Codes, arrangements and weight enumerators

Relinde Jurrius Ruud Pellikaan
Eindhoven University of Technology
Dept. of Mathematics and Computer Science, Coding and Crypto Group
P.O.Box 513, 5600 MB Eindhoven, The Netherlands
r.p.m.j.jurrius@tue.nl g.r.pellikaan@tue.nl

1 Introduction

This is part of the course "Applicable algebra and coding theory" in the Summer School on Applied Computational Algebraic Geometric Modelling (ACAGM) 13–24 July 2009 at Soria, Spain.

We treat error-correcting codes and several measures of error probability. In particular the weight enumerator and its generalization and extension will be considered from several points of view:
1) The theory of arrangements of hyperplanes,
2) The theory of posets with the Möbius function and characteristic polynomial,
3) The theory of matroids with the Tutte polynomial.
All these polynomials are intimately connected. The MacWilliams identities are considered that relate the polynomials of an object and its dual.

These notes are based on a book in preparation [31], the Master's thesis [18] and ongoing research [19, 20].

2 Error-correcting codes

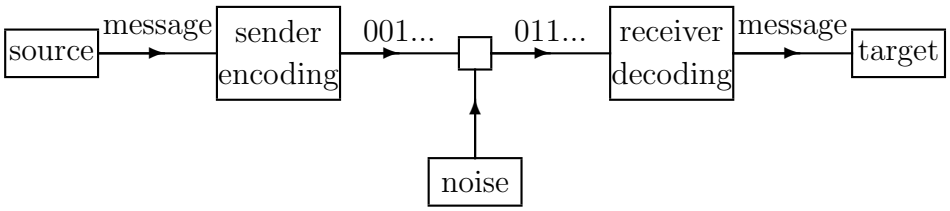
The basics of the theory of error-correcting codes one can find for instance in [26].

2.1 Codes and Hamming distance

The idea of *redundant* information is a well known phenomenon in reading a newspaper. Misspellings go usually unnoticed for a casual reader, while the meaning is still grasped. In Semitic languages such as Hebrew, and even older in the hieroglyphics in the tombs of the pharaohs of Egypt, only the consonants are written while the vowels are left out, so that we do not know for sure how to pronounce these words nowadays. The letter "e" is the most frequent occurring symbol in the English language, and leaving out all these letters would still give in almost all cases an understandable text to the expense of greater attention of the reader. The art and science of deleting redundant information in a clever way such that it can be stored in less memory or space and still can be expanded to the original message, is called *data compression* or *source coding*. It is not the topic of this paper. So we can compress data but an error made in a compressed text would give a different message that is most of the time completely meaningless. The idea in *error-correcting codes* is the converse. One adds redundant information in such a way that it is possible to detect or even correct errors after transmission.

Legend goes that Hamming was so frustrated the computer halted every time it detected an error after he handed in a stack of punch cards, he thought about a way

the computer would be able not only to detect the error but also to correct it automatically. He came with his nowadays famous code named after him. Whereas the theory of Hamming is about the actual construction, the encoding and decoding of codes and uses tools from *combinatorics* and *algebra*, the approach of Shannon lead to *information theory* and his theorems tell us what is and what is not possible in a *probabilistic* sense.



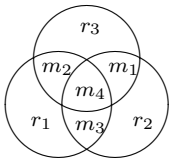
According to Shannon we have a message $\mathbf{m}$ in a certain alphabet and of a certain length, we encode $\mathbf{m}$ to $\mathbf{c}$ by expanding the length of the message and adding redundant information. One can define the *information rate* R that measures the slowing down of the transmission of the data. The encoded message $\mathbf{c}$ is sent over a noisy channel such that the symbols are changed, according to certain probabilities that are characteristic of the channel. The received word $\mathbf{r}$ is decoded to $\mathbf{m}'$. Now given the characteristics of the channel one can define the *capacity* C of the channel and it has the property that for every $R < C$ it is possible to find an encoding and decoding scheme such that the *error probability* that $\mathbf{m}' \neq \mathbf{m}$ is arbitrarily small. For $R > C$ such a scheme is not possible. The capacity is explicitly known as a function of the characteristic probability for quite a number of channels.

The notion of a channel must be taken in a broad sense. Not only the transmission of data via satellite or telephone but also the storage of information on a hard disk of a computer or a compact disc for music and film can be modeled by a channel.

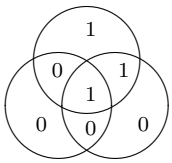
The theorem of Shannon tells us the existence of certain encoding and decoding schemes and one can even say that they exist in abundance and that almost all schemes satisfy the required conditions, but it does not tell us how to construct a specific efficient scheme.

Example 2.1 Replacing every symbol by a threefold repetition gives the possibility of correcting one error in every 3-tuple of symbols in a received word by a majority vote. The price one has to pay is that the transmission is three times slower. We see here the two conflicting demands of error-correction: to correct as many errors as possible and to transmit as fast a possible. Notice furthermore that in case two errors are introduced by transmission the majority decoding rule will introduce an *decoding error*.

