Coding Theory

F. Oggier

These notes were written for the CIMPA School in Ho Chi Minh City, Vietnam, June 2024.

- Section 1 is meant to provide some context and pointers. It is not meant to be exhaustive, and it is certainly biased by my own mathematical interests.
- Section 2 is based on the draft of a book chapter, for a book edited by Prof. Dinh Hai, as part of a project with the Vietnam Institute of Advanced Study in Mathematics (VIASM) in Hanoi. These notes are made available with his agreement. The book chapter also contains more general bounds, properties with respect to duality, and a section on quasi-cyclic codes not present in these notes.

1 Some Historical Notes and Context

The beginning of coding theory is often attributed to Claude Shannon (see Figure 1), who, in 1948, published "A Mathematical Theory of Communication", in the Bell System Technical Journal. This is not to say that there was no notion of coding before, but rather that his work provided a mathematical foundation on how to model communication. This work is also considered as the beginning of information theory. Important early contributors to coding theory include Marcel Golay (Golay codes), Richard Hamming (Hamming distance and Hamming codes), and Irving S. Reed and Gustave Solomon (Reed-Solomon codes, Reed-Mueller codes), see Figure 2.

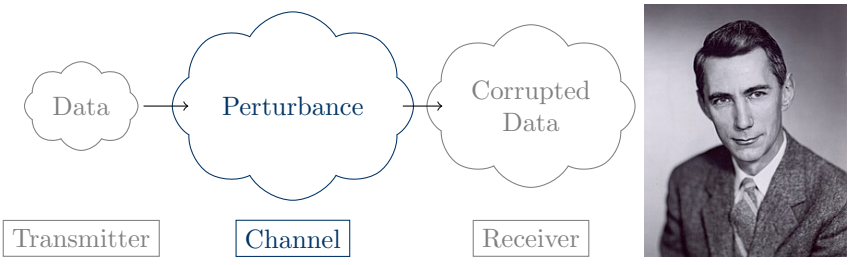


Figure 1: On the right, Claude Shannon. On the left, abstractions of communication channels.

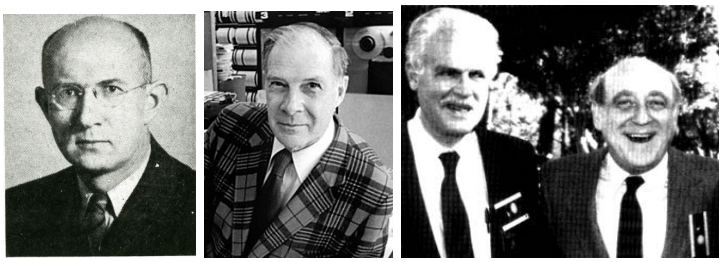


Figure 2: Marcel Golay, Richard Hamming, Reed and Solomon.

It is useful to remember that coding theory is a mathematical theory built on top of models for communication. The general rough framework of a communication system involves a transmitter (or several of them), a noisy communication channel (modelled typically in terms of probability), and a receiver (or several of them). The goal is for the transmitter and the receiver to communicate as reliably and efficiently as possible, despite the noise of the channel. Mathematical coding theory usually considers a single transmitter and a single receiver, and data to be transmitted is modelled as a vector with coefficients in some finite alphabet A, where A is typically a finite field, but it could also be a finite ring:

$$(x_1, x_2, \ldots, x_k) \in A^k \quad \longrightarrow \quad \text{Errors/Erasures} \quad \longrightarrow \quad (y_1, y_2, \ldots, y_k) \in A^k.$$

Both the transmitted and received vectors have coefficients over the same alphabet, but if erasures happen, the resulting vectors may contain erasures, sometimes denoted by $*$. The assumption is that there is no control on the channel, so we need to work at the transmitter and the receiver, by introducing an encoder at the transmitter and a decoder at the receiver. If the information symbols to be transmitted are $x_1, \ldots, x_k$, an encoder will first map them to a longer vector $(c_1, \ldots, c_n) \in A^n$, $n \geq k$, and it is this new vector, often called codeword, that will be transmitted. The question becomes, how to design the map from the information symbols to the codeword (and then how to decode, that is retrieve the information symbols). To help design codes providing reliability, metrics are often used. The most commone ones include the Hamming metric, the rank metric (both will be covered in this course), but other metrics of interest include e.g., the Lee metric.

Since coding theory is built on top of models of communication, these models have evolved over time, together with communication technologies. Therefore, depending on whether we are considering a wired or a wireless transmission technology, we may have different mathematical models, such as:

- The alphabet A is finite, e.g., a finite field or a finite ring, the channel adds errors and/or erasures.
- The alphabet A is $\mathbb{R}$, the channel adds Gaussian noise.

- The alphabet A is $\mathbb{C}$, the channel adds fading and Gaussian noise (this is for wireless communication).

Each of these scenarios may be approached using algebraic methods. For the second case, codes may be built from Euclidean lattices. For the third case, codes may be built using central simple algebras. Classical mathematical coding theory looks at the first case, but even in this case, several flavours exist:

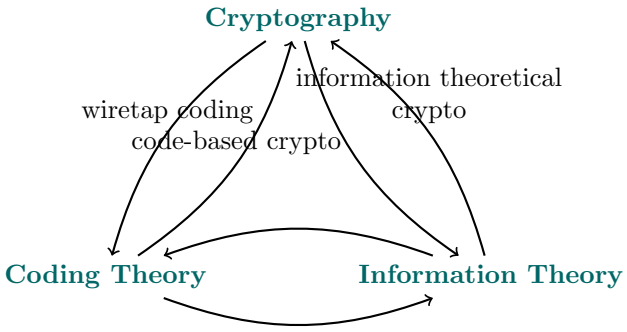
- Linear codes over finite fields (this will be the topic of both courses in the school), where the metric of interest is traditionally the Hamming metric, but we will also discuss the rank metric, which is a metric which got renewed interest recently (and is still a topic of active research).
- Codes with asymptotic length, built using algebraic geometry.
- Codes over rings.
- There is a whole area of coding theory which does not rely on algebraic methods, but rather on probabilistic ones, such as LDPC codes (though some LDPC codes do come with some algebraic structures), or polar codes.

Mathematical coding theory often uses the communication problem as a starting point, but then builds mathematical theories of their own.

The picture of coding theory would be incomplete if I were to skip the following two topics:

- Coding for storage: classical linear codes and their variations have been popular over the last 15 years for distributed storage systems (e.g., cloud storage), and earlier for other storage medium (e.g., CD).
- Quantum coding is an area of coding enjoying a growing interest with the progress in quantum computing.

Finally, coding theory is close to two other areas: information theory and cryptography. The goal of coding theory is reliability, while that of cryptography is security (confidentiality comes to mind, but there are many other forms of security, such as integrity and availability). Information theory typically studies limits of communication, and a variety of communication channels are considered (with or without memory, with or without feedback, for a different number of users), using probabilistic models:



Connections between cryptography and coding/information theory have become more popular over the past years, due to the rise of post-quantum cryptography.

2 On Generalized Rank Weights of Linear Codes

Classical coding theory assumes the alphabet is a finite field $\mathbb{F}_q$, where the index q refers to the size of the field and is a prime power. When $q = p$ is a prime, $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$ corresponds to integers modulo p. Then $\mathbb{F}_q$ for $q = p^r$ is a vector space over $\mathbb{F}_p$ of dimension r, meaning that once fixed an $\mathbb{F}_p$-basis, an element of $\mathbb{F}_q$ may be written as a vector of length r with coefficients in $\mathbb{F}_p$. In the same manner, $\mathbb{F}_{q^m}$ is a vector space over $\mathbb{F}_q$.

