determine $X \subseteq M_P(\operatorname{supp}(\mathbf{c}))$ that attains the minimum of the right-hand side of the last equality, we can start with any packing vector $\mathbf{x}$ of $\mathbf{c}$ with $\operatorname{supp}(\mathbf{x}) \subseteq \operatorname{supp}(\mathbf{c})$ and set $X = M_P(\operatorname{supp}(\mathbf{x})) \cap M_P(\operatorname{supp}(\mathbf{c}))$. This remark and Eq. (6.2) permit to translate the problem of finding the packing radius of $\mathbf{c}$ into a problem of determining a partition of ideals in a poset, namely

$$R_{d_P}(\mathbf{c}) = \min_{A, B \subseteq M_P(\operatorname{supp}(\mathbf{c}))} \{\max\{|\langle A \rangle_P|, |\langle B \rangle_P|\}\} - 1,$$

where $M_P(\operatorname{supp}(\mathbf{c}))$ is the disjoint union of $A$ and $B$.

Since we are partitioning the maximal elements of $\operatorname{supp}(\mathbf{c})$, finding the packing radius of $\mathbf{c}$ (or the packing radius of a one-dimensional code) is called the *generalized partitioning problem*. It is a generalization of the well-known *partitioning problem*, which can be stated as follows: given a collection $T = \{t_1, t_2, \ldots, t_s\}$ (allowing multiplicities) of positive real numbers, we wish to find a partition of $T$

into disjoint subsets, $T = T_1 \cup T_2$, such that the difference $|\sum_{t_i \in T_1} t_i - \sum_{t_j \in T_2} t_j|$ is minimized. Consider $s$ masses with weights corresponding to $T$. The partition problem amounts to distributing the masses between two plates of a scale so that it becomes as balanced as possible. In the particular case that the maximal elements of $\mathrm{supp}(\mathbf{c})$ generate disjoint ideals, if we denote $t_i = |\langle\{i\}\rangle_P|$, for $i \in M_P(\mathrm{supp}(\mathbf{c}))$, we see that the determination of the packing radius of a one-dimensional code is equivalent to the traditional partitioning problem. This is the case, for example, if $P$ is a disjoint union of chains, as in Chap. 4. The difficulty arises because the partitioning problem is NP-hard (see [80] for details). To stress this point, we restate it explicitly: determining the packing radius of a code $C = \{\mathbf{0}, \mathbf{c}\}$ is an NP-hard problem!

All this is needed to find the packing radius of a single vector. It means that for finding the packing radius of a $k$-dimensional code we need first to solve a collection of $q^k$ problems, each one an NP-hard problem. If the determination of the minimum distance of a code is an intractable problem [107], to compute the packing radius of a code with a general poset metric is incomparably intractable. This seems to be a dead end, but there is some hidden passage to the problem, since there is a heuristic algorithm for the traditional partitioning problem[1] that performs well in many cases, known as the Karmakar–Karp algorithm [56]. The algorithm makes use of the *differencing method*. The differencing method and the algorithm were both generalized by D'Oliveira in [26] and they are used to determine the packing radius of a code when considering a general poset metric. To evaluate the performance of this algorithm is an open problem, interesting on its own sake, especially to determine whether there exists a threshold behavior for this heuristic.

## 6.2 Bounds Using Hierarchical Posets

In Chap. 3, considering $P = ([n], \preceq)$ to be a hierarchical poset, the canonical $P$-decomposition of a code was introduced: up to a $P$-isometry a code $C$ can be decomposed as $\bigoplus_i C_i$, where $\mathrm{supp}(C_i) \subset H_i(P)$. This decomposition was crucial in establishing formulae for the main invariants of a code (minimum distance, weight distribution, packing and covering radii) and also it played a crucial role in the proof of Theorem 3.8.

This kind of decomposition is peculiar to the hierarchical case and in this sense, we have a dead end: for the general case, it is not possible to establish simple explicit formulae for the packing and covering radii, neither to have propositions as useful as the classical MacWilliams' identity. Nevertheless, there is a hidden passage to

[1] The partitioning problem is equivalent to determining the packing radius of a vector if the intersection of principal ideals in the poset is trivial, in the sense that for any two principal ideals $I, J$ either $I \subseteq J$, $J \subseteq I$ or $I \cap J = \emptyset$. This includes not only the case of multi-chains, but also many other cases. See Exercise 5.

attain qualitative results: bounds. Indeed, for a general poset, it is possible to find a decomposition of a code as direct sum of smaller codes, with a weaker unicity of parameters, which can be used at least to produce bounds for the invariants. The details of the construction and the proofs of the results can be found in [39], but we shall explain how it is done and how it can be used.

Consider a partition of $[n]$ as the disjoint union $[n] = \cup_{i=1,\ldots,r} J_i$ and denote $\mathcal{J} = (J_1, J_2, \ldots, J_r)$. We assume that each part $J_i \neq \emptyset$. We say that a linear code $C$ is *$\mathcal{J}$-decomposable* if it can be expressed as the direct sum

$$C = \bigoplus_{i=1}^{r} C_i,$$

with $\mathrm{supp}(C_i) \subseteq J_i$. We denote $n_i = |J_i|$ , $k_i = \dim(C_i)$ and the array $[(n_1, k_1), (n_2, k_2), \ldots, (n_r, k_r)]$ is called the *$\mathcal{J}$-profile* of $C$. We remark that a $\mathcal{J}$ decomposition may have trivial components, that is, we may have $k_i = 0$. This is a property of the vectorial structure, not depending on the metric structure.

Considering a poset metric $d_P$ we say that $C$ is $(\mathcal{J}, P)$-decomposable if it is $d_P$-equivalent to a code $T(C)$ which is $\mathcal{J}$-decomposable, for some $T \in GL_P(\mathbb{F}_q^n)$. Using this terminology, Corollary 3.3 may be restated by saying that if $P$ is hierarchical and $J_i = H_i$ is the $i$-th level of $P$, then every linear code is $(\mathcal{J}, P)$-decomposable.

A partition $\mathcal{J}' = (J'_1, J'_2, \ldots, J'_s)$ is called a *refinement* of a partition $\mathcal{J} = (J_1, J_2, \ldots, J_r)$ if each $J_i$ can be expressed as the union of some of the $J'_i$ 's. A $(\mathcal{J}, P)$-decomposition $C = \bigoplus_{i=1}^{r} C_i$ is called *reducible* (or refinable) if there is a non-trivial refinement $\mathcal{J}'$ of $\mathcal{J}$ such that $C$ is $(\mathcal{J}', P)$-decomposable. Here, to say that a refinement is non-trivial means that some $J_k$ is expressed as $J_k = J'_{i_1} \cup \cdots \cup J'_{i_l}$, with $l > 1$ and hence none of it equals $J_k$. In case $C = \bigoplus_{i=1}^{r} C_i$ does not admit a non-trivial refinement, we say it is *irreducible*. We say that $C' = \bigoplus_{i=1}^{r} C'_i$ is a *maximal $P$-decomposition* of a code $C$ if $C'$ is $P$-equivalent to $C$ and the decomposition of $C'$ is irreducible.

Maximal $P$-decomposition does exist:

**Theorem 6.1 ([39])** *Given a poset $P = ([n], \preceq)$ and a code $C \subseteq \mathbb{F}_q^n$ , there is a maximal $P$-decomposition.*

We do not present the proof of this theorem (proved in [39, Theorem 1]), since it is too lengthy and technical, but we shall focus on its consequences.

It is important to remark that in case $P$ is a hierarchical poset, a maximal $P$-decomposition is a refinement (perhaps trivial) of a canonical decomposition, but in the general case, we cannot ensure that each part $J_i$ will be contained in a single level $H_{t_i}$ of the poset and this prevents us from finding simple expressions for (and relations between) metric invariants. Nonetheless, we can use the maximal decompositions to compare it to canonical decompositions according to hierarchical posets and produce bounds for the invariants. Let us explain how to do it.

As seen in Sect. 2.1, the set $\mathcal{P}_n$ of all posets over $[n]$ is itself a poset, with the order $\leq$ defined by $P \leq Q$ if $i \preceq_P j$ implies $i \preceq_Q j$. We recall that $\mathcal{H}_n$ denotes the set of all hierarchical posets. A poset $P$ has a least upper bound and a greatest lower bound among the hierarchical posets (see Exercise 2):

$$P^+ = \min\{Q \in \mathcal{H}_n;\, P \leq Q\}, \qquad P^- = \max\{Q \in \mathcal{H}_n;\, P \leq Q\}.$$

Considering a maximal $P$-decomposition of a code $C$, we can get upper and lower bounds for the packing radius of a code.

**Proposition 6.2 ([39, Section IV.A])** *Let $C' = \bigoplus_{i=1}^{r} C'_i$ be a maximal $P$-decomposition of a code $C$. Then,*

$$R_{d_{P^-}}(C) \leq R_{d_P}(C) \leq \min_{i\in\{1,\ldots,r\}} R_{d_{P^+}}(C'_i).$$

*Proof* First of all we note that, considering the natural ordering on $\mathcal{P}_n$, $P^- \leq P \leq P^+$ (see Exercise 2). Also, whenever $P \leq Q$ we have that

$$B_{d_Q}(\mathbf{x}, r) \cap B_{d_Q}(\mathbf{0}, r) \subseteq B_{d_P}(\mathbf{x}, r) \cap B_{d_P}(\mathbf{0}, r) \tag{6.3}$$

and this implies that $R_{d_P}(C) \leq R_{d_Q}(C)$ for every code $C \subseteq \mathbb{F}_q^n$. Thus, since $P^- \leq P$, from 6.3 we get that $R_{d_{P^-}}(C) \leq R_{d_P}(C)$. On the other hand, since $C'_i$ is a subcode of $C'$, $R_{d_P}(C') \leq \min_i R_{d_P}(C'_i)$ and because $R_{d_P}(C'_i) \leq R_{d_{P^+}}(C'_i)$ and $R_{d_P}(C') = R_{d_P}(C)$, it follows that

$$R_{d_{P^-}}(C) \leq R_{d_P}(C) \leq \min_{i\in\{1,\ldots,r\}} R_{d_{P^+}}(C'_i).$$

■

*Remark 6.3* It is an open problem to determine whether $\min_i R_{d_{P^-}}(C'_i) \leq R_{d_P}(C)$ or not. We imagine it is true, since no counter-example is available. If this is indeed the case, Proposition 6.2 could be improved, obtaining a lower bound for $R_{d_P}(C)$ based on a maximal $P$-decomposition of $C$.

## 6.3 Perfect and MDS-Codes

Considering the Hamming metric, perfect codes and MDS codes, although very desirable from the point of view of error correction, they are rare (considering binary or small alphabets). One of the initial motivations to the study of poset codes is the proliferation of perfect and MDS codes. At the very beginning, in the seminal work [14], Brualdi et al. approached this subject, showing that: (1) there is a relative abundance of perfect codes when considering $P$ to be a chain; (2) the non-existence of perfect codes for a poset consisting of two disjoint chains of equal size (except

for the trivial codes), and (3) the extended binary Hamming and Golay codes[2] are perfect when considering a hierarchical poset (of appropriate size) having a unique minimal element.

There is a quite vast literature studying perfect and MDS codes with a poset metric. Here we give a short (and not exhaustive) overview on the subject.

### 6.3.1 Perfect Codes

Perfect codes with a poset metric were studied in three different approaches.

**(i) Fixing a Family of Codes:** Some works consider a code (or a family of codes), and try to classify the posets which turns the given code into a perfect code. This is the case of the extended Hamming codes, initially approached by Brualdi et al. The posets which turns the extended Hamming codes into a perfect code were classified in [50], including the families described by Brualdi et al. and adding two more families of posets, each with two levels and exactly two minimal elements. The posets for which the extended Golay code is perfect were classified in [55]. A more general (and difficult) approach was adopted in [62], where the authors established necessary conditions over a poset $P$ for an $[n, k]_2$ linear code to be a $d_P$-perfect code able to correct errors determined by a vector $\mathbf{e}$ with $\varpi_P(\mathbf{e})$ equals $n-k-2$, $n-k-1$ or $n-k$. The perfectness of the extended Hamming and Golay codes was also studied in [3, 20], considering the poset-block metrics (see Sect. 7.1), a variation of the poset metrics.