Example 2.2 An improvement of the repetition code of rate 1/3 is given by Hamming. Suppose we have a message (m_1, m_2, m_3, m_4) of 4 bits. Put them in the middle of the following *Venn-diagram* of three intersecting circles. Complete the three empty areas of the circles according to the rule that the number of ones in every circle is even. In this way we get 3 redundant bits (r_1, r_2, r_3) that we add to the message and which we transmit over the channel.



In every block of 7 bits the receiver can correct one error. Since the parity in every circle should be even. So if the parity is even we declare the circle correct, if the parity is odd we declare the circle incorrect. The error is in the incorrect circles and in the complement of the correct circles. We see that every pattern of at most one error can be corrected in this way. For instance, if $\mathbf{m} = (1, 1, 0, 1)$ is the message, then $\mathbf{r} = (0, 0, 1)$ is the redundant information added and $\mathbf{c} = (1, 1, 0, 1, 0, 0, 1)$ the codeword sent. If after transmission one symbol is flipped and $\mathbf{y} = (1, 0, 0, 1, 0, 0, 1)$ is the received word.



Then we conclude that the error is in the left and upper circle, but not in the right one. And we conclude that the error is at m_2. But in case of 2 errors and for instance the word $\mathbf{y}' = (1, 0, 0, 1, 1, 0, 1)$ is received, then the receiver would assume that the error occurred in the upper circle and not in the two lower circles, and would therefore conclude that the transmitted codeword was $(1, 0, 0, 1, 1, 0, 0)$. Hence the decoding scheme creates an extra error.

The redundant information $\mathbf{r}$ can be obtained from the message $\mathbf{m}$ by means of three linear equations or parity checks modulo two

$$\begin{cases} r_1 = m_2 + m_3 + m_4 \\ r_2 = m_1 + m_3 + m_4 \\ r_3 = m_1 + m_2 + m_4 \end{cases}$$

Let $\mathbf{c} = (\mathbf{m}, \mathbf{r})$ be the codeword. Then $\mathbf{c}$ is a codeword if and only if $H\mathbf{c}^T = 0$, where

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

The transmission rate is improved from 1/3 for the repetition code to 4/7 for the Hamming code.

In general the alphabets of the message word and the encoded word might be distinct. Furthermore the length of both the message word and the encoded word might vary such as in a *convolutional code*. We restrict ourselves to $[n, k]$ *block codes* that is the message words have a fixed length of k symbols and the encoded words have a fixed length of n symbols both from the same *alphabet* Q. For the purpose of error control, before transmission, we add redundant symbols to the message in a clever way.

Definition 2.3 Let Q be a set of q symbols called the *alphabet*. Let Q^n be the set of all n-tuples $\mathbf{x} = (x_1, \ldots, x_n)$, with entries $x_i \in Q$. A *block code* C of *length* n over Q is a non-empty subset of Q^n. The elements of C are called *codewords*. If C contains M codewords, then M is called the *size* of the code. We call a code with length n and size M an (n, M) code. If $M = q^k$, then C is called an $[n, k]$ code. For an (n, M) code defined over Q, the value $n - \log_q(M)$ is called the *redundancy*. The *information rate* is defined as $R = \log_q(M)/n$.

Example 2.4 The repetition code has length 3 and 2 codewords, so its information rate is $1/3$. The Hamming code has length 7 and 2^4 codewords, therefore its rate is $4/7$.

Example 2.5 Let C be the binary block code of length n consisting of all words with exactly two ones. This is an $(n, n(n-1)/2)$ code. In this example the number of codewords is not a power of the size of the alphabet.

Definition 2.6 Let C be an $[n, k]$ block code over Q. An *encoder* of C is a one-to-one map

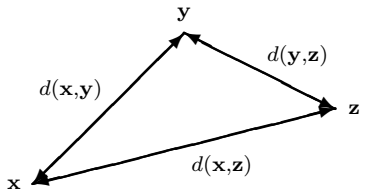
$$\mathcal{E} : Q^k \longrightarrow Q^n$$

such that $C = \mathcal{E}(Q^k)$. Let $\mathbf{c} \in C$ be a codeword. Then there exists a unique $\mathbf{m} \in Q^k$ with $\mathbf{c} = \mathcal{E}(\mathbf{m})$. This $\mathbf{m}$ is called the *message* or *source word* of $\mathbf{c}$.

In order to measure the difference between two distinct words and to evaluate the error-correcting capability of the code, we need to introduce an appropriate metric to Q^n. A natural metric used in Coding Theory is the *Hamming distance*.

Definition 2.7 For $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n) \in Q^n$, the Hamming distance $d(\mathbf{x}, \mathbf{y})$ is defined as the number of places where they differ, that is

$$d(\mathbf{x}, \mathbf{y}) = |\{i \mid x_i \neq y_i\}|.$$



Proposition 2.8 *The Hamming distance is a* metric *on Q^n, that means that it has the following properties:*
1) $d(\mathbf{x}, \mathbf{y}) \geq 0$ *and equality hods if and only if* $\mathbf{x} = \mathbf{y}$,
2) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ *(*symmetry*),*
3) $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ *(*triangle inequality*),*
for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in Q^n$.