Given the extension $\mathbb{F}_{q^m}/\mathbb{F}_q$ of finite fields, for q a prime power and $m \geq 2$, an $[n, k]$ linear code $\mathcal{C}$ in $\mathbb{F}_{q^m}^n$ is a k-dimensional subspace of the vector space $\mathbb{F}_{q^m}^n$ over $\mathbb{F}_{q^m}$. The extension $\mathbb{F}_{q^m}/\mathbb{F}_q$ is Galois, with cyclic Galois group generated by $\sigma : \mathbb{F}_{q^m} \to \mathbb{F}_{q^m}$, $a \mapsto a^q$. Since $\mathcal{C}$ is a subspace, we may fix a basis, and a generator matrix G for $\mathcal{C}$ contains as rows the chosen basis vectors. Alternatively, since $\mathcal{C}$ is a subspace, it is the kernel of some linear transformation (the projection onto $\mathcal{C}$), and fixing again a basis, $\mathcal{C}$ is the kernel of H, called a parity check matrix.

To a linear code is usually attached a distance. The Hamming weight of a nonzero codeword $\mathbf{c} \in \mathcal{C}$ counts the number of nonzero coefficients of $\mathbf{c}$, and the Hamming distance of $\mathcal{C}$ is the minimum Hamming weight across all nonzero codewords of $\mathcal{C}$. For $\mathcal{D}$ a subcode of $\mathcal{C}$, that is a subspace of $\mathcal{C}$, its support $\mathrm{supp}(\mathcal{D})$ is the set of coordinates at which not all codewords of $\mathcal{D}$ are zero. The notion of support allows to generalize the Hamming distance in the following sense [16]. For $1 \leq r \leq k$

$$d_r(\mathcal{C}) = \min\{|\mathrm{supp}(\mathcal{D})|,\ \mathcal{D} \text{ an } [n, r] \text{ subcode of } \mathcal{C}\}$$

is called the rth generalized Hamming weight of $\mathcal{C}$. When $r = 1$, d_1 is the Hamming distance.

The reason to specify $\mathbb{F}_{q^m}/\mathbb{F}_q$ with $m \geq 2$ here is because we will be interested in a different distance than the Hamming distance, namely that of the rank [7, 14]. The rank weight of a nonzero codeword $\mathbf{c} \in \mathcal{C}$ counts the maximal number of linearly independent coefficients of $\mathbf{c}$ over $\mathbb{F}_q$, and the rank distance of $\mathcal{C}$ is the minimum rank weight across all nonzero codewords of $\mathcal{C}$. In a sense, the rank distance is a refinement of the Hamming distance, since the Hamming distance drops when coefficients are zero, but the rank distance further drops when coefficients are multiples, or more generally linear combinations of each other. The topic of this course, namely generalized rank weights, is about developing an analogy of rth generalized Hamming weights in the context of the rank distance instead.

2.1 Definitions and First Example

Let V be a subspace of $\mathbb{F}_{q^m}^n$. Recall the map $\sigma : \mathbb{F}_{q^m} \to \mathbb{F}_{q^m}$, $a \mapsto a^q$. For $\mathbf{v} = (v_1, \ldots, v_n) \in V$, $\sigma(\mathbf{v})$ is understood componentwise, namely $\sigma(\mathbf{v}) = (\sigma(v_1), \ldots, \sigma(v_n))$. Similarly, $\sigma(V)$ means that σ is applied to every vector $\mathbf{v} \in V$. We collect all subspaces V such that $\sigma(V) = V$ in the set $\Gamma = \Gamma(\mathbb{F}_{q^m}^n)$.

Definition 1. [9] Let $\mathcal{C}$ be an $[n, k]$ linear code in $\mathbb{F}_{q^m}^n$. For $1 \leq r \leq k$, the rth generalized rank weight of $\mathcal{C}$ is

$$M_r(\mathcal{C}) = \min_{\substack{V \in \Gamma \\ \dim(\mathcal{C} \cap V) \geq r}} \dim(V).$$

More precisely, the definition given in [9] is that of relative generalized rank weight, namely instead of $\dim(\mathcal{C} \cap V) \geq r$, the constraint $\dim(\mathcal{C}_1 \cap V) - \dim(\mathcal{C}_2 \cap V) \geq r$ is used, where the code C_1 contains the subcode C_2 [9, Definition 5]; the above definition uses $C_1 = \mathcal{C}$ and $C_2 = \mathbf{0}$ and is thus a particular case.

Since we are considering the dimension $\dim(\mathcal{C} \cap V)$ of the intersection $\mathcal{C} \cap V$, the dimension $\dim(V)$ is minimized by considering subspaces that are included in $\mathcal{C}$. We then however need to ensure that such V belong to Γ. Given a subspace V, the smallest subspace that contains V and is also stable by σ is the subspace $V^* = V + \sigma(V) + \ldots + \sigma^{m-1}(V)$, called the Galois closure of V. By definition, $V^* \in \Gamma$. If $V \in \Gamma$, then $\sigma(V) = V$ and $V^* = V$. Conversely, if $V^* = V$, then $V = \sigma(V)$.

This leads to the following equivalent definition (the formal proof is given below):

Definition 2. [13] Let $\mathcal{C}$ be an $[n, k]$ linear code over $\mathbb{F}_{q^m}$. For $1 \leq r \leq k$, the rth generalized rank weight of $\mathcal{C}$ is

$$M_r(\mathcal{C}) = \min_{\substack{\mathcal{D} \text{ a subcode of } \mathcal{C} \\ \dim(\mathcal{D}) = r}} \dim(\mathcal{D}^*).$$

For $r = k$, $M_r(\mathcal{C}) = \dim(\mathcal{C}^*)$. Across these notes, we may use the term "rank weight" for short, instead of "generalized rank weight".

Proposition 1. *[13, Theorem 21] Definitions 1 and 2 are equivalent, namely:*

$$\min_{\substack{V \in \Gamma \\ \dim(\mathcal{C} \cap V) \geq r}} \dim(V) = \min_{\substack{\mathcal{D} \text{ a subcode of } \mathcal{C} \\ \dim(\mathcal{D}) = r}} \dim(\mathcal{D}^*).$$

Proof. Take $V \in \Gamma$ with $\dim(\mathcal{C} \cap V) \geq r$ and minimum dimension. Such a V always exists, since $V = \mathcal{C}^*$ belongs to Γ by definition, and $\mathcal{C} \cap \mathcal{C}^*$ contains at least $\mathcal{C}$ which has dimension $k \geq r$. To show that the left-hand side is greater or equal to the right-hand side, we need to exhibit a subcode $\mathcal{D}$ of $\mathcal{C}$ of dimension r such that $\dim(\mathcal{D}^*) \leq \dim(V)$. That $V \in \Gamma$ implies $V = V^*$. Choose $\mathcal{D} \subseteq \mathcal{C} \cap V$, a subcode of $\mathcal{C}$ of dimension r, which exists since $\mathcal{C} \cap V$ is a subcode of $\mathcal{C}$ of dimension at least r. Since $V = V^*$ and $\mathcal{D} \subseteq V$, we must have $\mathcal{D}^* \subseteq V$. So $\mathcal{D}^*$ is a subspace of V, it has a dimension smaller or equal to that of V.

Conversely, to show that the right-hand side is greater or equal to the left-hand side, we need a subspace $V \in \Gamma$ such that $\dim(\mathcal{C} \cap V) \geq r$ and $\dim(V) \leq \dim(\mathcal{D}^*)$ for all subcodes $\mathcal{D} \subseteq \mathcal{C}$ such that $\dim(\mathcal{D}) = r$. For all such subcodes $\mathcal{D}$, take $V = \mathcal{D}^*$. Indeed, we have $\dim(\mathcal{C} \cap \mathcal{D}^*) \geq r$ since $\mathcal{D} \subseteq \mathcal{D}^*$, and $\mathcal{D}^* \in \Gamma$ by definition (and trivially $\dim(V) \leq \dim(\mathcal{D}^*)$). □

When $r = 1$, a subcode $\mathcal{D}$ of dimension 1 is a subcode $\langle \mathbf{c} \rangle$ generated by a nonzero codeword $\mathbf{c} \in \mathcal{C}$.