**(ii) Fixing a Family of Posets:** In this approach, a poset (or family of posets) is fixed and then one looks for the codes (or families of codes) that are perfect for that poset (or family of posets). Fixing a family of posets was first considered in [54], where the authors consider a family of posets over $[n]$ having two levels, with $n-1$ maximal elements and $j$ minimal elements, with $|\langle i\rangle|$ equals 1 or 2, for all $i \in [n]$. Several examples of perfect codes are given, for different values of $n$ and $j$. Considering the so-called *crown poset*, it was proved in [65] that no code is perfect. Considering the poset-block structure (see Sect. 7.1), for the poset consisting of a single chain, perfect codes were classified in [89].

**(iii) Operations on a Given Perfect Code:** The first work using this approach was [69]. The author considers a $d_P$-perfect $[n, k]_q$ code $C$ and then looks for conditions

[2]There are many ways to construct the extended binary Golay code. One of the most simple ways to describe (but not to work with) is the following: we consider the coordinates of each vector $\mathbf{x} \in \mathbb{F}_2^{24}$ as the coefficients of the binary expansion of the natural numbers smaller than $2^{24}$, that is, we identify $\mathbf{x} = (x_1, \ldots, x_{24})$ with $\sum_{i=1}^{24} x_i 2^{i-1}$; we start with the vector $\mathbf{v}_0 = \mathbf{0}$ and given $\mathbf{v}_0, \ldots, \mathbf{v}_{k-1}$, we choose $\mathbf{v}_k$ as the vector representing the smallest integer which differs from each of the previous $\mathbf{v}_i$ in at least 8 coordinates; the code $\mathcal{G}_{12}$ is the code generated by $\{\mathbf{v}_0, \ldots, \mathbf{v}_{12}\}$ [19].

for a poset $Q$ over $[n+1]$ (actually establishing the valid relations involving the coordinate $n+1$) so that $C$ extends (in a natural given way) to a $d_Q$-perfect code. Similar conditions are given when puncturing the code $C$. In a later work [70], J.G. Lee considered the ordinal sum $P \oplus Q$ of posets (see the definition on Sect. 6.4) and established necessary conditions, concerning the size of the code and a kind of "projection" onto the coordinates corresponding to $P$ and $Q$, for a code to be $d_{P\oplus Q}$-perfect. Some necessary conditions for perfectness are given also when considering the less usual *standard sum* of posets (instead of the ordinal sum).

### 6.3.2 MDS Codes

Also MDS codes were studied in different approaches. The first thing that worth noting is that in general, we should expect to have more MDS codes when considering a poset metric different from the Hamming. It follows from [51, Lemma 4.1] that if $C$ is MDS with respect to $d_P$ and $P \leq Q$ (inequality in the sense of Sect. 6.2), then $C$ is also MDS with respect to $d_Q$. Since an anti-chain is a minimal poset, an MDS code for the Hamming metric will be MDS with respect to any poset metric.

A significant part of the literature is devoted to the case of NRT posets, which was explored in more detail in Sect. 4.3.2. We recall some of the results explained in that section and we quote some others that were not yet mentioned. The correspondence between MDS codes and optimal (with respect to discrepancy) distributions on a unit cube $[0,1)^n$ was established in [101] (considering the alphabet to be a finite field) and nearly simultaneously explored for general alphabets (with no algebraic structure) in [33]. Existence of MDS codes for NRT metric was shown in [93] (for small minimal distances). Near-MDS codes (that is, $\delta_{1,d_P}(C) = n-k$ and $\delta_{2,d_P}(C) = n-k+2$) were constructed in [8], where the authors also describe the weight distribution of Near-MDS codes. A more general tool to construct MDS-evaluation codes (a generalization of Reed-Salomon codes) for the NRT metric was described in [52] (see Sect. 4.3.2).

Excluding the case of hierarchical posets where MDS codes are characterized in terms of smaller MDS codes with the Hamming metric, as seen in Chap. 3 and the NRT case, the study of MDS codes for a general poset metric is much harder. The outmost understanding on the subject appears in [51], using the concept of $I$-perfect. Given a linear code $C \subseteq \mathbb{F}_q^n$ and an ideal $I \in \mathcal{I}(P)$, we say that $C$ is *$I$-perfect* if

$$\mathbb{F}_q^n = \overset{\circ}{\bigcup_{\mathbf{c}\in C}} B_P(\mathbf{c}, I),$$

where the union is disjoint and

$$B_P(\mathbf{c}, I) := \{\mathbf{x} \in \mathbb{F}_q^n; \operatorname{supp}(\mathbf{c}-\mathbf{x}) \subset I\}$$

is called the $I$-ball centered at $\mathbf{c}$.

The main result in [51] states that an $[n,k]_q$ linear code is MDS if, and only if, it is $I$-MDS for every $i \in \mathcal{I}(P)$ with $|I| \leq n-k$. This result involves a deep understanding of the role of ideals and its orbits, and it is the key to the MacWilliams' relations and MacWilliams-type identities presented in Sect. 5.2.

## 6.4 Chapter Notes

At this point we need to make some remarks about the case of a general poset.

**About Syndrome Decoding**

An important issue concerning the $P$-decomposition of a code is related to the difficulty of syndrome decoding. As seen in Sect. 1.3.2, syndrome decoding depends on the choice of coset leaders. For an $[n,k]_q$ linear code $C$, the coset leaders form a lookup table, having $q^{n-k}$ elements. Let us consider a $P$-decomposition $C' = \bigoplus_{i=1}^{r} C'_i$, with $n_i = |\mathrm{supp}(C'_i)|$ and $k_i = \dim(C'_i)$. We denote $J_0 = [n] \setminus \mathrm{supp}(C')$ and $n_0 = |J_0|$.

It is easy to see that $k = \sum_{i=1}^{r} k_i$ and $n = n_0 + \sum_{i=1}^{r} n_i$. The lookup table of $C'_i$ has

$$\prod_{i=1}^{r} q^{n_i - k_i} = q^{n-k-n_0},$$

elements. This means that finding a maximal decomposition (in which $n_0$ is maximal) divides the size of the lookup table by $q^{n_0}$. Finding a maximal $P$-decomposition is a lengthy task[3] and the improvement attainable in syndrome decoding is not as good as in the case of a hierarchical poset (see Sect. 3.5), but since it is a one-time job, it is worthy to be done. Moreover, in case $\langle \mathrm{supp}(C'_i) \rangle \cap \langle \mathrm{supp}(C'_j) \rangle = \emptyset$ for all $i \neq j$, we need only $\sum_{i=1}^{r} q^{n_i - k_i}$ coset leaders to perform syndrome decoding (see Exercise 3).

**Relation and Operations with Posets**

There are some traditional operations on posets in which we can construct a new poset from two given posets. Given posets $P = ([n], \preceq_P)$ and $Q = ([m], \preceq_Q)$, we define:

1. The *direct sum* $P + Q$ is a poset over $[n+m]$ defined by $i \preceq_{P+Q} j$ if either $i, j \in [n]$ and $i \preceq_P j$ or $i, j \in [n+m] \setminus [n]$ and $i - n \preceq_Q j - n$.
2. The *ordinal sum* $P \oplus Q$ is a poset over $[n+m]$ defined by $i \preceq_{P \oplus Q} j$ if either $i, j \in [n]$ and $i \preceq_P j$ or $i, j \in [n+m] \setminus [n]$ and $i - n \preceq_Q j - n$ or $i \in [n]$ and $j \in [n+m] \setminus [n]$.

[3] In [39] there is an algorithm to construct a maximal $P$-decomposition of a code.

3. The *direct product* $P \times Q$ is a poset over $[n] \times [m]$ defined by $(i, j) \preceq_{P\times Q} (k, l)$ if $i \preceq_P k$ and $j \preceq_Q l$.
4. The *ordinal product* $P \otimes Q$ is a poset over $[n] \times [m]$ defined by $(i, j) \preceq_{P\otimes Q} (k, l)$ if $i = k$ and $j \preceq_Q l$ or if $i \preceq_P k$.

The behavior of the invariants of coding theory (and of the desirable properties of the invariants) under these operations is nearly unexplored in the literature. One exception is the minimum distance for partially ordered multisets (see Sect. 7.3) and the extension property for ideals, explored in [102] (see Exercise 4).

**Approximation by Hierarchical Posets**
In Sect. 6.2 we saw how to produce bounds for metric invariants of a general poset metric $P$ by considering the corresponding invariants of the hierarchical posets $P^-$ and $P^+$. To understand the limitations of those bounds, we shall consider a natural metric on the set $\mathcal{P}_n$, the set of all posets over $[n]$. This metric is defined as follows: the metric $d_{\text{ins-del}}(P, Q)$ is the minimum number of relations that is needed to remove or add to reach a poset $P$ from a poset $Q$ (the *insertion-deletion metric* on posets). Using the metric $d_{\text{ins-del}}$ we have that, in general, $P^-$ and $P^+$ may be very far from $P$. Let us consider a poset over $P$ over $[3m]$, having $m$ elements on each of its three levels defined as follows:

1. Given $i \in H_1(P)$ and $j \in H_2(P)$, then $i \preceq_P j$ if, and only if, $j = i + m$.
2. For every $i \in H_2(P)$ and $j \in H_3(P)$ we have that $i \preceq_P j$, except for the case $i = 2m$ and $j = 3m$.

It is easy to see that $P^-$ is obtained from $P$ by removing $m^2 + m - 1$ relations and $P^+$ by adding $m^2 - m + 1$ relations. However, the closest hierarchical poset $\widetilde{P}$ to $P$ is obtained by removing $m$ relations and adding 1. In other words, $d_{\text{ins-del}}(P, P^-) = m^2 + m - 1$ and $d_{\text{ins-del}}(P, P^+) = m^2 - m + 1$ grow as $m^2$, while $d_{\text{ins-del}}(P, \widetilde{P}) = m + 1$ grows linearly on $m$. We believe that approximating $P$ by $\widetilde{P}$ can lead to better bounds, but to work with it one needs to compare $P$ and $\widetilde{P}$ level-by-level, a very delicate task.

## 6.5 Exercises

1. Let $n = 5$. Determine all the posets $P$ (up to isomorphism) for which the code $C = \{00000, 11111\}$ has minimum distance 5 and packing radius 2, 3, or 4.
2. Let $P \in \mathcal{P}_n$ and consider the natural ordering on $\mathcal{P}_n$.
   (a) Prove that $P^+ = \min\{Q \in \mathcal{H}_n; P \leq Q\}$ and $P^- = \max\{Q \in \mathcal{H}_n; Q \leq P\}$ are well defined, that is, prove the existence and unicity of the minimum and the maximum. **Hint:** *Consider the Hasse diagram of $P$ and, looking at levels $H_i(P)$ and $H_{i+1}(P)$, either add all possible relations or remove all the relations, in order to obtain $P^+$ and $P^-$, respectively.*
   (b) Prove that $P \in \mathcal{P}_n$ is hierarchical if, and only if, $P^- = P = P^+$.

3. Let $C' = \bigoplus_{i=1}^{r} C'_i$ be a $P$-decomposition of $C$ and suppose that $\langle \mathrm{supp}(C'_i)\rangle \cap \langle \mathrm{supp}(C'_j)\rangle = \emptyset$ for all $i \neq j$.
   (a) Prove that, for any $\mathbf{x} \in \mathbb{F}_q^n$,
   $$\min_{\mathbf{c}\in C} d_P(\mathbf{c}, \mathbf{x}) = \sum_{i=1}^{r} d_{\pi_i}(\pi_i(\mathbf{c}), \pi_i(\mathbf{x})),$$
   where $\pi_i : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^{n_i}$, $n_i = |\mathrm{supp}(C'_i)|$, $\pi_i$ is the projection and $d_{\pi_i}$ the metric on $\mathbb{F}_q^{n_i}$ induced by $\pi_i$.
   (b) Use the fact that there is $T \in GL_P(\mathbb{F}_q^n)$ such that $T(C) = C'$ and the previous item to conclude that syndrome decoding can be performed using $\sum_{i=1}^{r} q^{n_i - k_i}$ coset leaders.

4. Let $C \subseteq \mathbb{F}_q^n$ and $D \subseteq \mathbb{F}_q^m$ be linear codes. Let $P$ and $Q$ be posets over $[n]$ and $[m]$, respectively.
   (a) Show that $P + Q$ is hierarchical if, and only if, both $P$ and $Q$ are anti-chains. Show that $P \oplus Q$ is hierarchical if, and only if, both $P$ and $Q$ are hierarchical.
   (b) Let $C \oplus D \subseteq \mathbb{F}_q^{n+m}$ be the direct sum of the codes. Determine the minimal weight $\delta_{d_{P+Q}}(C \oplus D)$ and $\delta_{d_{P\oplus Q}}(C \oplus D)$ relative to the direct and ordinal sum of posets.
   (c) Let $C \otimes D \subset \mathbb{F}_q^n \otimes \mathbb{F}_q^m$ be the tensor product of the codes. Determine the minimal weight $\delta_{d_{P\times Q}}(C \otimes D)$ and $\delta_{d_{P\otimes Q}}(C \otimes D)$ relative to the direct and ordinal product of posets.

5. A poset $P$ is said to satisfy the *trivial intersection property (TIP)* if for any two ideals $I, J \in \mathcal{I}(P)$, either $I \subseteq J$, $J \subseteq I$ or $I \cap J = \emptyset$.
   (a) Let $P$ be a poset whose Hasse diagram is a rooted tree with the root as a maximal element. Show that $P$ satisfies the TIP.
   (b) Let $P$ and $Q$ be posets satisfying the TIP. Prove that the direct sum $P + Q$ satisfies the TIP.
   (c) Let $P$ and $Q$ be posets. Prove that the ordinal sum $P \oplus Q$ satisfies the TIP if, and only if, $P$ satisfies the TIP and $Q$ is a chain.

6. Let $P$ be a poset over $[3m]$, having three levels and $m$ elements on each of its levels, defined as follows: *(i)* Given $i \in H_1(P)$ and $j \in H_2(P)$, then $i \preceq_P j$ if, and only if, $j = i + m$. *(ii)* For every $i \in H_2(P)$ and $j \in H_3(P)$ we have that $i \preceq_P j$, except for the case $i = 2m$ and $j = 3m$.
   (a) Describe the hierarchical poset $\widetilde{P}$ closest to $P$ (with regard to the insertion deletion metric $d_{\text{ins-del}}(P, Q)$).
   (b) Prove that $d_{\text{ins-del}}(P, \widetilde{P}) = m + 1$, $d_{\text{ins-del}}(P, P^-) = m^2 + m - 1$ and $d_{\text{ins-del}}(P, P^+) = m^2 - m + 1$.

# Chapter 7
# Generalizations, Variations and Perspectives


In this chapter we present different generalizations and variations of the poset metrics. The approach adopted here is very concise. For each of these generalizations we introduce the proper definitions and only explain the concepts needed to understand the statement of the main results (whose proofs are not even sketched).

There are some reasons to make such a brief overview.

The first reason to do this review is that these generalizations are very recent, they were publicized after 2015 (and some not yet published after a peer review), so it gives an overview of recent topics, with many open questions for interested researchers. In this sense, despite being short, this is a very up-to-date survey of the area. Studying the references, the interested reader will find many open questions, since not much is known about any of these generalizations.

Secondly, all those generalizations may be considered as bricks to the construction of a more comprehensive study program of the space of all channels, which is presented as an "*Epilogue*" in the last section.

A brief overview of the Chapter is as follows.

In Sect. 7.1 we present the poset-block metrics, that combine a poset with blocks. In Sect. 7.2 we introduce the metrics determined by directed graphs (of which the Hasse diagram of a poset is a special case). In each of these two sections we follow, somehow, the main script undergone in the main part of the text, which includes determining explicit expressions for the main invariants, a description of the group of linear isometries, conditions for the existence of a MacWilliams-type identity and the extension property.

On Sect. 7.3 we introduce the recent concept of partially ordered multisets, allowing us to generalize the classical Lee metrics. This is also a generalization of the poset metrics because it coincide with a poset metric when the Hamming and Lee metric do coincide: when the base field is $\mathbb{F}_2$ or $\mathbb{F}_3$.

On Sect. 7.4 we move into a different direction and consider a family of metrics which intersects the family of poset metric only in one single (but not fortuitous)



M. Firer et al., *Poset Codes: Partial Orders, Metrics and Coding Theory*, SpringerBriefs in Mathematics, https://doi.org/10.1007/978-3-319-93821-9_7

case, the Hamming metric. These are the combinatorial metrics and the general script played in Sects. 7.1 and 7.2 is followed also here.

Finally on Sect. 7.5, we present, nearly as a storyboard, what we envision as a possible direction to the study of the space of all channels (or all metrics) from a coding theory perspective, study that still is in an embryonic state.

## 7.1 Poset-Block Metrics

Block metrics, in the context of coding theory, were introduced in 1973 by Gabidulin [41]. As we shall see they are a particular case of Combinatorial Metrics, but it rested untouched since 2006, when Feng et al. in [37] gave a complete understanding of these metrics. After that, posets and block structures were mixed by Alves et al. [3], giving rise to the so-called poset-block metrics.

We let

$$\mathbb{F}_q^N := V_1 \oplus V_2 \oplus \cdots \oplus V_n$$

be a decomposition of $\mathbb{F}_q^N$ as a direct sum of subspaces. We denote $k_i = \dim(V_i) > 0, \pi = (k_1, k_2, \ldots, k_n)$ and we call this decomposition a *block structure*. We remark that $N = k_1 + k_2 + \cdots + k_n$. Being a direct sum we have that each $\mathbf{x} \in \mathbb{F}_q^N$ has a unique decomposition as $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n$, with $\mathbf{x}_i \in V_i$. Considering such a decomposition we have that the $\pi$*-support* of $\mathbf{x}$ is defined as $\mathrm{supp}_\pi(\mathbf{x}) := \{i \in [n]; \mathbf{x}_i \neq \mathbf{0}\}$. Counting the $\pi$-support gives rise to the $\pi$*-weight* $\varpi_\pi(\mathbf{x}) := |\mathrm{supp}_\pi(\mathbf{x})|$. For example, if we consider $N = 3$, $V_1 = \{(x_1, x_2, 0); x_1, x_2 \in \mathbb{F}_q\}$ and $V_2 = \{(0, 0, x_3); x_3 \in \mathbb{F}_q\}$, then we have that $\varpi_\pi(110) = \varpi_\pi(100) = 1$.

If $P = ([n], \preceq)$ is a poset over $[n]$, we may combine the block structure $\pi$ and the poset structure $P$ into a single one. Instead of counting the $\pi$-support we consider the $P$-ideal generated by the $\pi$-support and then count it, obtaining the *poset-block weight* and the *poset-block metric*:

$$\varpi_{(P,\pi)}(\mathbf{x}) := |\langle \mathrm{supp}_\pi(\mathbf{x}) \rangle_P| \quad \text{and} \quad d_{(P,\pi)}(\mathbf{x}, \mathbf{y}) := \varpi_{(P,\pi)}(\mathbf{x} - \mathbf{y}).$$

It is not difficult to see that $d_{(P,\pi)}$ is indeed a metric on $\mathbb{F}_q^N$. What is known about codes with poset-block metrics? Not too much, but we describe the main known results.

### *7.1.1 Linear Isometries*

We denote by $GL_{(P,\pi)}(\mathbb{F}_q^n)$ the group of linear isometries of $\mathbb{F}_q^n$ endowed with the metric $d_{(P,\pi)}$. The description of $GL_{(P,\pi)}(\mathbb{F}_q^n)$ was done in [3] and it is similar to what happens in the poset case. There are two distinguished types of isometries.

**(i) Operating on Ideals:** A linear isometry $T$ which, to each vector $\mathbf{x} \in V_i$, associates a vector

$$T(\mathbf{x}) = T_i(\mathbf{x}) + \sum_{\substack{j \in \langle i \rangle_P \\ j \neq i}} \mathbf{y}_j,$$

where $T_i : V_i \to V_i$ is any linear isomorphism. We remark that $T_i$ depends on the block structure while the $\mathbf{y}_j$'s depend on the poset structure. The set of all such isometries is a group, denoted by $G_{(P,\pi)}$.

**(ii) Induced by Poset Blocks Automorphisms:** A *poset-block automorphism* is a poset automorphism $\sigma$ such that $k_{\sigma(i)} = k_i$ for every $i \in [n]$, in other words, it is a poset automorphism that permutes blocks of same size. This is a group, denoted by $\mathrm{Aut}(P, \pi)$ and, once we fix a basis $\beta_i = \{\mathbf{e}_1^i, \mathbf{e}_2^i, \ldots, \mathbf{e}_{k_i}^i\}$ for each $V_i$, it acts on $\mathbb{F}_q^N$ by sending $\mathbf{e}_j^i$ to $\mathbf{e}_j^{\sigma(i)}$. At this point it is essential the fact that $k_{\sigma(i)} = k_i$.

Similarly to the poset case, we have that $GL_{(P,\pi)}(\mathbb{F}_q^n)$ may be described [3, Theorem 4.11] as the semi-direct product

$$GL_{(P,\pi)}(\mathbb{F}_q^n) \simeq G_{(P,\pi)} \ltimes \mathrm{Aut}(P, \pi).$$

### 7.1.2 MacWilliams' Identity

MacWilliams' Identity for poset-block metric codes was treated in [90] where the poset-block metrics admitting a MacWillliams identity where classified. As it happens in the poset case, we want to establish a relation between the weight distribution $W_C^{(P,\pi)}(z)$ of a code $C$ and the weight distribution $W_{C^\perp}^{(P^*,\pi)}(z)$ of the dual code $C^\perp$ with respect to the $(P^*, \pi)$-metric, where $P^*$ is the opposite poset.

Since poset-block metrics generalizes both poset and block metrics, the conditions which are necessary in either case must be necessary in the poset-block case. Hence, as we saw in Theorem 3.8 and Chap. 5, $P$ being hierarchical is a necessary condition for $(P, \pi)$ to admit a MacWilliams identity. Considering the case that $P$ is an anti-chain, it is easy to see that all the blocks having the same size ($k_i = k_j$ for all $i, j \in [n]$) is also a necessary condition. Indeed, we may consider the smallest instance where this does not happen.

Let $N = 3$, $V_1 = \{(x_1, x_2, 0); x_1, x_2 \in \mathbb{F}_q\}$, $V_2 = \{(0, 0, x_3); x_3 \in \mathbb{F}_q\}$ and $P$ be an anti-chain over [2]. Consider the one-dimensional codes $C_1 = \{000, 100\}$ and $C_2 = \{000, 001\}$. Hence, we have that

$$W_{C_1}^{(P,\pi)}(z) = 1 + z = W_{C_2}^{(P,\pi)}(z).$$

The dual codes $(C_1)^\perp = \{000, 010, 001, 011\}$ and $(C_2)^\perp = \{000, 100, 010, 110\}$ have different weight distribution, namely,

$$W_{(C_1)^\perp}^{(P^*,\pi)}(z) = 1 + 2z + z^2 \neq 1 + 3z = W_{(C_2)^\perp}^{(P^*,\pi)}(z).$$

These conditions, namely, $P$ being a hierarchical poset and all the blocks in the same level of the poset being of the same size ($k_i = k_j$ if $i, j \in H_r(P)$, for every $1 \leq r \leq h(P)$) are not only necessary but also sufficient conditions for the $(P, \pi)$ metric to admit a MacWilliams identity. The proof in [90, Theorem 2] uses characters and follows, with some additional technical difficulties, the original proof of MacWilliams.