Lemma 1. *A code $\mathcal{C}$ has generalized rank weight $M_1(\mathcal{C}) = 1$ if and only if there exists a nonzero codeword $\mathbf{c} \in \mathcal{C}$ with coefficients in $\mathbb{F}_q$.*

Proof. If $\mathbf{c} \in \mathcal{C}$ has coefficients in $\mathbb{F}_q$, then $\mathcal{D} = \langle \mathbf{c} \rangle = \mathcal{D}^*$ and using Definition 2, $M_1(\mathcal{C}) = 1$. Conversely, if $M_1(\mathcal{C}) = 1$, there exists $\mathcal{D}$ a subcode of $\mathcal{C}$ such that $\dim(\mathcal{D}^*) = 1$, that is $\mathcal{D} = \mathcal{D}^* = \langle \mathbf{c} \rangle$ for $c \in \mathcal{C}$, and $\sigma(\langle \mathbf{c} \rangle) = \langle \mathbf{c} \rangle$. Since $\mathcal{C}$ is linear and $\mathbf{c} \neq \mathbf{0}$, we may assume without loss of generality that $\mathbf{c} = (1, c_2, \ldots, c_n)$ (the coordinate at 1 can be placed in any arbitrary position, it will always exist since $\mathbf{c}$ is nonzero). Thus $\sigma(\mathbf{c}) = (1, \sigma(c_2), \ldots, \sigma(c_n)) = a(1, c_2, \ldots, c_n)$ for some $a \in \mathbb{F}_{q^m}$, showing that $a = 1$ and $\mathbf{c} \in \mathbb{F}_q^n$. □

Example 1. Consider the $[6, 3]$ hexacode $\mathcal{C}$ over $\mathbb{F}_4 = \{0, 1, \omega, \omega^2 = \omega + 1\}$, whose generator matrix in systematic form is given by

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

The sum of the three rows of G is the codeword $(1, 1, 1, 1, 1, 1)$, thus by Lemma 1

$$M_1(\mathcal{C}) = \min_{\substack{\mathcal{D} \text{ a subcode of } \mathcal{C} \\ \dim(\mathcal{D})=1}} \dim(\mathcal{D}^*) = 1.$$

To compute $M_2(\mathcal{C})$, we next consider all 2-dimensional subcodes of $\mathcal{C}$. Take again $(1, 1, 1, 1, 1, 1)$ and for example $(1, 0, 0, 1, \omega, \omega)$, the first row of G. These two vectors generate a 2-dimensional subcode

$$\mathcal{D} = \{a(1, 1, 1, 1, 1, 1) + b(1, 0, 0, 1, \omega, \omega),\ a, b \in \mathbb{F}_4\}.$$

Then $\sigma(\mathcal{D}) = \{a(1, 1, 1, 1, 1, 1) + b(1, 0, 0, 1, \omega^2, \omega^2),\ a, b \in \mathbb{F}_4\}$ and $\mathcal{D}^* = \mathcal{D} + \sigma(\mathcal{D})$ has dimension 3, which shows that

$$M_2(\mathcal{C}) = \min_{\substack{\mathcal{D} \text{ a subcode of } \mathcal{C} \\ \dim(\mathcal{D})=2}} \dim(\mathcal{D}^*) \leq 3.$$

To show it is 3, we need to show $M_2(\mathcal{C})$ cannot be 2 (since $\dim(\mathcal{D}) = 2$, $\dim(\mathcal{D}^*) \geq 2$). For it to be 2, given any two codewords $\mathbf{c}_1$, $\mathbf{c}_2$, that are linearly independent, we would need both $\sigma(\mathbf{c}_1)$ and $\sigma(\mathbf{c}_2)$ to be linear combinations of $\mathbf{c}_1, \mathbf{c}_2$, that is in particular, they both need to be codewords. For a generic

codeword, we apply σ on it and check whether it belongs to the kernel of the parity check matrix corresponding to G. For

$$\begin{bmatrix} 1 & \omega & \omega & 1 & 0 & 0 \\ \omega & 1 & \omega & 0 & 1 & 0 \\ \omega & \omega & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ a + b\omega^2 + c\omega^2 \\ a\omega^2 + b + c\omega^2 \\ a\omega^2 + b\omega^2 + c \end{bmatrix} = \mathbf{0}$$

to hold, the third equation implies that $a = b$, after which the second equation implies that $b = c$. This holds only for the subcode $\{a(1, 1, 1, 1, 1, 1),\ a \in \mathbb{F}_4\}$ and the trivial subcode $\mathbf{0}$.

We are left with computing $M_3(\mathcal{C})$ ($k = 3$ means that we look at $\mathcal{C}$ itself). Since $\sigma((1, 0, 0, 1, \omega, \omega)) = (1, 0, 0, 1, \omega^2, \omega^2)$ is not a codeword, we add it to the basis, and check whether $\sigma((0, 1, 0, \omega, 1, \omega)) = (0, 1, 0, \omega^2, 1, \omega^2)$ is generated by

$$\begin{bmatrix} 1 & 0 & 0 & 1 & \omega & \omega \\ 0 & 1 & 0 & \omega & 1 & \omega \\ 0 & 0 & 1 & \omega & \omega & 1 \\ 1 & 0 & 0 & 1 & \omega^2 & \omega^2 \end{bmatrix}.$$

This is not the case, because the 3rd row cannot be used (we want a 0 as 3rd coefficient), the second row is necessarily used as such (we want a 1 as 2nd coefficient), and the same multiple of the first and 4rth rows must be used (we want a 0 as 1st coefficient), which gives

$$(0, 1, 0, \omega, 1, \omega) + a(0, 0, 0, 0, 1, 1),\ a \in \mathbb{F}_4$$

and we cannot generate the desired vector. We thus add it to get

$$\begin{bmatrix} 1 & 0 & 0 & 1 & \omega & \omega \\ 0 & 1 & 0 & \omega & 1 & \omega \\ 0 & 0 & 1 & \omega & \omega & 1 \\ 1 & 0 & 0 & 1 & \omega^2 & \omega^2 \\ 0 & 1 & 0 & \omega^2 & 1 & \omega^2 \end{bmatrix}$$

and summing all the rows generates $\sigma((0, 0, 1, \omega, \omega, 1)) = (0, 0, 1, \omega^2, \omega^2, 1)$. This shows that

$$M_3(\mathcal{C}) = \dim(\mathcal{C}^*) = 5.$$

In summary, the generalized rank weight hierarchy of the $[6, 3]$ hexacode $\mathcal{C}$ is

$M_1(\mathcal{C})$	$M_2(\mathcal{C})$	$M_3(\mathcal{C})$
1	3	5

Exercise 1. Check numerically that for the code $\mathcal{C}$ of Example 1, $M_2(\mathcal{C}) = 3$.

We remark that to compute $\mathcal{C}^*$ given $\mathcal{C}$ and a generator matrix G, it is enough to compute $\sigma(G), \sigma^2(G), \ldots, \sigma^{m-1}(G)$ (σ is applied componentwise). Indeed, for a codeword $c = \sum_{i=1}^{k} c_i \mathbf{g}_i$, $\sigma(c) = \sum_{i=1}^{k} \sigma(c_i)\sigma(\mathbf{g}_i)$ for $\mathbf{g}_i$ the ith row of G. If $\sigma(\mathbf{g}_i)$ belongs to $\mathcal{C}^*$, then so will $\sigma(c_i)\sigma(\mathbf{g}_i)$ by linearity and because σ is an automorphism of $\mathbb{F}_{q^m}$.

The above example illustrates that a direct approach to compute generalized rank weights is not so straightforward, even for a simple small code. Even for the first rank, it requires some amount of computations. Then for higher ranks, an exhaustive search is not quite possible in general.

Open Problem 1. Propose an algorithm to reduce the computational complexity of computing generalized rank weights.

2.2 Basic Properties

We start with proving a few properties that will lead us to the monotonicity of generalized rank weights.

Proposition 2. *A subspace V is in Γ if and only if there exists a basis of V formed by vectors in $\mathbb{F}_q^n$.*

Proof. Let $l \leq n$ be the dimension of V.