### 7.1.3 Canonical Form

A canonical decomposition is available for poset-block codes in case the poset $P$ is a chain. In this case, it follows from [89, Theorem 4] that, up to an isometry, a linear code $C$ may be decomposed as $C = C_1 \oplus C_2 \oplus \cdots \oplus C_l$, with $C_i \subseteq V_i$ where $V_i$ is the $i$-th block, corresponding also to the $i$-th level of the poset $P$. We note that some of the $C_i$'s may be trivial ($C_i = \{\mathbf{0}\}$) and to simplify we write $C = C_{j_1} \oplus C_{j_2} \oplus \cdots \oplus C_{j_r}$, where these are all the non-trivial codes in the decomposition of $C$.

### 7.1.4 Perfect Codes

The study of perfect codes with a poset-block metric is still at initial stages but the known results follow the same variety of approaches used in the case of posets, namely: (1) to fix a particular (family) of posets; (2) to study some families of codes and classify the poset-block metrics that turn it into a perfect code; (3) to find obstructions and bounds for a code being perfect. We briefly sketch the known results in these three approaches.

If $P$ is a chain, we consider the canonical decomposition we just describe. In [89, Theorem 7] it was proved that a code is $d_{(P,\pi)}$-perfect if, and only if, it is MDS and this happens if, and only if, from the Singleton bound [89, Proposition 3], we have that, in its canonical decomposition, $C = V_{j_1} \oplus V_{j_1+1} \oplus \cdots \oplus V_{l-1} \oplus V_l$.

Approaching specific families of codes, we start with the extended Hamming binary code $\overline{\mathcal{H}}(3)$. It is a linear code, with parity check matrix obtained from the usual Hamming code's parity check matrix by adding to it a line with all entries equal to 1, so that it is an $[8, 4]_2$ linear code. We consider a partition $\mathbb{F}_2^n = V_1 \oplus V_2 \oplus \cdots \oplus V_n$ with $k_1 + k_2 + \cdots + k_n = 8$. Then, a poset-block $(P, \pi)$ turns $\overline{\mathcal{H}}(3)$ into a 1-perfect code if, and only if, $k_1 = 4$ and the minimum distance satisfies $\delta_{d_{(P,\pi)}}(\overline{\mathcal{H}}(3)) \geq 2$ [3, Theorem 3.3]. The same work presents a non-trivial family of poset-block metrics that turns the extended [24, 12]-Golay into a perfect code (using designs theory).

Consider now a general poset block metric $d_{(P,\pi)}$ determined by a poset $P = ([n], \preceq)$ and a block structure $\sum_{i=1}^{n} k_i = N$. Suppose that $C$ is a perfect code with packing radius $R := R_{d_{(P,\pi)}}(C)$. Suppose that we add to the structure a new block with $k_{n+1}$ positions and we add order relations involving $n+1$ and the previous elements, so that we get a poset structure $P' = ([n+1], \preceq')$ in which $n+1$ is a maximal element. We consider the code $C'$ obtained from $C$ by adding $k_{n+1}$ new coordinates which may assume any value in the base field. Dass et al. proved that $C'$ is perfect if, and only if, the $P'$-ideal generated by the element $n+1$ satisfies $\langle\{n+1\}\rangle_{P'} \geq r+1$ [20, Theorem 3.2]. Also, working on the opposite direction, that is, by removing a maximal block (let say $n$) and puncturing the code in the corresponding coordinates, one still gets a perfect code with packing radius $R$, unless the block structure is trivial ($k_i = 1$ for each $i \in [n]$) and the code $C$ is one-dimensional.

In this same work the authors determine the height of a poset-block structure that may turn a code $C$ into a perfect code, depending on the minimum distance $\delta(C) = \delta_{d_{(P,\pi)}}(C)$ and the packing radius $R(C) = R_{d_{(P,\pi)}}(C)$: the height $h$ of $P$ satisfies $h = 2R(C) = \delta(C) + 2$ [20, Proposition 4.2]. Using this result, it follows that there is no poset-block structure that turns the extended Hamming code $\overline{\mathcal{H}}(3)$ a 2-perfect code. Considering the ternary Golay code $\mathcal{G}_{12}$, they classified all the poset structures that may turn $\mathcal{G}_{12}$ into a 3-perfect poset code: $P$ has at most 3 elements in the first level and, up to isomorphism, if $j \in H_2(P)$, then $\langle j\rangle_P = \{1, j\}$.

## 7.2 Graph Metrics

We defined the Hasse diagram in Sect. 2.1 as a directed graph, where an edge $\overrightarrow{ab}$ connects $a$ to $b$ if, and only if, $a$ covers $b$. Looking at a poset from this point of view, it is surprising that it took a long lapse of time until poset metrics were generalized to graph metrics. This family of metrics were introduced in [34].

We consider a finite *directed graph* (or simply *digraph*) $G(V, E)$ consisting of a finite set of *vertices* $V = \{v_1, \ldots, v_n\}$ and a set of *directed edges* $E \subset V \times V$ (parallel edges are not allowed), an edge connecting $u$ (the *head*) to $v$ (the *tail*) is denoted by $e := \overrightarrow{uv}$. A *trail* of *length* $k$ is a sequence of edges $\overrightarrow{u_0u_1}, \overrightarrow{u_1u_2}, \ldots, \overrightarrow{u_{k-1}u_k}$ in which all edges are distinct. When $u_0 = u_k$, the trail is called a *circuit*. In case all the vertices $u_i$'s in a trail are distinct, except for the possibility that $u_0 = u_k$, the trail is called a *simple directed path of length* $k$, or simply a path. If $u_0 = u_k$ then the path is called a *directed cycle*.

If there is a trail from $u$ to $v$, we say that $u$ *dominates* $v$, and denote it by $u \to v$. A set $X \subset V$ is called a *closed set* if $u \in X$ and $u$ dominates $v \in V$ implies that $v \in X$. The *closure* $\langle X\rangle_G$ of a set $X \subset V$ is the smallest closed subset containing $X$. It is immediate to realize that, if $G(V, E)$ is the Hasse diagram of a poset over $V = [n]$, then the closure $\langle X\rangle_G$ of $X \subseteq [n]$ is the ideal generated by $X$. If $X = \{v\}$,

we denote $\langle\{v\}\rangle_G = \langle v\rangle_G$. This suggests the definition of a weight and a distance based on $G$.

We identify $V = \{v_1, \ldots, v_n\}$ with $[n] = \{1, 2, \ldots, n\}$ and define the $G$-weight $\varpi_G(\mathbf{x})$ of $\mathbf{x} \in \mathbb{F}_q^n$ as the number of vertices in $G$ dominated by the vertices in the support of $\mathbf{x}$:

$$\varpi_G(\mathbf{x}) := |\, \langle \mathrm{supp}(\mathbf{x})\rangle_G \,|.$$

The *G-distance* between $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$ is defined by $d_G(\mathbf{x}, \mathbf{y}) := \varpi_G(\mathbf{y} - \mathbf{x})$.

An edge $\overrightarrow{uv} \in E$ is called a *shortcut* if there exists a simple directed path from $u$ to $v$ which contains at least two edges. Adding or removing shortcuts to the graph do not affect the metric and this leads us to two canonical forms of a graph.

The *expanded canonical form* of $G(V, E)$ is the graph $G(V, E')$ obtained from $G$ by adding edges, one by one, such that each added edge is a shortcut. As an example, the expanded canonical form of a graph is a complete graph if, and only if, there is a circuit in $G$ containing all the vertices.

An *L-valued directed graph* $G_L(V, E)$ consist of a directed graph $G(E, V)$ and a *value function* $L : V \longrightarrow \mathbb{N}$. This gives rise to a further (unexplored) generalization, if we define

$$\varpi_{G_L}(\mathbf{x}) = \sum_{i \in \langle \mathrm{supp}(\mathbf{x})\rangle_G} L(i).$$

When is trivial, in the sense that $L(i) = 1$ for any $i \in [n]$, we get that $\varpi_{G_L}(\mathbf{x}) = \varpi_G(\mathbf{x})$, for all $\mathbf{x} \in \mathbb{F}_q^n$. Also, if $G$ is the Hasse diagram of a poset $P$ and $L$ is trivial, then $\varpi_{G_L}(\mathbf{x}) = \varpi_P(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{F}_q^n$, so we do have a generalization of poset metrics. The $L$-valued directed graph are necessary to determine a second canonical form.

The *reduced canonical form* of a graph $G(V, E)$ is a valued acyclic graph $G' = G_{L'}(V', E')$ obtained from $G$ as follows:

(1) Each maximal cycle of vertices in $V$ becomes a vertex in $V'$.

(2) To determine the set of edges we work in two steps. First we construct an edge connecting vertices $u_1, u_2 \in V'$ if there is an edge in $E$ connecting some vertex in the cycle represented by $u_1$ to some vertex in the cycle represented by $u_2$. Then we remove, one by one, all the shortcuts and obtain the set of edges $E'$.

(3) The value $L'$ of a vertex in $V'$ is the number of vertices in the corresponding cycle.

We remark that, since $G(V', E')$ has no cycle and no shortcuts, it is a Hasse diagram of a poset. The $G$-weight of $\mathbf{x} \in \mathbb{F}_q^n$ can be computed considering the canonical reduced form:

$$\varpi_G(\mathbf{x}) = \sum_{u \in \langle \pi(\mathrm{supp}(\mathbf{x}))\rangle_{G'}} L'(u).$$

It is not difficult to see that each canonical form is unique and defines the same metric as the original graph $G$. The reciprocal is also true:

**Theorem 7.1** *[34, Theorem 3 and Corollary 1] Given two graph metrics $G_1(V, E_1)$, $G_2(V, E_2)$, we have that $(\mathbb{F}_q^n, d_{G_1})$ and $(\mathbb{F}_q^n, d_{G_2})$ are isometric if, and only if, the canonical extended and reduced forms of $G_1$ and $G_2$ are isomorphic (as directed graphs).*

The canonical forms allow to explore some of the difficult combinatorial questions concerning graph metrics (see [34, Section IV]) and also to achieve some understanding about metric coding properties. The main questions worked up to the moment are the description of the structure of the group of linear isometries, necessary and sufficient conditions for the existence of a MacWilliams isometry and, somehow surprising, these conditions are not necessary for the validity of the extension property. We will now introduce, with some explanations but no proofs, these results.

### 7.2.1 Linear Isometries

We denote by $GL_G(\mathbb{F}_q^n)$ the group of linear isometries of $\mathbb{F}_q^n$ endowed with the metric $d_G$ induced by a digraph $G(V, E)$, with canonical extended form $\tilde{G}$. As we did for posets, we define two types of isometries:

**(i) Operating on Closures:** We say that an isometry $T$ operates on closures if each closed set in $V$ is mapped into itself, that is, if $\langle T(\mathrm{supp}(X))\rangle_G \subseteq \langle \mathrm{supp}(X)\rangle_G$ for every $X \subseteq \mathbb{F}_q^n$. If we write $T(\mathbf{e}_i) = \sum_{j=1}^n \alpha_{ij}\mathbf{e}_j$ we have that $T$ operates on closures if, and only if, *(i)* $\alpha_{ii} \neq 0$ for every $i \in [n]$ and *(ii)* $\alpha_{ij} \neq 0$ implies that $v_j \in \langle v_i\rangle_G$. The set of all isometries operating on closures is a group, denoted by $N_G$.

**(ii) Induced by Graph Automorphisms:** When considering isometries that operate on closures, we could consider either $G$ or the extended form $\tilde{G}$, since $\langle X\rangle_G = \langle X\rangle_{\tilde{G}}$ for every $X \subseteq V$. This is not the case for graph automorphisms. For example, if $G$ is a cycle, its extended form $\tilde{G}$ is a complete graph and $\mathrm{Aut}(G)$ is a cyclic group while $\mathrm{Aut}(\tilde{G})$ is the permutation group $S_n$. So, we consider the group $\mathrm{Aut}(\tilde{G})$ acting by permutation of the coordinates, as in the poset case. It is not difficult to see that each $\phi \in \mathrm{Aut}(\tilde{G})$ determines an isometry $T_\phi$ of $\mathbb{F}_q^n$ with the metric $d_G$.

Similarly to the poset case, it is possible to prove [34, Theorem 8] that $GL_G(\mathbb{F}_q^n) \simeq N_G \ltimes \mathrm{Aut}(\tilde{G})$.

### 7.2.2 Transitivity on Spheres

Let $G'(V', E')$ be the reduced canonical form of $G$. Since $G'$ is the Hasse diagram of a poset, we may say that $G$ is of *hierarchical type* if $G'$ determines a hierarchical poset. Being of a hierarchical type is enough to ensure a kind of canonical decomposition of a code, but not for a transitive action of $GL_G(\mathbb{F}_q^n)$ on spheres, for this we need another property. Let $V_1' \cup \ldots \cup V_l'$ be the level decomposition of $V'$. We say that $G$ satisfies the *unique decomposition property* (UDP) if, given subsets $S, S' \subset V_i'$ satisfying $\sum_{a\in S} L'(a) = \sum_{b\in S'} L'(b)$, then there is a bijection $g : S \to S'$ such that $L'(a) = L'(g(a))$ for all $a \in S$.

We conclude that $GL_G(\mathbb{F}_q^n)$ acts transitively on spheres if, and only if, $G$ is of hierarchical type and satisfies the UDP [34, Proposition 2]. To see that the UDP is a necessary condition, we may consider the graph $G(V, E)$ with $V = \{v_1, v_2, v_3, v_4\}$ and $E = \{\overrightarrow{v_1v_2}, \overrightarrow{v_2v_1}\}$. It is a graph of hierarchical type which does not satisfy the UDP. Considering $\mathbf{x} = (1000)$ and $\mathbf{y} = (0011)$, we have that $\varpi_G(\mathbf{x}) = \varpi_G(\mathbf{y}) = 2$ but there is no $T \in GL_G(\mathbb{F}_q^n)$ such that $T(\mathbf{x}) = \mathbf{y}$.

### 7.2.3 MacWilliams' Identity

As in the case of poset, to obtain a MacWilliams identity we consider the $G$-weight distribution of a code $C$ and, for its dual code $C^\perp$ we need to consider the *opposite graph* $G^* = (V, E^*)$, where $E^*$ is defined by $\overrightarrow{uv} \in E^*$ if, and only if, $\overrightarrow{vu} \in E$. There is no general condition for the validity of a MacWilliams identity, but we do know one for digraphs of hierarchical type: if $G$ is of hierarchical type, then the metric $d_G$ admits a MacWilliams identity if, and only if, $G$ satisfies the UDP. The proof is technical (using characters theory) and can be found in [73, Theorem 3]. To see that the UDP is indeed necessary, we may consider $G$ as in Sect. 7.2 and let $C_1 = \{0000, 1000\}$ and $C_2 = \{0000, 0011\}$. Direct computations shows that $W_{C_1}^G(x) = W_{C_2}^G(x) = 1 + x^2$ while $W_{(C_1)^\perp}^{G^*}(x) = 1 + 2x + 2x^2 + 2x^3 + x^4$ and $W_{(C_2)^\perp}^{G^*}(x) = 1 + 4x^2 + 3x^3$.

### 7.2.4 Extension Property

As for the MacWilliams property, there is no general condition for the validity of the extension property and we restrict ourself to graphs of hierarchical type. The UDP is still a necessary condition for the extension property, but to get sufficiency we need an additional (and very restrictive) condition (called the $\Omega$ condition in [35]): given $k > 1$, there are at most two elements $u, v \in V'$ such that $L'(u) = L'(v) = k$, where $G'(V', E')$ is the canonical reduced form of $G(V, E)$. One should remark that this contradicts the conjecture presented at the end of [73]: admission of a MacWilliams

identity and extension property are not equivalent properties. Furthermore, the statement made in that work about the extension property (Theorem 4, with no proof) is not valid, lately corrected in [35] with the addition of the $\Omega$-condition.

As a final remark, Hyun et al. started working on specific of codes with graph metrics. In [53] they classified all graph metrics that turn the $[8, 4]$-extended Hamming code into a 2-perfect code and found a family of graph metrics that turns any extended Hamming code into a 2-perfect code

## 7.3 Pomset Metrics

In 1958, C.Y. Lee, in [67, 1958] introduced what became known as the *Lee metric*: given $\alpha \in \mathbb{Z}_m$, the Lee-norm is $\|\alpha\|_L = \min\{\alpha, m - \alpha\}$ and the Lee weight $\varpi_L$ of $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_m^n$ is $\varpi_L(\mathbf{x}) = \sum_i \|x_i\|_L$. In 1971, Chiang and Wolf described all the discrete, memoryless, symmetric channels matched to the Lee metric [16]. To avoid unnecessary new definitions, in this section we consider $q$ to be prime, so that $\mathbb{F}_q \sim \mathbb{Z}_q$.

The poset codes does not generalize the Lee metric, but partial orders over multiset does it, as understood by Sudha and Selvaraj [45, 2017]. All the content of this section refers to [45].

Multisets is a generalization of the concept of a set that, unlike a set, allows multiple instances of the multiset's elements, and it had been used earlier in the context of coding theory [24, 97] and also with poset codes [81], mainly to explore duality issues, as an alternative to the use of matroids.

We can formalize it by considering a *multiset* (or simply *mset*) as a pair $M = (X, \mathfrak{c})$, where $\mathfrak{c} : X \to \mathbb{N}$ is the *counting function*: $\mathfrak{c}(i)$ counts the number of occurrences of $i$ in $M$. We remark that we allow $\mathfrak{c}(i) = 0$, and we define the cardinality of $M$ as $|M| = \sum_{x \in X} \mathfrak{c}(x)$. We denote by $M = \{k_1/a_1, \ldots, k_n/a_n\}$ the multiset with $\mathfrak{c}(a_i) = k_i$. Let us denote by $\mathcal{M}^m(X)$ the set of all multisets underlying $X$ such that any element occurs at most $m$ times.

To define an order relation on a multiset one should consider the *cartesian product of multisets* $M_1(X, \mathfrak{c}_1)$ and $M_2(X, \mathfrak{c}_2)$, it is the mset with underlying set $X \times X$ and counting function $\mathfrak{c}$ defined as the product: $\mathfrak{c}((a, b)) = \mathfrak{c}_1(a)\mathfrak{c}_2(b)$, that is,

$$M_1 \times M_2 := \{mn/(m/a, n/b); m/a \in M_1, n/b \in M_2\}.$$

The *sum* $M_1 \oplus M_2 = (X, \mathfrak{c})$ is the mset with underlying set $X$ defined by the counting function $\mathfrak{c}(a) = \mathfrak{c}_1(a) + \mathfrak{c}_2(a)$.

A submultiset (*submset*) of $M = (X, \mathfrak{c})$ is a multiset $S = (X, \mathfrak{c}')$ such that $\mathfrak{c}'(x) \leq \mathfrak{c}(x)$, for all $x \in X$. We denote it by $S \ll M$. A submset $R = (M \times M, \mathfrak{c}_\times) \ll M \times M$ is called a *mset relation* on $M$ if $\mathfrak{c}_\times(m/a, n/b) = mn$.

**Definition 7.2** Let $M$ be a multiset. A *partially ordered mset relation* $R$ (or *pomset relation*) on $M$ is a mset relation satisfying:

1. $(m/a)R(m/a)$, for all $m/a \in M$ (*reflexivity*);
2. $(m/a)R(n/b)$ and $(n/b)R(m/a)$ implies $m = n$, $a = b$ (*antisymmetry*);
3. $(m/a)R(n/b)R(k/c)$ implies $(m/a)R(k/c)$ (*transitivity*).

The pair $\mathbb{P} = (M, R)$, where $M$ is a mset and $R$ is a pomset relation is called a *partially ordered mset* (or *pomset*). Given $\mathbb{P} = (M, R)$, a subset $I \ll M$ is called a *pomset ideal* if $m/a \in I$ and $(n/b)R(k/a)$, with $k > 0$ and $b \neq a$ implies $n/b \in I$. Given a submset $S \ll M$, we denote by $\langle S \rangle_{\mathbb{P}}$ the ideal generated by $S$, that is, the smallest ideal of $\mathbb{P}$ containing $S$.

*Example 7.3* Let $X = \{1, 2, 3, 4\}$ and $M = \{3/1, 3/2, 3/3, 3/4\}$ be a pomset with underlying set $X$. The set

$$R = \left\{ \begin{array}{ll} 9/(\,3/1\,,\,3/1\,), & 9/(\,3/2\,,\,3/2\,), \\ 9/(\,3/3\,,\,3/3\,), & 9/(\,3/4\,,\,3/4\,), \\ 9/(\,3/2\,,\,3/1\,), & 9/(\,3/3\,,\,3/4\,) \end{array} \right\}$$

is a pomset on $M$. The set $S = \{1/1, 2/3\}$ is a submset of $M$. To turn it into a pomset ideal we should look on $R$ for every element of the form $1 \cdot k/(1/1, k/a)$ or $2 \cdot k/(2/3, k/a)$. There is a unique element like that, namely $9/(3/2, 3/1)$, so the pomset generated by $S$ is $\langle S \rangle_{\mathbb{P}} = \{1/1, 2/3, 3/1\}$.

We can now define a pomset metric on $\mathbb{F}_q^n$. We consider the mset $M = \{r/1, r/2, \ldots, r/n\} \in \mathcal{M}^r([n])$ where $r := \lfloor q/2 \rfloor$ is the integer part of $q/2$. With this notation, we define the *Lee support* of a vector $\mathbf{x} \in \mathbb{F}_q^n$ as $\mathrm{supp}_L(\mathbf{x}) = \{k/i;\, k = \|x_i\|_L, k \neq 0\}$. Let $\mathbb{P} = (M, R)$ be a pomset. The $\mathbb{P}$-weight and $\mathbb{P}$-distance on $\mathbb{F}_q^n$ are defined as

$$\varpi_{\mathbb{P}}(\mathbf{x}) := |\langle \mathrm{supp}_L(\mathbf{x}) \rangle_{\mathbb{P}}| \quad \text{and} \quad d_{\mathbb{P}}(\mathbf{x}, \mathbf{y}) := \varpi_{\mathbb{P}}(\mathbf{x} - \mathbf{y}),$$

respectively. It is not difficult to prove that $d_{\mathbb{P}}$ is a metric. The positivity condition is trivial. The symmetry condition follows from the fact that $\|\alpha\| = \| - \alpha\|$. To prove the triangular inequality we need to observe that, given ideals $I, J \ll \mathcal{M}^r([n])$, we have that $I \oplus J$ is an ideal and the ideal $\langle I \oplus J \rangle_{\mathbb{P}}$ is contained in $I \oplus J$. It follows that $\varpi_H \mathbb{P}(\mathbf{x} + \mathbf{y}) \leq |\langle \mathrm{supp}_L(\mathbf{x}) \oplus \mathrm{supp}_l(\mathbf{y}) \rangle_{\mathbb{P}}| \leq |\langle \mathrm{supp}_L(\mathbf{x}) \rangle_{\mathbb{P}} \oplus \langle \mathrm{supp}_l(\mathbf{y}) \rangle_{\mathbb{P}}| \leq |\langle \mathrm{supp}_L(\mathbf{x}) \rangle_{\mathbb{P}}| + |\langle \mathrm{supp}_l(\mathbf{y}) \rangle_{\mathbb{P}}| = \varpi_{\mathbb{P}}(\mathbf{x}) + \varpi_{\mathbb{P}}(\mathbf{y})$.

A linear code $C \subseteq \mathbb{F}_q^n$, considered with a pomset metric $d_{\mathbb{P}}$, is called a $\mathbb{P}$-linear code with parameters $[n, k, \delta_{d_{\mathbb{P}}}]_q$, where $\delta_{d_{\mathbb{P}}}$ is the minimum $d_{\mathbb{P}}(C)$-distance between different codewords.

In case the pomset is an *anti-chain*, that is, any two distinct pair of points $m/a, n/b \in M$ with $a \neq b$ are not comparable (neither $(m/a)R(n/b)$ nor $(n/b)R(m/a)$), we have that $\langle \mathrm{supp}_L(\mathbf{x}) \rangle_{\mathbb{P}} = \mathrm{supp}_L(\mathbf{x})$ and so, $\varpi_{\mathbb{P}}(\mathbf{x}) =$

$\sum_i \|x_i\|_L = \varpi_L(\mathbf{x})$, therefore the pomset metric is a generalization of the Lee metric.