Suppose V has a basis $\mathbf{b}_1, \ldots, \mathbf{b}_l \in \mathbb{F}_q^n$. Then every vector $\mathbf{v}$ in V is of the form $\sum a_i \mathbf{b}_i$, $a_i \in \mathbb{F}_{q^m}$. Then $\sigma(\mathbf{v}) = \sum \sigma(a_i)\sigma(\mathbf{b}_i) = \sum \sigma(a_i)\mathbf{b}_i$ which belongs to V because V is a subspace and σ is an automorphism of $\mathbb{F}_{q^m}$. This shows that $\sigma(V) = V$ and thus $V \in \Gamma$.

Conversely, suppose $V \in \Gamma$ and let $\mathbf{b}_1, \ldots, \mathbf{b}_l \in \mathbb{F}_{q^m}^n$ be a basis, which we may assume is in reduced row echelon form, meaning that after stacking the vectors as l rows of a matrix: (1) the leading coefficient of a non-zero row is always strictly to the right of the leading coefficient of the row above it and it is 1, and (2) each column containing a leading 1 has zeroes in all its other entries. Let $\mathbf{b}$ be any of these rows. By (2), $\mathbf{b}$ is zero at every coordinate corresponding to a leading 1 of every other row. Compute $\mathbf{b} - \sigma(\mathbf{b})$, which thus is also zero at every coordinate corresponding to a leading 1 (including at the position of the leading 1 of $\mathbf{b}$, since in this position $1 - \sigma(1) = 0$). Now $\mathbf{b} - \sigma(\mathbf{b})$ belongs to V (since $V \in \Gamma$, $\sigma(V) = V$), thus it is a linear combination of $\mathbf{b}_1, \ldots, \mathbf{b}_l$. But $\mathbf{b} - \sigma(\mathbf{b}) = \sum_{i=1}^{l} a_i \mathbf{b}_i$ is zero at every leading 1, forcing $a_i = 0$ for all i thus $\mathbf{b} - \sigma(\mathbf{b}) = \mathbf{0} \iff \mathbf{b} = \sigma(\mathbf{b})$ and $\mathbf{b} \in \mathbb{F}_q^n$. □

Lemma 2. *Let $\mathcal{C}$ be an $[n, k]$ linear code over $\mathbb{F}_{q^m}$. We have*

$$0 \leq \max_{\substack{V \in \Gamma \\ \dim(V) = i+1}} \dim(\mathcal{C} \cap V) - \max_{\substack{W \in \Gamma \\ \dim(W) = i}} \dim(\mathcal{C} \cap W) \leq 1, \ 0 \leq i \leq n-1.$$

Proof. If $i = 0$, $0 \leq \max_{\substack{V \in \Gamma \\ \dim(V) = 1}} \dim(\mathcal{C} \cap V) \leq 1$. We thus assume $1 \leq i \leq n-1$. Consider a subspace $\tilde{W} \in \Gamma$ of dimension i maximizing $\dim(\mathcal{C} \cap W)$. By Proposition 2, it has a basis of vectors in $\mathbb{F}_q^n$. Create a subspace W' by

adding to this basis a linearly independent vector in $\mathbb{F}_q^n$, W' thus belongs to Γ by Proposition 2, and $\dim(\mathcal{C} \cap W') \geq \dim(\mathcal{C} \cap \tilde{W})$, thus

$$\max_{\substack{V \in \Gamma \\ \dim(V)=i+1}} \dim(\mathcal{C} \cap V) \geq \dim(\mathcal{C} \cap W') \geq \dim(\mathcal{C} \cap \tilde{W}) = \max_{\substack{W \in \Gamma \\ \dim(W)=i}} \dim(\mathcal{C} \cap W).$$

This proves the first inequality.

Then choose a subspace $\tilde{V} \in \Gamma$ of dimension $i+1$ maximizing $\dim(\mathcal{C} \cap V)$. By Proposition 2, it has a basis of vectors in $\mathbb{F}_q^n$. Create a subspace V' by removing from this basis one vector in $\mathbb{F}_q^n$, V' thus belongs to Γ by Proposition 2. If the removed vector was in $\mathcal{C} \cap \tilde{V}$, then $\dim(\mathcal{C} \cap V') = \dim(\mathcal{C} \cap \tilde{V}) - 1$, else $\dim(\mathcal{C} \cap V') = \dim(\mathcal{C} \cap \tilde{V})$. Either way, $\dim(\mathcal{C} \cap V') \geq \dim(\mathcal{C} \cap \tilde{V}) - 1$. Thus

$$\dim(\mathcal{C} \cap \tilde{V}) - 1 = \max_{\substack{V \in \Gamma \\ \dim(V)=i+1}} \dim(\mathcal{C} \cap V) - 1 \leq \dim(\mathcal{C} \cap V') \leq \max_{\substack{W \in \Gamma \\ \dim(W)=i}} \dim(\mathcal{C} \cap W).$$

This proves the second inequality. □

What Lemma 2 tells us is that from i to $i+1$, the quantity $\max_{\substack{W \in \Gamma \\ \dim(W)=i}} \dim(\mathcal{C} \cap W)$ increases by at most 1. This ensures that given $1 \leq r \leq k$, the set $\{i, \max_{\substack{W \in \Gamma \\ \dim(W)=i}} \dim(\mathcal{C} \cap W) = r\}$ is not empty, since if it were for some r, then either all values of $\max_{\substack{W \in \Gamma \\ \dim(W)=i}} \dim(\mathcal{C} \cap W)$ would be strictly less than r, or strictly more, a contradiction since, when $i = 0$, it is 0, and when $i = n$, it is k (take $W = \mathbb{F}_{q^m}^n$).

This also shows that

$$\min\{i, \max_{\substack{W \in \Gamma \\ \dim(W)=i}} \dim(\mathcal{C} \cap W) \geq r\} = \min\{i, \max_{\substack{W \in \Gamma \\ \dim(W)=i}} \dim(\mathcal{C} \cap W) = r\} \quad (1)$$

since the set on the right-hand side is not empty.

Now the minimum i on the left-hand side is obtained by having a subspace $\tilde{W} \in \Gamma$ of dimension i, such that among all $W \in \Gamma$ of dimension i, $\tilde{W}$ has the largest intersection with $\mathcal{C}$, and in particular $\dim(\mathcal{C} \cap \tilde{W}) \geq r$. Thus

$$\min\{i, \max_{\substack{W \in \Gamma \\ \dim(W)=i}} \dim(\mathcal{C} \cap W) \geq r\} = \min\{i, \exists W \in \Gamma, \ \dim(W) = i, \ \dim(\mathcal{C} \cap W) \geq r\}$$

since we are interested in the dimension i of W, which not only gives back Definition 1, namely

$$M_r(\mathcal{C}) = \min_{\substack{W \in \Gamma \\ \dim(\mathcal{C} \cap W) \geq r}} \dim(W),$$

but also, using (1), shows that

$$M_r(\mathcal{C}) = \min_{\substack{W \in \Gamma \\ \dim(\mathcal{C} \cap W) = r}} \dim(W).$$

Theorem 1 (Monotonicity). *[9, Lemma 9] Let $\mathcal{C}$ be an $[n,k]$ linear code over $\mathbb{F}_{q^m}^n$. Then*

$$1 \leq M_1(\mathcal{C}) < M_2(\mathcal{C}) < \cdots < M_k(\mathcal{C}) \leq n.$$

Proof. We just saw above that

$$\begin{aligned} M_r(\mathcal{C}) &= \min\{i, \max_{\substack{V \in \Gamma \\ \dim(V)=i}} \dim(\mathcal{C} \cap V) = r\} \\ M_{r+1}(\mathcal{C}) &= \min\{j, \max_{\substack{V \in \Gamma \\ \dim(V)=j}} \dim(\mathcal{C} \cap V) = r+1\}. \end{aligned}$$

The two sets are disjoint, because given i, if the maximum dimension of the intersection is r, it cannot be $r+1$, thus $M_r(\mathcal{C}) < M_{r+1}(\mathcal{C})$. □

Next we prove the Singleton bound.

Lemma 3. *Let $\mathcal{C}$ be an $[n,k]$ linear code over $\mathbb{F}_{q^m}$. Then for $1 \leq r \leq k$, $M_r(\mathcal{C}) \leq M_k(\mathcal{C}) - k + r$.*

Proof. Set $a = k - r$. Then we need to prove that $M_{k-a}(\mathcal{C}) \leq M_k(\mathcal{C}) - a$ for $0 \leq a \leq k-1$. We proceed by induction on a. The base case for $a = 0$ is $M_k(\mathcal{C}) \leq M_k(\mathcal{C}) - 0$, which trivially holds. Suppose thus that $M_{k-a}(\mathcal{C}) \leq M_k(\mathcal{C}) - a$ for $0 \leq a < k-1$. We want to show that $M_{k-(a+1)}(\mathcal{C}) \leq M_k(\mathcal{C}) - (a+1)$ holds. We know that $M_{k-(a+1)}(\mathcal{C}) < M_{k-a}(\mathcal{C})$ by Theorem 1, which combined with the inductive hypothesis yields $M_{k-(a+1)}(\mathcal{C}) < M_k(\mathcal{C}) - a$, which is equivalent to $M_{k-(a+1)}(\mathcal{C}) \leq M_k(\mathcal{C}) - (a+1)$ as desired. □

Theorem 2 (Singleton bound). *Let $\mathcal{C}$ be an $[n,k]$ linear code over $\mathbb{F}_{q^m}$. Then for $1 \leq r \leq k$, the rth Singleton bound holds:*

$$M_r(\mathcal{C}) \leq n - k + r.$$

Proof. Observe that $M_k(\mathcal{C}) \leq n$, since a subspace of $\mathbb{F}_{q^m}^n$ cannot have dimension more than n. By Lemma 3, we have $M_r(\mathcal{C}) \leq M_k(\mathcal{C}) - k + r \leq n - k + r$ for $1 \leq r \leq k$. □

For $r = k$, we have $M_k(\mathcal{C}) \leq n$ which is trivial.

Definition 3. A linear code $\mathcal{C}$ whose rth generalized rank weight $M_r(\mathcal{C})$ meets the Singleton bound is called r-maximum rank distance, or r-MRD.

MRD codes for the rank metric are somewhat analogous to MDS (maximum distance separable) codes for the Hamming metric.

Example 2. For a $[6,3]$ code, the Singleton bound gives

$$M_1(\mathcal{C}) \leq 4,\ M_2(\mathcal{C}) \leq 5,\ M_3(\mathcal{C}) \leq 6.$$

We compare the weights of the hexacode with the Singleton bound:

	$M_1(\mathcal{C})$	$M_2(\mathcal{C})$	$M_3(\mathcal{C})$
Singleton	4	5	6
hexacode	1	3	5

.

The hexacode has thus weights pretty far from the Singleton bound.

2.3 Connection with the Rank Distance

We connect $M_1(\mathcal{C})$ to the rank distance [7, 14] of $\mathcal{C}$.

Definition 4. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a vector in $\mathbb{F}_{q^m}^n$. Its rank weight is the maximal number of linearly independent coordinates x_i over $\mathbb{F}_q$. For $\mathcal{C}$ an $[n, k]$ linear code over $\mathbb{F}_{q^m}$, its rank distance is the minimum over the rank weight of all its nonzero codewords.

We recall that the rank weight is always smaller or equal to the Hamming weight. Indeed, the Hamming weight counts the number of nonzero coordinates, but among them, some could be multiples or more generally linear combinations of each others (see Subsection 2.4 for a more precise discussion).

Example 3. Consider the vector $(1, 0, 0, 1, w, w) \in \mathbb{F}_4^6$ with $w^2 = w + 1$. Then its Hamming weight is 4, while its rank weight is 2, because $1, w$ are linearly independent over $\mathbb{F}_2$.

Proposition 3. *Let* $\mathbf{v}$ *be a nonzero vector in* $\mathbb{F}_{q^m}^n$. *Then its rank distance* d *is given by*

$$d = \dim\langle \mathbf{v} \rangle^*.$$

Proof. Let $\mathbf{v}$ be a nonzero vector in $\mathbb{F}_{q^m}^n$. Its rank distance d is the maximum number of its coordinates which are linearly independent over $\mathbb{F}_q$. Thus we can write $\mathbf{v} = \mathbf{e}P$ where $\mathbf{e}$ is a vector containing d linearly independent entries of $\mathbf{v}$ in $\mathbb{F}_{q^m}$ followed by $n - d$ zeros, and P is a non-singular matrix with coefficients in $\mathbb{F}_q$. Then

$$\begin{aligned} \dim\langle \mathbf{v} \rangle^* &= \dim\langle \mathbf{v}, \sigma(\mathbf{v}), \ldots, \sigma^{m-1}(\mathbf{v}) \rangle \\ &= \dim\langle \mathbf{e}P, \sigma(\mathbf{e})P, \ldots, \sigma^{m-1}(\mathbf{e})P \rangle \\ &= \dim\langle \mathbf{e}, \sigma(\mathbf{e}), \ldots, \sigma^{m-1}(\mathbf{e}) \rangle. \end{aligned}$$

By stacking these vectors as rows of an $m \times n$ matrix, and recalling the construction of $\mathbf{e}$, we get that $\dim\langle v \rangle^* = d$. □

Example 4. Consider again the vector $\mathbf{v} = (1, 0, 0, 1, w, w) \in \mathbb{F}_4^6$ with $w^2 = w + 1$ and rank weight 2. Since $1, w$ are linearly independent over $\mathbb{F}_2$, we write

$$\mathbf{v} = (1, w, 0, 0, 0, 0) \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ & & \tilde{P} & & & \end{bmatrix}.$$

where $\tilde{P}$ is chosen so the overall matrix P is non-singular, e.g.,

$$\tilde{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} \mathbf{v} \\ \sigma(\mathbf{v}) \end{bmatrix} = \begin{bmatrix} 1 & w & 0 & 0 & 0 & 0 \\ 1 & \sigma(w) & 0 & 0 & 0 & 0 \end{bmatrix} P.$$

Let $\mathcal{C}$ be an $[n,k]$ linear code over $\mathbb{F}_{q^m}$. Let $\mathbf{c}$ be a codeword of $\mathcal{C}$ of rank d. By the above proposition, $\dim\langle\mathbf{c}\rangle^* = d$. Now

$$\begin{aligned} M_1(\mathcal{C}) &= \min\{\dim(V),\ V\in\Gamma,\ \exists\ c\in\mathcal{C}\cap V,\ c\neq\mathbf{0}\} \\ &= \min\{\dim\langle\mathbf{c}\rangle^*,\ \mathbf{c}\in\mathcal{C},\ \mathbf{c}\neq\mathbf{0}\} \end{aligned}$$

which shows that $M_1(\mathcal{C})$ indeed corresponds to the rank distance of $\mathcal{C}$ (see also [9, II.D]).

2.4 Connection with the Hamming Distance

The goal next is to prove that for $\mathcal{C}$ an $[n,k]$ linear code over $\mathbb{F}_{q^m}$, we have

$$M_r(\mathcal{C}) \leq d_r(\mathcal{C}),\ 1\leq r\leq k,$$

namely, the rth generalized rank weight $M_r(\mathcal{C})$ is always a lower bound to the rth generalized Hamming weight $d_r(\mathcal{C})$, which we recall is given by

$$d_r(\mathcal{C}) = \min\{|\mathrm{supp}(\mathcal{D})|,\ \mathcal{D}\text{ an }[n,r]\text{ subcode of }\mathcal{C}\}.$$

In order to compare both weights, we first need to express $d_r(\mathcal{C})$ in terms of $\mathbb{F}_{q^m}$-subspaces. Define $\Lambda = \Lambda(\mathbb{F}_{q^m}^n)$ to be the set of subspaces V of $\mathbb{F}_{q^m}^n$ which are an $\mathbb{F}_{q^m}$-span of i distinct rows of the $n\times n$ identity matrix, for $0\leq i\leq n$.