In the seminal work [45], the authors generalize for pomsets the basic operations known for posets: direct and ordinal sum, direct and ordinal products, denoted by $\mathbb{P}_1 \oplus \mathbb{P}_2$, $\mathbb{P}_1 + \mathbb{P}_2$, $\mathbb{P}_1 \otimes \mathbb{P}_2$ and $\mathbb{P}_1 \times \mathbb{P}_2$, respectively. Then, given $C_i$ an $[n_i, k_i, \delta_{d_{\mathbb{P}_i}}]_q$ linear $\mathbb{P}_i$-code where $i \in \{1, 2\}$, they established the following:

1. *Direct sum of codes*: the minimum distance of $C_1 \oplus C_2$ in terms of $\delta_{d_{\mathbb{P}_1}}$ and $\delta_{d_{\mathbb{P}_2}}$, considering the posets $\mathbb{P}_1 \oplus \mathbb{P}_2$ and $\mathbb{P}_1 + \mathbb{P}_2$;
2. $(\mathbf{u}|\mathbf{u}+\mathbf{v})$-*construction*: the minimum distance of the code

$$C_1 \,{}_{(\mathbf{u}|\mathbf{u}+\mathbf{v})}^{\oplus} C_2 = \{(\mathbf{u}, \mathbf{u}+\mathbf{v}); \mathbf{u} \in C_1, \mathbf{v} \in C_2\}$$

   obtained by the $(\mathbf{u}|\mathbf{u}+\mathbf{v})$-construction, considering the posets $\mathbb{P}_1 \oplus \mathbb{P}_2$ and $\mathbb{P}_1 + \mathbb{P}_2$;
3. *Tensor product*: bounds for the minimum distance of the tensor product $C_1 \otimes C_2$, for the direct product $\mathbb{P}_1 \otimes \mathbb{P}_2$ and the ordinal product $\mathbb{P}_1 \times \mathbb{P}_2$ of posets.

This is essentially what is known about pomset codes: the road is open to be paved by interested researchers.

## 7.4 Combinatorial Metrics: An Opposite Direction

Besides the poset metrics and its generalizations presented in this chapter (poset-block, graph and pomset metrics), there is another family of metrics which satisfies the conditions stated in Sect. 1.1, that is, metrics assuming integer values, determined by a weight and respecting the support of vectors: the family of combinatorial metrics. This family of metrics was introduced in 1973 by Gabidulin [41] and rested untouched until it was recently recalled in a survey of 2012 [42].

It is important to note that the family of combinatorial metrics is a very large family, but its intersection with the family of poset metrics (and its generalizations) is nearly trivial: the only poset metric that is also a combinatorial metric is the Hamming metric. This is the reason we say this is a generalization of the Hamming metric in a direction opposite to the one given by the poset metrics.

In this section we shall define the combinatorial metric and just quote some results concerning the main invariants we explored in this text (minimum distance and its bounds, MacWilliams identity and extension property). As we shall see from the strongly restrictive conditions needed to ensure a MacWilliams type identity or the extension property, working with this metrics suggests to be a hard task.

Let $\mathcal{P}([n]) := \{A; A \subset [n]\}$ be the power set of $[n]$. We say that a family $\mathcal{A} \subset \mathcal{P}([n])$ is a *covering* of a set $X \subset [n]$ if $X \subset \cup_{A \in \mathcal{A}} A$. If $\mathcal{F}$ is a covering of $[n]$, then the $\mathcal{F}$-*combinatorial weight* of $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{F}_q^n$ is the integer-valued map $\varpi_{\mathcal{F}}$ defined by

$$\varpi_{\mathcal{F}}(\mathbf{x}) := \min\{|\mathcal{A}|;\ \mathcal{A} \subset \mathcal{F} \text{ and } \mathcal{A} \text{ is a covering of } \operatorname{supp}(\mathbf{x})\}.$$

Each element $A \in \mathcal{F}$ is called a *basic set* of the covering.

As showed in [41], the function $d_{\mathcal{F}} : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{N}$ defined by

$$d_{\mathcal{F}}(\mathbf{x}, \mathbf{y}) := \varpi_{\mathcal{F}}(\mathbf{x} - \mathbf{y})$$

satisfies the metric axioms and is called an $\mathcal{F}$-*combinatorial metric*.

Some particular cases of combinatorial metrics, depending on the covering $\mathcal{F}$, had been studied in the literature under a different name. These include:

1. The *b-burst metric*, where $\mathcal{F}_{[b]} = \{[b], [b]+1, [b]+2, \ldots, [b]+(n-b)\}$, with $[b]+i := \{1+i, 2+i, \ldots, b+i\}$; suitable for decoding burst errors [41].
2. The block metrics defined in [37], where $\mathcal{F}$ is a partition of $[n]$. We remark that this is a particular case of a block metric (as we introduced in Sect. 7.1). It is worth noting that this is all the intersection between the family of poset-block (or graph) metrics and combinatorial metrics. Among the partitions, those which have all parts of equal size is a relatively rare but significant case, known as *k-partition*.
3. The metrics determined by a covering $\mathcal{F}(n,k)$ consisting of all the subsets of $[n]$ with cardinality $k$. As we shall see, these coverings will play a relevant role concerning the extension property.

### 7.4.1 Singleton Bound

As we noted, particular instances of combinatorial metrics have been studied before the general approach proposed by Gabidulin. The Singleton bound for the class of burst metric was given by Reiger in 1960 [95]. A general approach (concerning combinatorial metrics) is given in [11]: considering an $[n,k]_q$ linear code $C$ and a covering $\mathcal{F}$, its minimum distance $\delta_{d_{\mathcal{F}}}$ is bounded by

$$n\frac{\delta_{d_{\mathcal{F}}} - 1}{D} \leq \left\lceil n\frac{\delta_{d_{\mathcal{F}}} - 1}{D} \right\rceil \leq n - k,$$

where $\lceil \cdot \rceil$ is the ceiling function and $D$ is the minimum number of basic sets needed to cover $[n]$, that is, the maximum of $\varpi_{\mathcal{F}}(\mathbf{x})$ for $\mathbf{x} \in \mathbb{F}_q^n$. We remark that $D \leq n$ and equality in the bounds holds if, and only if, $\mathcal{F}$ consists only of singletons, or equivalently, $d_{\mathcal{F}}$ is the Hamming metric, and, in this case, we get the usual Singleton bound. We also note that the result in [11] is stated for the case of general codes, not necessarily linear.

### 7.4.2 MacWilliams' Identity

The MacWilliams identity was studied in the particular case of block metrics, i.e., when $\mathcal{F}$ is a partition ([37]), where it was proved that $d_{\mathcal{F}}$ satisfies a MacWilliams identity if and only if $\mathcal{F}$ is a $k$-partition, that is, all the parts has size equal to some $k$. In [91] it was proved that if there are basic sets with non-empty intersection, then there is a code for which the MacWilliams identity does not hold. So, in order to satisfy a MacWilliams type identity, $\mathcal{F}$ must be a partition, hence a $k$-partition. We remark that, for poset metric, to admit a MacWilliams identity means that the poset is hierarchical and to construct hierarchical posets is much easier than to construct $k$-partitions: a hierarchical poset is given once we decompose $n$ as a sum $n = n_1 + n_2 + \cdots + n_l$ while a $k$-partition arises once we decompose $n$ as a product $n = k \cdot m$.

### 7.4.3 Extension Property

The extension property for combinatorial metrics carries an interesting surprise, since the conditions for its validity are not the same as for the MacWilliams identity. These conditions are fully understood only in the case that $\mathcal{F}$ is not connected and $\mathbb{F}_q = \mathbb{F}_2$.

Given a covering $\mathcal{F}$, a *path of basic sets* is a sequence of basic sets, each of them intersecting the previous one. The covering is said to be *$\mathcal{F}$-connected* if given $r, s \in [n]$ there is a path of basic sets $A_1, A_2, \ldots, A_l \in \mathcal{F}$ such that $r \in A_1$ and $s \in A_l$, otherwise, it is called *$\mathcal{F}$-disconnected*. The connected components of $\mathcal{F}$ are the maximal connected subsets of $\mathcal{F}$. A disconnected covering $\mathcal{F}$ satisfies the extension property if, and only if, either $\mathcal{F}$ has exactly two connected components and it is a $k$-partition or $l > 2$ and $d_{\mathcal{F}}$ is the Hamming metric [91]. For the connected case, all is known is that the validity of the extension property implies that $\mathcal{F} = \mathcal{F}(n, k)$, that is, $\mathcal{F}$ contains every subset of $[n]$ with $k$ elements. Whether $\mathcal{F}(n, k)$ satisfies the extension property for every $k$ it is still unknown.

## 7.5 Epilogue: Approximation of Channels and Metrics

From the engineering point of view, there are many reasons why the use of a poset metric (or its generalizations) may be interesting. The most immediate reason is the possibility to match a metric to a channel. The relation between channels and metrics, which opened this book (and later was mentioned only in Chap. 5) should also close the book, since the possibility to use poset-metrics (and other families of metrics) to match or approximate a channel is the key to a prospective promising way that motivates us, part of a broader program that we briefly outline here.

A channel, as we saw in the beginning of this text, is a probabilistic model, which defines a probabilistic decoding criteria (maximum likelihood or maximum a priori decoding). In the same way, a metric also determines a decoding criteria (minimum distance). However, due to imprecisions in the measurement of the channel properties (physical experiments that should be performed) or due to algorithm issues (ML decoding may be too difficult for the available resources), it may be interesting to use another channel model to establish the decoding process. This is known in the literature as *mismatched decoding* (see for example [43] for the subject). This is an interesting research subject, much studied from the point of view of information theory, interested mainly in what is possible to achieve asymptotically, that is, when the length of a code (the dimension of $\mathbb{F}_q^n$) goes to infinity.

We are concerned with the asymptotic aspects, but on the finite length regime, that is, what can be done given $n$. This approach poses some structures and questions that are new from the mathematical point of view and which we believe to be interesting.

Considering a channel as a set of possible messages (that is, we are considering the input and output sets $\mathcal{X}$ and $\mathcal{Y}$ to be $\mathbb{F}_q^n$ and not $\mathbb{F}_q$), we say that two channels $\mathbb{P}_1$ and $\mathbb{P}_2$ are *decoding-equivalent* if they determine the same decoding criteria for every code $C \subset \mathcal{X}$, that is,

$$\operatorname*{argmax}_{\mathbf{x}\in C} P_1(\mathbf{y}|\mathbf{x}) = \operatorname*{argmax}_{\mathbf{x}\in C} P_2(\mathbf{y}|\mathbf{x}), \forall\, C \subseteq \mathcal{X} \text{ and } \mathbf{y} \in \mathcal{X}.$$

We denote $N = |\mathcal{X}|$. Since $\mathbb{P} = (P_{\mathbf{y},\mathbf{x}})_{\mathbf{y},\mathbf{x}\in\mathcal{X}}$, we may identify a channel with an $N \times N$ matrix with non-negative real entries (we can, if necessary, normalize each column to attain $\sum_{\mathbf{y}\in[N]} P_{\mathbf{y},\mathbf{x}} = 1$). Up to rescaling, we can call any such matrix a channel, so that the space of all channels over a set $\mathcal{X}$ with $N$ elements is identified with $\mathbb{R}_+^{N\times N}$ where $\mathbb{R}_+$ is the set of non-negative reals. We denote the decoding-equivalence relation by $\underset{dec}{\sim}$, the equivalence class of a matrix $A \in \mathbb{R}_+^{N\times N}$ is denoted by $[A]_{dec}$ and the space of equivalence classes is denoted by $\mathrm{Chan}(N) = \mathbb{R}_+^{N\times N}/\underset{dec}{\sim}$.