We then have [9, Remark 6]

$$d_r(\mathcal{C}) = \min_{\substack{V\in\Lambda \\ \dim(\mathcal{C}\cap V)=r}} \dim(V).$$

Indeed, for a given r, the condition $\dim(\mathcal{C}\cap V) = r$ identifies subcodes of $\mathcal{C}$, namely $\mathcal{C}\cap V$, of dimension r, and $V\in\Lambda$ enforces the condition on the code support.

Now since subspaces in Λ have unit vectors as basis vectors, in particular these vectors belong to $\mathbb{F}_q^n$ and thus $\Lambda\subset\Gamma$. Therefore optimizing over Γ gives more subspace candidates than optimizing over Λ, so the minimum over Γ is always smaller or equal to that of over Λ, which shows that

$$M_r(\mathcal{C}) \leq d_r(\mathcal{C}),\ 1\leq r\leq k.$$

2.5 Related Works

This course is dedicated to generalized rank weights of two main families of linear codes, and is mostly based on [9, 4, 6, 13, 5]. There are a number of related topics, as illustrated in Figure 3. A series of works [9, 11, 13, 3, 12] looked at defining generalized rank weights, and studied some of their properties. A natural motivation is the parallelism with generalized Hamming weights, which are known to characterize leakage of information for wiretap codes [16], and a natural application is wiretap network coding [9, 11, 12].

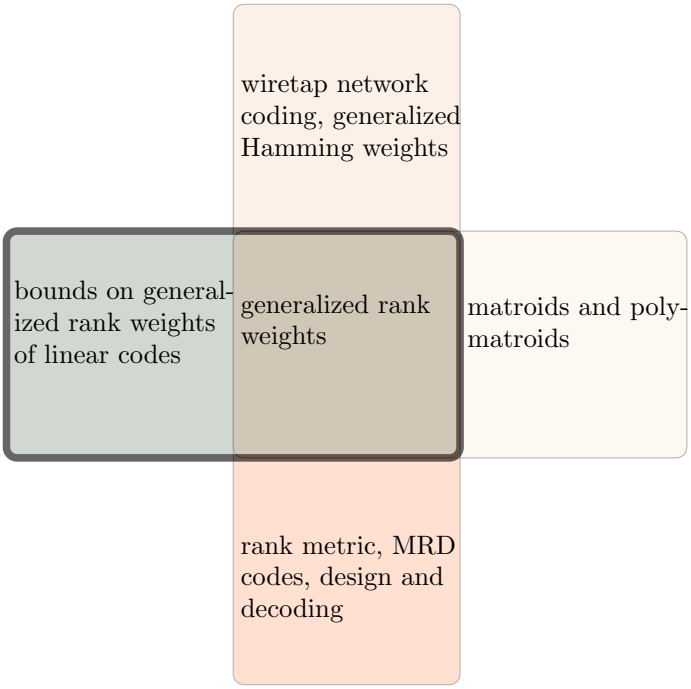


Figure 3: Topics around generalized rank weights: this chapter focuses on the definitions, basic properties and bounds on generalized rank weights of linear codes.

Generalized rank weights extend the notion of rank metric, similarly to how generalized Hamming weights extend the Hamming metric. The rank metric itself has attracted a lot of attention over the years, starting with the works [7, 14]. Topics of interests are the design of codes with respect to the rank metric, and their decoding, in particular, their list decoding (e.g. [17], this is really just one pointer, there are many works done in this direction).

Another direction of study is the connection to matroids and polymatroids (see [15, 8] and references therein).

We also note the generalization to fields of characteristic zero instead of finite fields (see [3] for definitions and [2, 1] for code constructions and decoding).

Finally, the notion of sum-rank metric has been gaining traction recently, as a generalization of both the Hamming and rank metric, see e.g. [10].

2.6 Cyclic Codes

Let $f(x) \in \mathbb{F}_q[x]$ be a polynomial of degree n. The quotient ring $\mathbb{F}_{q^m}[x]/(f)$ is an $\mathbb{F}_{q^m}$-vector space of dimension n. An $[n, k]$ cyclic code is then any ideal

$\mathcal{C}$ of $\mathbb{F}_{q^m}[x]/(f)$, defined by a generator polynomial $g(x) \in \mathbb{F}_{q^m}[x]$ dividing $f(x) = x^n - 1$ of degree $n-k$. To get constacyclic codes, or more precisely λ-cyclic codes, take instead $f(x) = x^n - \lambda,\ \lambda \in \mathbb{F}_q,\ \lambda \neq 0$. We will make the assumption throughout this section that $f(x), g(x)$ have simple roots.

Suppose that $g(x)$ is split in $\mathbb{F}_{q^m}$ (there is a corresponding theory for the case where $g(x)$ is not split in $\mathbb{F}_{q^m}$) with simple roots $\alpha_1, \ldots, \alpha_{n-k}$, and reorder these roots such that roots are grouped together when they share the same minimal polynomial over $\mathbb{F}_q$. Let $\alpha_1, \ldots, \alpha_\nu$ be the roots of $f(x)$ in $\mathbb{F}_{q^m}$. Then

$$\begin{aligned} f(x) &= (x-\alpha_1)\cdots(x-\alpha_\nu)\prod_{i=\nu+1}^{n}(x-\alpha_i) \\ &= (x-\alpha_1)\cdots(x-\alpha_{n-k})(x-\alpha_{n-k+1})\cdots(x-\alpha_\nu)\prod_{i=\nu+1}^{n}(x-\alpha_i) \\ &= (x-\alpha_1)\cdots(x-\alpha_{m_1})\cdot \\ &\quad (x-\alpha_{m_1+1})\cdots(x-\alpha_{m_2})\cdots \\ &\quad (x-\alpha_{m_{s-1}+1})\cdots(x-\alpha_{m_s})\cdot \\ &\quad (x-\alpha_{n-k+1})\cdots(x-\alpha_\nu)\prod_{i=\nu+1}^{n}(x-\alpha_i) \end{aligned}$$

where $m_s = n-k$, and each of the roots $\alpha_1, \alpha_{m_1+1}, \ldots, \alpha_{m_{s-1}+1}$ have minimal polynomial

$$\mu_{\alpha_{m_r+1}}(x) = \prod_{m_r+1 \le t \le m_{r+1}}(x-\alpha_t)\prod_{j\in J_r}(x-\alpha_j), \tag{2}$$

for $0 \le r \le s$ (setting $m_0 = 0$). The first product contains the linear factors that appear in $g(x)$, the second product contains those that appear in $f(x)$ only.

Let G be a generator matrix of $\mathcal{C}$. Then a codeword is given by

$$[c_0, c_1, \ldots, c_{k-1}]G,\ c_0, \ldots, c_{k-1} \in \mathbb{F}_{q^m}.$$

Written in terms of polynomial, we get

$$c(x)g(x),\ c(x) = c_0 + c_1 x + \ldots + c_{k-1}x^{k-1},$$

where $g(x)$ is of degree $n-k$, yielding a polynomial of degree $\le n-1$, whose n coefficients correspond to one codeword.

Since $g(x)$ splits, any codeword can be written as

$$c(x)\prod_{1\le j\le n-k}(x-\alpha_j).$$

Recall from Lemma 1 that a code $\mathcal{C}$ has rank weight 1 if and only if there exists a codeword with coefficients in $\mathbb{F}_q$, which translates here into saying that the corresponding polynomial $c(x)\prod_{1\le j\le n-k}(x-\alpha_j)$ lives in $\mathbb{F}_q[x]$.

Definition 5. Set

$$\eta_q(\mathcal{C}) = \sum_{0 \le r \le s-1} [\mathbb{F}_q(\alpha_{m_r+1}) : \mathbb{F}_q],$$

where $[\mathbb{F}_q(\alpha_{m_r+1}) : \mathbb{F}_q] = \deg(\mu_{m_r+1})$ and μ_{m_r+1} is the minimal polynomial of α_{m_r+1}.

Lemma 4. *We have $\eta_q(\mathcal{C}) \le n$.*

Proof. The minimal polynomial $\mu_{\alpha_{m_r+1}}(x)$ divides $f(x)$ in $\mathbb{F}_q[x]$ for every $0 \le r \le s-1$ and since the polynomials $\mu_{\alpha_{m_r+1}}$ are pairwise coprime, the polynomial

$$\prod_{0 \le r \le s-1} \mu_{\alpha_{m_r+1}}(x)$$

which has degree $\eta_q(\mathcal{C})$ divides $f(x)$ in $\mathbb{F}_q[x]$ whose degree is n. □

Recall that we are interested in $f(x) = x^n - \lambda$, with $\lambda \neq 0$, and $\lambda = 1$ corresponds to cyclic codes.

Proposition 4. *[4, Proposition 2] Let $\mathcal{C}$ be an $[n, k]$ code over $\mathbb{F}_{q^m}$ such that $f(x), g(x)$ have only simple roots and $g(x)$ is split in $\mathbb{F}_{q^m}[x]$. Then $\mathcal{C}$ has rank weight 1 if and only if $\eta_q(\mathcal{C}) \le n-1$.*

Proof. Assume first that $\eta_q(\mathcal{C}) \le n-1$. Construct a codeword $c(x)g(x)$ of $\mathcal{C}$ by "reconstructing" the minimal polynomial μ_{α_j} according to (2):

$$c(x) = \prod_{0 \le r \le s-1} \prod_{j \in J_r} (x - \alpha_j).$$

Then by design, the polynomial

$$\begin{aligned} c(x)g(x) &= \prod_{0 \le r \le s-1} \prod_{j \in J_r} (x - \alpha_j) \prod_{1 \le j \le n-k} (x - \alpha_j) \\ &= \prod_{0 \le r \le s-1} \prod_{j \in J_r} (x - \alpha_j) \prod_{0 \le r \le s-1} \prod_{m_r+1 \le t \le m_{r+1}} (x - \alpha_t) \\ &= \prod_{0 \le r \le s-1} \mu_{\alpha_{m_r+1}}(x) \end{aligned}$$

has coefficients in $\mathbb{F}_q$. Since the degree of this polynomial is exactly $\eta_q(\mathcal{C}) \le n-1$, $c(x)g(x)$ corresponds to a codeword of $\mathcal{C}$ with coefficients in $\mathbb{F}_q$.

Conversely, assume that $\mathcal{C}$ has rank weight 1. Then there exists a polynomial $c(x)$ with degree $\le k-1$ such that $c(x)g(x)$ has coefficients in $\mathbb{F}_q$. Since α_{m_r+1} is a root of $c(x)g(x) \in \mathbb{F}_q[x]$ for $0 \le r \le s-1$, its minimal polynomial $\mu_{\alpha_{m_r+1}}(x)$ divides $c(x)g(x)$ in $\mathbb{F}_q[x]$. This being true for every $0 \le r \le s-1$ and the polynomials $\mu_{\alpha_{m_r+1}}$ being pairwise coprime, the polynomial

$$\prod_{0 \le r \le s-1} \mu_{\alpha_{m_r+1}}(x)$$

divides $c(x)g(x)$ in $\mathbb{F}_q[x]$. Taking the degrees, we get $\eta_q(\mathcal{C}) \leq n-1$, as desired. □

Example 5. Consider $\mathbb{F}_{5^4}$, with primitive element ω satisfying $\omega^4+4\omega^2+4\omega+2=0$. We consider different $f(x)$ $(x^3-1,\ x^3-2,\ x^4-2,\ x^4-4)$ all of which split in $\mathbb{F}_{5^4}$.

$x^3-1=(x+4)(x+\omega^{104})(x+\omega^{520})$
$g(x)=(x+4),\ J_1=1$ $g(x)=(x+\omega^{104}),\ J_2=(x+\omega^{520})$
$x^3-2=(x+2)(x+\omega^{364})(x+\omega^{572})=(x-\alpha_1)(x-\alpha_2)(x-\alpha_3), \nu=3$
$g(x)=(x+2)=(x-\alpha_1),\ J_1=1$ $g(x)=(x+\omega^{364})=(x-\alpha_2),\ J_2=(x+\omega^{572})$ $g(x)=(x+2)(x+\omega^{364}),\ J_1=1,\ J_2=(x+\omega^{572})$
$x^4-2=(x+\omega^{39})(x+\omega^{195})(x+\omega^{351})(x+\omega^{507})=(x-\alpha_1)\cdots(x-\alpha_4), \nu=4$
$g(x)=(x+\omega^{39}),\ J_1=(x+\omega^{195})(x+\omega^{351})(x+\omega^{507})$
$x^4-4=(x+\omega^{78})(x+\omega^{234})(x+\omega^{390})(x+\omega^{546})=(x-\alpha_1)\cdots(x-\alpha_4), \nu=4$
$g(x)=(x+\omega^{78}),\ J_1=(x+\omega^{390})$ $g(x)=(x+\omega^{78})(x+\omega^{234}),\ J_1=(x+\omega^{390}),\ J_2=(x+\omega^{546})$

When $f(x)=x^3-1$, for $g(x)=x+4$, the polynomial lives in $\mathbb{F}_5[x]$ thus it generates a code of first rank weight 1. For $g(x)=x+\omega^{104}$, we use J_2 to construct $c(x)=g(x)(x+\omega^{520})=x^2+x+1$, $c=(1,1,1)$, showing the first rank weight is again 1.

When $f(x)=x^3-2$, $g(x)=x+2$ and $g(x)=(x+\omega^{364})$ again generate codes of first rank weight 1. For $g(x)=(x+2)(x+\omega^{364})$ however, we have $\eta_5(\mathcal{C})=1+[\mathbb{F}_5(-\omega^{364}):\mathbb{F}_5]=3$ and the corresponding first rank weight is at least 2.

When $f(x)=x^4-2$, $g(x)=x+\omega^{39}$, the corresponding first rank weight is at least 2.

When $f(x)=x^4-4$, $g(x)=x+\omega^{78}$ generates a code whose first rank weight is 1, while $g(x)=(x+\omega^{78})(x+\omega^{234})$ yields a code whose first rank weight is at least 2.

Exercise 2. Check numerically the above factorization and rank weights.

The next result applies for cyclic codes (whether $g(x)$ is split or not).

Proposition 5. *[5, Corollary 1] Let $\mathcal{C}$ be an $[n,k]$ cyclic code over $\mathbb{F}_{q^m}$ such that $(x-1)$ does not divide $g(x)$. Then the minimum rank distance $M_1(\mathcal{C})$ is 1.*

Proof. Since $\mathcal{C}$ is cyclic, $f(x)=x^n-1=(x-1)(x^{n-1}+x^{n-2}+\ldots+x+1)$. Since $g(x)|f(x)$ but does not contain $x-1$, then $g(x)|(x^{n-1}+x^{n-2}+\ldots+x+1)$, say $(x^{n-1}+x^{n-2}+\ldots+x+1)=g(x)c(x)$ for some polynomial $c(x)$. This gives the codeword $(1,1,\ldots,1)$ which shows that $M_1(\mathcal{C})=1$. □

Exercise 3. Illustrate the above proposition with a cyclic code of your choice.

For more consequences of the above setting, see [5]. We will focus next on bounds.

We saw earlier that $\mathcal{C}$ has rank weight 1 if and only if $\eta_q(\mathcal{C}) \leq n-1$ when $g(x)$ splits in $\mathbb{F}_{q^m}$ (in Proposition 4) and a similar result holds when $g(x)$ does not split in $\mathbb{F}_{q^m}$. We provide a natural generalization of this result.