If we wish to develop an "approximation theory" of channels, we should address three main aspects: (1) the structure of $\mathrm{Chan}(N)$; (2) the knowledge about the behavior of specific subsets $A_{\mathtt{good}} \subset \mathrm{Chan}(N)$ (the subset is as "good" as deep is our understanding of its properties) which can be used to approximate channels that are more difficult to handle and (3) to estimate how large $A_{\mathtt{good}}$ is in $\mathrm{Chan}(N)$.

We remark that this definition of equivalence relation, when restricted to the subset of matrices in $\mathbb{R}_+^{N\times N}$ corresponding to the distances of a metric (a specific type of channel) differs from the usual definition of *scalar-equivalence*: two metrics $d_1$ and $d_2$ over $\mathcal{X}$ are said to be scalar-equivalent if they differ by a non-zero multiplicative constant, that is, $d_1 \underset{scal}{\sim} d_2$ if there is $\lambda > 0$ such that $d_1(\mathbf{x},\mathbf{y}) = \lambda d_2(\mathbf{x},\mathbf{y})$ for all $\mathbf{x},\mathbf{y} \in \mathcal{X}$. In our case, each equivalence relation is much larger and the quotient space much smaller. As an example, if $N = 3$, the possible

distances determine the sides of a triangle. In the usual situation there are infinitely many scalar equivalence classes: two metrics are scalar-equivalent if, and only if, the triangles are similar (in the usual euclidean sense). In our case there are only four equivalence classes: the class determined by equilateral triangles, two classes determined by isosceles (non-equilateral) triangles (depending on the equal edges being the larger or the smaller one) and the class determined by triangles with three different edges.

We give a very brief review of what is known about those issues (actually not too much).

### 7.5.1 The Structure of Chan$(N)$

It is known [27] that each equivalence class $[A] \in \mathrm{Chan}(N)$ is a simplicial cone in $\mathbb{R}_+^{N\times N}$. Moreover, $\mathrm{Chan}(N)$ carries a structure known in the literature as an *hyperplane arrangement* ([30]).

In order to have a figure of merit about how close two channels are, in ([30]) a metric on $\mathrm{Chan}(N)$ playing this role was defined, in the sense that the closer the channels, the smaller is the probability of improper decoding. A manageable expression (as manageable as an expression involving $N \times N = q^n \times q^n$ elements can be) for this distance is presented, separating the role of the equivalence classes of channels and of the actual channel itself.

Also, each $[A] \in \mathrm{Chan}(N)$ have a representative $B \in [A]$ that can be isometrically embedded into a Hamming cube $\mathbb{F}_2^M$ for $M$ sufficiently large (see [28, 29]). Moreover, if $B$ is invariant by translations, this embedding is a linear map. It is not known how to determine such a minimal $M$, since the construction done in [28] does not ensures the minimality of $M$. However, it seems that the closer $[A]$ is to the Hamming metric, the smaller is $M$. This can be an interesting issue if we wish to approximate a channel by the symmetric memoryless channel (or equivalently, by the Hamming minimum distance decoder).

Finally, the study of metrics with the $\underset{scal}{\sim}$ equivalence is a subject vastly studied in the literature. Many questions, similar to the ones explored in [23] are interesting and relevant on its own sake.

### 7.5.2 Good Channels and Metrics

The main reason we approximate something by something else is that the "something else" is better understood then the original object. In this sense, the subject of this text is understanding (and qualifying) the candidates for $A_{\mathtt{good}}$. Of course, the hierarchical poset metric are very good candidates, but any poset metric (or its generalizations) may be seen as reasonable ones.

### 7.5.3 How Large the Manageable Metric Channels Are

The concept of mangeable metric is a fuzzy one, depending on the goals and difficulties posed by concrete problems. Nevertheless, having $[A] \in \mathrm{Chan}(N)$ determined by a metric gives, in advance, an advantage: the fact that a metric is symmetric automatically reduces all computations to half the original size. However, it is known that not every $[A] \in \mathrm{Chan}(N)$ is determined by a metric.[1]

The family of poset metrics (and its generalizations) is a very large family (with exponential growth.[2]) Even the family of hierarchical posets grows very fast.[3] Also the set of combinatorial metrics grows exponentially with $n$. Nevertheless, none of these sets of metrics is sufficient to describe all the equivalence classes in $\mathrm{Chan}(N)$, not even those determined by a metric.

Besides considering the union of poset and combinatorial metrics, one can combine them, for example by summing the distances: $\varpi_{P\oplus\mathcal{F}}(\mathbf{x}) := \varpi_P(\mathbf{x}) + \varpi_{\mathcal{F}}(\mathbf{x})$. This is also not enough: there are examples of metrics over $\mathbb{F}_q^n$ (with very small values of $n$) which can not be attained by such a procedure. Another option is to make conditional sums, for example $\varpi_{P\oplus_r\mathcal{F}}(\mathbf{x}) := \varpi_P(\mathbf{x})$ if $|\langle \mathrm{supp}(\mathbf{x})\rangle_P| \leq r$ and $\varpi_{P\oplus\mathcal{F}}(\mathbf{x}) := \varpi_P(\mathbf{x}) + \varpi_{\mathcal{F}}(\mathbf{x})$ if $|\langle \mathrm{supp}(\mathbf{x})\rangle_P| > r$. Other conditions can be imposed somehow similar to the conditions determining the different types of MacWilliams' relations in Sect. 5.2. Another possibility follows the guidelines of studying the geometry of cuts (see [23]): given a subset $A \subset \mathrm{Chan}(N)$, one should consider all the representatives in $\mathbb{R}_+^{N\times N}$ of the classes in $A$ and then, to consider the convex hull and look at the linear combination of elements represented in $A$.

All this is a long and winding road, to be traced by many researchers.

[1] See [40] for a necessary condition and [29] for an algorithm that either determines a metric matched to a channel or determines the nonexistence of such a metric.

[2] The number $f(n)$ of posets over $n$ grows as $f(n) = 2^{\frac{n^2}{4}+o(n^2)}$.

[3] The number $p(n)$ of hierarchical posets is the number of partitions of $n$ which behaves as $\frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$ for $n \to \infty$.

# References


1. J. Ahn, H.K. Kim, J.S. Kim, M. Kim, Classification of perfect linear codes with crown poset structure. Discrete Math. **268**(1–3), 21–30 (2003)
2. M.M.S. Alves, A standard form for generator matrices with respect to the Niederreiter-Rosenbloom-Tsfasman metric, in *2011 IEEE Information Theory Workshop, ITW 2011* (2011), pp. 486–489
3. M.M.S. Alves, L. Panek, M. Firer, Error-block codes and poset metrics. Adv. Math. Commun. **2**(1), 95–111 (2008)
4. A. Barg, M. Firer, Translation association schemes, poset metrics, and the shape enumerator of codes, in *IEEE International Symposium on Information Theory Proceedings* (2012), pp. 101–105
5. A. Barg, W. Park, On linear ordered codes. Moscow Math. J. **15**, 679–702 (2015)
6. A. Barg, P. Purkayastha, Bounds on ordered codes and orthogonal arrays, in *IEEE International Symposium on Information Theory Proceedings* (2007), pp. 331–335
7. A. Barg, P. Purkayastha, Bounds on ordered codes and orthogonal arrays. Moscow Math. J. **9**(2), 211–243 (2009)
8. A. Barg, P. Purkayastha, Near MDS poset codes and distributions, in *IEEE International Symposium on Information Theory - Proceedings* (2010), pp. 1310–1314
9. A. Barg, L.V. Felix, M. Firer, M.V.P. Spreafico, Linear codes on posets with extension property. Discrete Math. **317**, 1–13 (2014)
10. A. Barra, H. Gluesing-Luerssen, MacWilliams extension theorems and the local–global property for codes over Frobenius rings. J. Pure Appl. Algebra **219**(4), 703–728 (2015)
11. M. Bossert, V. Sidorenko, Singleton-type bounds for blot-correcting codes. IEEE Trans. Inf. Theory **42**(3), 1021–1023 (1996)
12. T. Britz, K. Shiromoto, A MacWilliams type identity for matroids. Discrete Math. **308**(20), 4551–4559 (2008)
13. T. Britz, T. Johnsen, D. Mayhew, K. Shiromoto, Wei-type duality theorems for matroids. Des. Codes Crypt. **62**(3), 331–341 (2012)
14. R.A. Brualdi, J.S. Graves, K.M. Lawrence, Codes with a poset metric. Discrete Math. **147**(1–3), 57–72 (1995)
15. A.G. Castoldi, E.L.M. Carmelo, The covering problem in Rosenbloom-Tsfasman spaces. Electron. J. Comb. **22**(3), 1–18 (2015)
16. J.C. Chiang, J.K. Wolf, On channels and codes for the Lee metric. Inf. Control **19**(2), 159–173 (1971)




M. Firer et al., *Poset Codes: Partial Orders, Metrics and Coding Theory*, SpringerBriefs in Mathematics, https://doi.org/10.1007/978-3-319-93821-9

17. S.H. Cho, D.S. Kim, Automorphism group of the crown-weight space. Eur. J. Comb. **27**(1), 90–100 (2006)
18. S. Choi, J.Y. Hyun, H.K. Kim, D.Y. Oh, MacWilliams-type equivalence relations (2012, preprint). arXiv:1205.1090
19. J.H. Conway, N.J.A. Sloane, E. Bannai, *Sphere-packings, Lattices, and Groups* (Springer, New York, 1987)
20. B.K. Dass, N. Sharma, R. Verma, Perfect codes in poset spaces and poset block spaces. Finite Fields Appl. **46**, 90–106 (2017)
21. P. Delsarte, Bounds for unrestricted codes by linear programming. Philips Res. Rep. **27**, 272–289 (1972)
22. P. Delsarte, An algebraic approach to the association schemes of coding theory. Philips Res. Rep. **10** (1973)
23. M.M. Deza, M. Laurent, *Geometry of Cuts and Metrics* (Springer, Berlin, 1997)
24. S. Dodunekov, J. Simonis, Codes and projective multisets. Electron. J. Comb. **5**, 1–23 (1998)
25. R.G.L. D'Oliveira, M. Firer, The packing radius of a code and partitioning problems: the case for poset metrics, in *IEEE International Symposium on Information Theory - Proceedings* (2014), pp. 2954–2958
26. R.G.L. D'Oliveira, M. Firer, The packing radius of a code and partitioning problems: the case for poset metrics on finite vector spaces. Discrete Math. **338**(12), 2143–2167 (2015)
27. R.G.L. D'Oliveira, M. Firer, Geometry of communication channels: metrization and decoding. Symmetry Culture Sci. **27**, 279–289 (2016)
28. R.G.L. D'Oliveira, M. Firer, Minimum dimensional Hamming embeddings. Adv. Math. Commun. **11**, 359–366 (2017)
29. R.G.L. D'Oliveira, M. Firer, Channel metrization. Eur. J. Comb. (2018) Available online 5 March 2018
30. R.G.L. D'Oliveira, M. Firer, A distance between channels: the average error of mismatched channels (2018). arXiv e-prints
31. S. Dougherty, K. Shiromoto, Maximum distance codes in $Mat_{n,s}(Z_k)$ with a non-Hamming metric and uniform distributions. Des. Codes Crypt. **33**(1), 45–61 (2004)
32. S.T. Dougherty, M.M. Skriganov, MacWilliams duality and the Rosenbloom-Tsfasman metric. Moscow Math. J. **2**(1), 81–97 (2002)
33. S.T. Dougherty, M.M. Skriganov, Maximum distance separable codes in the $\rho$ metric over arbitrary alphabets. J. Algebraic Comb. **16**, 71–81 (2002)
34. T. Etzion, M. Firer, Metrics based on finite directed graphs, in *2016 IEEE International Symposium on Information Theory (ISIT)* (2016), pp. 1336–1340
35. T. Etzion, M. Firer, R.A. Machado, Metrics based on finite directed graphs and coding invariants (2017). arXiv:1609.08067v3
36. L.V. Felix, M. Firer, Canonical-systematic form for codes in hierarchical poset metrics. Adv. Math. Commun. **6**(3), 315–328 (2012)
37. K. Feng, L. Xu, F.J. Hickernell, Linear error-block codes. Finite Fields Appl. **12**(4), 638–652 (2006)
38. C. Feyling, Punctured maximum distance separable codes. Electron. Lett. **29**(5), 470–471 (1993)
39. M. Firer, J.A. Pinheiro, Bounds for complexity of syndrome decoding for poset metrics, in *2015 IEEE Information Theory Workshop (ITW)* (2015), pp. 1–5
40. M. Firer, J.L. Walker, Matched metrics and channels. IEEE Trans. Inf. Theory **62**(3), 1150–1156 (2016)
41. E.M. Gabidulin, Combinatorial metrics in coding theory, in *2nd International Symposium on Information Theory* (Akadémiai Kiadó, 1973)
42. E.M. Gabidulin, A brief survey of metrics in coding theory, in *Mathematics of Distances and Applications (MDA)* (2012), pp. 66–84
43. A. Ganti, A. Lapidoth, I.E. Telatar, Mismatched decoding revisited: general alphabets, channels with memory, and the wide-band limit. IEEE Transa. Inf. Theory **46**(7), 2315–2328 (2000)