Proposition 6. *[5] Let $\mathcal{C}$ be an $[n,k]$ code over $\mathbb{F}_{q^m}$ such that $f(x) = x^n - \lambda$, $\lambda \neq 0$, and $g(x)$ have simple roots. Then for $1 \leq r \leq k$, $M_r(\mathcal{C}) = r$ if and only if $\eta_q(\mathcal{C}) \leq n-r$.*

Proof. Assume first that $\eta_q(\mathcal{C}) \leq n-r$. Then, taking the polynomial $c(x)$ defined in the proof of Proposition 4 (we give the polynomial for the case where $g(x)$ does not split without proof):

$$c(x) = \begin{cases} \prod_{0 \leq r \leq s-1} \prod_{j \in J_r} (x - \alpha_j) & g(x) \text{ splits} \\ \prod_{i \in I} h_i(x) & \text{else,} \end{cases}$$

we set, for every $0 \leq u \leq r-1$, $c_u(x) = x^u c(x)$. Then, for all $0 \leq u \leq r-1$, $c_u(x)g(x)$ is a polynomial lying in $\mathbb{F}_q[x]$ with degree

$$\begin{aligned} \deg(g(x)c_u(x)) &= \deg(g(x)x^u c(x)) \\ &= u + \deg(g(x)c(x)) \\ &= u + \eta_q(\mathcal{C}) \leq u + n - r \leq n-1 \end{aligned}$$

since $u \leq r-1$.

It then corresponds to a codeword c_u with rank weight 1. Moreover, the subspace V of $\mathcal{C}$ generated by the c_u's has dimension exactly r (for all $0 \leq u \leq r$, the polynomial $c_u(x)g(x)$ has degree $n-r+u$, so the family of the codewords c_u is linearly independent) and V belongs to Γ (since the basis vectors c_u lie in $\mathbb{F}_q^n$). Therefore, $M_r(\mathcal{C}) \leq \dim V = r$. Moreover, as a direct consequence of the monotonicity property (Theorem 1), $r \leq M_r(\mathcal{C})$ and we get the desired equality.

Conversely, assume that $M_r(\mathcal{C}) = r$. Then by definition, there exists a subcode $\mathcal{D} \in \mathcal{C}$ such that $\dim(\mathcal{D}) = r$ and $\mathcal{D} = \mathcal{D}^* \in \Gamma$. Moreover, we know from Proposition 2 that $\mathcal{D}$ has a basis of vectors having coefficients in $\mathbb{F}_q$: there exists some polynomials $c_1(x), \ldots, c_r(x) \in \mathbb{F}_{q^m}[x]$ with degree $\leq k-1$ such that $c_i(x)g(x) \in \mathbb{F}_q[x]$ and the family $\{c_i(x)g(x) | 1 \leq i \leq r\}$ is linearly independent over $\mathbb{F}_{q^m}$. Therefore, there exists a non-zero polynomial $c(x) \in \mathbb{F}_{q^m}[x]$ with degree $\leq k-r$ lying in the subspace spanned by the $c_i(x)g(x)$ over $\mathbb{F}_q$. The minimal polynomial of any root α (say in an algebraic closure of $\mathbb{F}_q$) over $\mathbb{F}_q$ divides $c(x)g(x)$, hence

$$\left(\prod \mu_{\alpha_j}(x)\right) | c(x)g(x),$$

and taking degrees,

$$\eta_q(\mathcal{C}) \leq \deg c(x) + \deg g(x) \leq k - r + n - k = n - r,$$

which completes the proof. □

Example 6. Consider $\mathbb{F}_{5^4}$, with primitive element ω satisfying $\omega^4+4\omega^2+4\omega+2=0$. Take the $[4,2]$ constacyclic code generated by $g(x)=(x+\omega^{39})(x+\omega^{195})$ in $\mathbb{F}_{5^4}[x]/(x^4-2)$, with

$$x^4-2=(x+\omega^{39})(x+\omega^{195})(x+\omega^{351})(x+\omega^{507}),$$

so x^4-2 is the minimal polynomial of all its roots, and thus of $-\omega^{39}$. Then $\eta_5(\mathcal{C})=[\mathbb{F}_5(-\omega^{39})]=4$ and the first rank of $\mathcal{C}$ cannot be 1. If we wanted the first rank of $\mathcal{C}$ to be 2, we would need $c(x)\in\mathbb{F}_{5^4}$ such that

$$g(x)c(x)=(x+\omega^{39})(x+\omega^{195})c(x)\in\mathbb{F}_{5^2}[x],\ \deg(c(x))=1.$$

This would mean that $g(x)c(x)$ is fixed by $\tau: a\mapsto a^{5^2}$, in other words τ is permuting the factors of $g(x)c(x)$. But $\tau(\omega^{39})=\omega^{351}$, and $\tau(\omega^{195})=\omega^{507}$, so the first rank must be 3 (since the second rank cannot be more than 4). This code is thus MRD.

Open Problem 2. The rank weight hierarchy of cyclic codes and constant cyclic codes is open in general.

References

[1] Daniel Augot, Alain Couvreur, Julien Lavauzelle, and Alessandro Neri. Rank-metric codes over arbitrary galois extensions and rank analogues of reed-muller codes. *SIAM J. Appl. Algebra Geom.*, 5(2):165–199, 2021.

[2] Daniel Augot, Pierre Loidreau, and Gwezheneg Robert. Generalized gabidulin codes over fields of any characteristic. *Des. Codes Cryptogr.*, 86(8):1807–1848, 2018.

[3] Grégory Berhuy, Jean Fasel, and Odile Garotta. Rank weights for arbitrary finite field extensions. *Advances in Mathematics of Communications*, 0, 2020.

[4] J. Ducoat and F. Oggier. Rank weight hierarchy of some classes of cyclic codes. In *2014 IEEE Information Theory Workshop (ITW 2014)*, pages 142–146, 2014.

[5] J. Ducoat and F. Oggier. Rank weight hierarchy of some classes of polynomial codes. *Designs, Codes and Cryptography*, 5, 2023.

[6] Jérôme Ducoat. Generalized rank weights : a duality statement. *Contemporary Mathematics*, 632:101–109, 2015.

[7] Ernst Gabidulin. Theory of codes with maximum rank distance (translation). *Problems of Information Transmission*, 21:1–12, 01 1985.

[8] Sudhir R Ghorpade and Trygve Johnsen. A polymatroid approach to generalized weights of rank metric codes. *Designs, Codes and Cryptography*, 88(12):2531–2546, 2020.

[9] J. Kurihara, R. Matsumoto, and T. Uyematsu. Relative generalized rank weight of linear codes and its applications to network coding. *IEEE Transactions on Information Theory*, 61(7):3912–3936, 2015.

[10] Umberto Martínez-Peñas, Mohannad Shehadeh, and Frank R. Kschischang. *Codes in the Sum-Rank Metric: Fundamentals and Applications*. 2022.

[11] F. Oggier and A. Sboui. On the existence of generalized rank weights. In *International Symposium on Information Theory and Its Applications (ISITA 2012)*, 2012.

[12] Alberto Ravagnani. Generalized weights: An anticode approach. *Journal of Pure and Applied Algebra*, 220(5):1946–1962, 2016.

[13] Ruud Pellikaan Relinde Jurrius. On defining generalized rank weights. *Advances in Mathematics of Communications*, 11(1):225–235, 2017.

[14] R.M. Roth. Maximum-rank array codes and their application to crisscross error correction. *IEEE Transactions on Information Theory*, 37(2):328–336, 1991.

[15] Keisuke Shiromoto. Codes with the rank metric and matroids. *Designs, Codes and Cryptography*, 87(8):1765–1776, 2019.

[16] V. K. Wei. Generalized hamming weights for linear codes. *IEEE Transactions on Information Theory*, 37(5):1412–1418, 1991.

[17] Chaoping Xing and Chen Yuan. A new class of rank-metric codes and their list decoding beyond the unique decoding radius. *IEEE Trans. Inf. Theory*, 64:3394–3402, 2018.