44. H. Gluesing-Luerssen, Fourier-reflexive partitions and MacWilliams identities for additive codes. Des. Codes Crypt. **75**(3), 543–563 (2015)
45. I. Gnana Sudha, R.S. Selvaraj, Codes with a pomset metric and constructions. Des. Codes Crypt. (2017)
46. J.N. Gutiérrez, H. Tapia-Recillas, A MacWilliams identity for poset-codes, in *Congressus Numerantium* (1998), pp. 63–74
47. H. Hasse, Theorie der höheren differentiale in einem algebraischen funktionenkörper mit vollkommenem konstantenkörper bei beliebiger charakteristik. J. Reine Angew. Math. **175**, 50–54 (1936)
48. W.C. Huffman, V. Pless, *Fundamentals of Error-Correcting Codes* (Cambridge University Press, Cambridge, 2003)
49. J.Y. Hyun, A subgroup of the full poset-isometry group. SIAM J. Discrete Math. **24**(2), 589–599 (2010)
50. J.Y. Hyun, H.K. Kim, The poset structures admitting the extended binary Hamming code to be a perfect code. Discrete Math. **288**(1–3), 37–47 (2004)
51. J.Y. Hyun, H.K. Kim, Maximum distance separable poset codes. Des. Codes Crypt. **48**(3), 247–261 (2008)
52. J.Y. Hyun, Y. Lee, MDS poset-codes satisfying the asymptotic Gilbert-Varshamov bound in Hamming weights. IEEE Trans. Inf. Theory **57**(12), 8021–8026 (2011)
53. J.Y. Hyun, H.K. Kim, J. Rye Park, The weighted poset metrics and directed graph metrics (2017). arXiv e-prints
54. Y. Jang, J. Park, On a MacWilliams type identity and a perfectness for a binary linear $(n, n-1, j)$-poset code. Discrete Math. **265**(1–3), 85–104 (2003)
55. C. Jang, H.K. Kim, D.Y. Oh, Y. Rho, The poset structures admitting the extended binary Golay code to be a perfect code. Discrete Math. **308**(18), 4057–4068 (2008)
56. N. Karmarkar, R.M. Karp, The diferencing method of set partitioning. Technical Report 810, Computer Science Division (EECS), University of California, Berkley (1982)
57. D.S. Kim, Dual MacWilliams pair. IEEE Trans. Inf. Theory **51**(8), 2901–2905 (2005)
58. D.S. Kim, MacWilliams-type identities for fragment and sphere enumerators. Eur. J. Comb. **28**(1), 273–302 (2007)
59. D.S. Kim, S.H. Cho, Weight distribution of the crown-weight space. Eur. J. Comb. **28**(1), 356–370 (2007)
60. D.S. Kim, D.C. Kim, Character sums and MacWilliams identities. Discrete Math. **287**(1–3), 155–160 (2004)
61. H.K. Kim, D.S. Krotov, The poset metrics that allow binary codes of codimension $m$ to be $m$, $(m-1)$, or $(m-2)$-perfect, in *2007 IEEE International Symposium on Information Theory* (2007), pp. 1371–1375
62. H.K. Kim, D.S. Krotov, The poset metrics that allow binary codes of codimension $m$ to be $m$, $(m-1)$, or $(m-2)$-perfect. IEEE Trans. Inf. Theory **54**(11), 5241–5246 (2008)
63. D.S. Kim, J.G. Lee, A MacWilliams-type identity for linear codes on weak order. Discrete Math. **262**(1–3), 181–194 (2003)
64. H.K. Kim, D.Y. Oh, A classification of posets admitting the MacWilliams identity. IEEE Trans. Inf. Theory **51**(4), 1424–1431 (2005)
65. H.K. Kim, D.Y. Oh, On the nonexistence of triple-error-correcting perfect binary linear codes with a crown poset structure. Discrete Math. **297**(1–3), 174–181 (2005)
66. D.J. Kleitman, B.L. Rothschild, Asymptotic enumeration of partial orders on a finite set. Trans. Am. Math. Soc. **205**, 205–220 (1975)
67. C. Lee, Some properties of nonbinary error-correcting codes. IRE Trans. Inf. Theory **4**(2), 77–82 (1958)
68. K. Lee, The automorphism group of a linear space with the Rosenbloom-Tsfasman-Metric. Eur. J. Comb. **24**(6), 607–612 (2003)
69. Y. Lee, Projective systems and perfect codes with a poset metric. Finite Fields Appl. **10**(1), 105–112 (2004)
70. J.G. Lee, Perfect codes on some ordered sets. Bull. Korean Math. Soc. **43**(2), 293–297 (2006)

71. R. Lidl, H. Niederreiter, *Introduction to finite fields and their applications* (Cambridge University Press, Cambridge, 1994)
72. W.C. Lidl, H. Niederreiter, *Finite Fields*, 2nd edn. (Cambridge University Press, Cambridge, 1996)
73. R.A. Machado, M. Firer, MacWilliams' identity for metrics determined by directed graphs, in *2016 IEEE Information Theory Workshop (ITW)* (2016), pp. 96–100
74. R.A. Machado, J.A. Pinheiro, M. Firer, Characterization of metrics induced by hierarchical posets. IEEE Trans. Inf. Theory **63**(6), 3630–3640 (2017)
75. F.J. MacWilliams, Combinatorial properties of elementary Abelian groups. Thesis, Radcliffe College, Cambridge, MA, 1962
76. F.J. MacWilliams, A theorem on the distribution of weights in a systematic code. Bell Syst. Tech. J. **42**(1), 79–94 (1963)
77. F.J. MacWilliams, N.J.A. Sloane, *The Theory of Error-Correcting Codes* (Elsevier, Amsterdam, 1977)
78. F.J. MacWilliams, N.J.A. Sloane, J.M. Goethals, The MacWilliams identities for nonlinear codes. Bell Labs Tech. J. **51**(4), 803–819 (1972)
79. J.L. Massey, Notes on Coding Theory, Class Notes for Course 6.575 (spring) (MIT, Cambridge, 1967)
80. S. Mertens, The easiest hard problem: number partitioning (2003). eprint arXiv:cond-mat/0310317
81. A.O. Moura, M. Firer, Duality for poset codes. IEEE Trans. Inf. Theory **56**(7), 3180–3186 (2010)
82. H. Niederreiter, Point sets and sequences with small discrepancy. Monatshefte für Mathematik **104**(4), 273–337 (1987)
83. H. Niederreiter, A combinatorial problem for vector spaces over finite fields. Discrete Math. **96**(3), 221–228 (1991)
84. H. Niederreiter, Orthogonal arrays and other combinatorial aspects in the theory of uniform point distributions in unit cubes. Discrete Math. **106–107**, 361–367 (1992)
85. D.Y. Oh, Poset metrics admitting association schemes and a new proof of MacWilliams identity. J. Korean Math. Soc. **50**(5), 917–931 (2013)
86. M. Ozen, I. Siap, Linear codes over $\mathbb{F}_q[u]/(u^s)$ with respect to the Rosenbloom-Tsfasman metric. Des. Codes Crypt. **38**(1), 17–29 (2006)
87. L. Panek, M. Firer, H.K. Kim, J.Y. Hyun, Groups of linear isometries on poset structures. Discrete Math. **308**(18), 4116–4123 (2008)
88. L. Panek, M. Firer, M.M. Silva Alves, Symmetry groups of Rosenbloom-Tsfasman spaces. Discrete Math. **309**(4), 763–771 (2009)
89. L. Panek, M. Firer, M.M.S. Alves, Classification of Niederreiter-Rosenbloom-Tsfasman block codes. IEEE Trans. Inf. Theory **56**(10), 5207–5216 (2010)
90. J.A. Pinheiro, M. Firer, Classification of poset-block spaces admitting MacWilliams-type identity. IEEE Trans. Inf. Theory **58**(12), 7246–7252 (2012)
91. J.A. Pinheiro, R.A. Machado, M. Firer, Combinatorial metrics: MacWilliams-type identities, isometries and extension property (2017). CoRR abs/1703.08271
92. A. Poplawski, On matched metric and channel problem. CoRR abs/1606.02763 (2016)
93. J. Quistorff, On Rosenbloom and Tsfasman's generalization of the Hamming space. Discrete Math. **307**(21), 2514–2524 (2007)
94. C.M. Qureshi, Matched metrics to the binary asymmetric channels. CoRR abs/1606.09494 (2016)
95. S. Reiger, Codes for the correction of clustered errors. IRE Trans. Inf. Theory **6**(1), 16–21 (1960)
96. M.Y. Rosenbloom, M.A. Tsfasman, Codes for the $m$-metric. Problems Inf. Transm. **33**, 45–52 (1997)
97. H.G. Schaathun, Duality and support weight distributions. IEEE Trans. Inf. Theory **50**(5), 862–867 (2004)

98. G. Séguin, On metrics matched to the discrete memoryless channel. J. Franklin Inst. **309**(3), 179–189 (1980)
99. E.C. Shannon, A mathematical theory of communication. Bell Syst. Tech. J. **27**, 379–423 (1948)
100. R.C. Singleton, Maximum distance $q$-nary codes. IEEE Trans. Inf. Theory **10**(2), 116–118 (1964)
101. M.M. Skriganov, Coding theory and uniform distributions. St. Petersburg Math. J. **13**, 301–337 (2002)
102. M.V.P. Spreafico, Extensão de isomorfismos de ideais em conjuntos parcialmente ordenados. Ph.D. Thesis, Unicamp, 2016
103. R.P. Stanley, *Enumerative Combinatorics*. Cambridge Studies in Advanced Mathematics, vol. 49 (Cambridge University Press, Cambridge, 1997)
104. A. Tietäväinen, On the nonexistence of perfect codes over finite fields. SIAM J. Appl. Math. **24**(1), 88–96 (1973)
105. W.T. Tutte, Lectures on matroids. J. Res. Nat. Bur. Standards Sect. B **69**(1–47), 468 (1965)
106. J.H. van Lint, On the nonexistence of certain perfect codes, in *Computers in Number Theory* (1969), pp. 227–282
107. A. Vardy, The intractability of computing the minimum distance of a code. IEEE Trans. Inf. Theory **43**(6), 1757–1766 (1997)
108. V.K. Wei, Generalized Hamming weights for linear codes. IEEE Trans. Inf. Theory **37**(5), 1412–1418 (1991)
109. H. Whitney, On the abstract properties of linear dependence. Am. J. Math. **57**(3), 509–533 (1935)
110. R.J. Wilson, An introduction to matroid theory. Am. Math. Mon. **80**(5), 500–525 (1973)
111. J.A. Wood, Lecture notes on the MacWilliams identities and the extension theorem, in *CIMAT International School and Conference on Coding Theory* (2008)
112. V.A. Zinoviev, V.K. Leontiev, The nonexistence of perfect codes over Galois fields. Problems Control Inf. Theory **2**(2), 123–132 (1973)

# Index





M. Firer et al., *Poset Codes: Partial Orders, Metrics and Coding Theory*, SpringerBriefs in Mathematics, https://doi.org/10.1007/978-3-319-93821-9