Proof. Properties 1) and 2) are trivial from the definition. We leave 3) to the reader as an exercise. ◇

Definition 2.9 The *minimum* (*Hamming*) *distance* of a code C of length n is defined as

$$d = d(C) = \min\{\ d(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in C,\ \mathbf{x} \neq \mathbf{y}\ \}$$

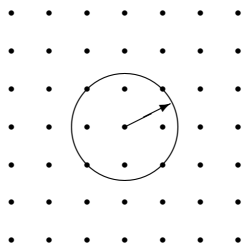
if C consists of more than one element, and is by definition $n + 1$ if C consists of one word. We denote by (n, M, d) a code C with length n, size M and minimum distance d.

The main problem of error-correcting codes from "Hamming's point view" is to construct for a given length and number of codewords a code with the largest possible minimum distance, and to find efficient encoding and decoding algorithms for such a code.

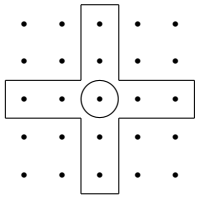
Example 2.10 The triple repetition code consists of two codewords: $(0, 0, 0)$ and $(1, 1, 1)$, so its minimum distance is 3. The Hamming code corrects one error. So the minimum distance is at least 3, by the triangle inequality. The Hamming code has minimum distance 3. Notice that both codes have the property that $\mathbf{x} + \mathbf{y}$ is again a codeword if $\mathbf{x}$ and $\mathbf{y}$ are codewords.

Definition 2.11 Let $\mathbf{x} \in Q^n$. The *ball of radius r around* $\mathbf{x}$, denoted by $B_r(\mathbf{x})$, is defined by $B_r(\mathbf{x}) = \{\ \mathbf{y} \in Q^n \mid d(\mathbf{x}, \mathbf{y}) \leq r\ \}$. The *sphere of radius r around* $\mathbf{x}$ is denoted by $S_r(\mathbf{x})$ and defined by $S_r(\mathbf{x}) = \{\ \mathbf{y} \in Q^n \mid d(\mathbf{x}, \mathbf{y}) = r\ \}$.

The following picture shows the ball in the Euclidean plane. This is misleading in some respects, but gives an indication what we should have in mind.



The following picture shows Q^2, where the alphabet Q consists of 5 elements. The ball $B_0(\mathbf{x})$ consists of the points in the circle, $B_1(\mathbf{x})$ is depicted by the points inside the cross, and $B_2(\mathbf{x})$ consists of all 25 dots.



Proposition 2.12 *Let Q be an alphabet of q elements and $\mathbf{x} \in Q^n$. Then*

$$|S_i(\mathbf{x})| = \binom{n}{i}(q-1)^i \quad \textit{and} \quad |B_r(\mathbf{x})| = \sum_{i=0}^{r} \binom{n}{i}(q-1)^i.$$

Proof. Let $\mathbf{y} \in S_i(\mathbf{x})$. Let I be the subset of $\{1, \ldots, n\}$ consisting of all positions j such that $y_j \neq x_j$. Then the number of elements of I is equal to i. And $(q-1)^i$ is the number of words $\mathbf{y} \in S_i(\mathbf{x})$ that have the same fixed I. The number of possibilities to choose the subset I with a fixed number of elements i is equal to $\binom{n}{i}$. This shows the formula for the number of elements of $S_i(\mathbf{x})$.
Furthermore $B_r(\mathbf{x})$ is the disjoint union of the subsets $S_i(\mathbf{x})$ for $i = 0, \ldots, r$. This proves the statement about the number of elements of $B_r(\mathbf{x})$. ◇

2.2 Linear codes

If the alphabet Q is a finite field, which is the case for instance when $Q = \{0, 1\} = \mathbb{F}_2$, then Q^n is a vector space. Therefore it is natural to look at codes in Q^n that have more structure, in particular that are linear subspaces.

Definition 2.13 A *linear code* C is a linear subspace of $\mathbb{F}_q^n$, where $\mathbb{F}_q$ stands for the finite field with q elements. The *dimension* of a linear code is its dimension as a linear space over $\mathbb{F}_q$. We denote a linear code C over $\mathbb{F}_q$ of length n and dimension k by $[n, k]_q$, or simply by $[n, k]$. If furthermore the minimum distance of the code is d, then we call by $[n, k, d]_q$ or $[n, k, d]$ the *parameters* of the code.

It is clear that for a linear $[n, k]$ code over $\mathbb{F}_q$, its size $M = q^k$. The information rate is $R = k/n$ and the redundancy is $n - k$.

Definition 2.14 For a word $\mathbf{x} \in \mathbb{F}_q^n$, its *support*, $\mathrm{supp}(\mathbf{x})$, is defined as the set of nonzero coordinate positions, so $\mathrm{supp}(\mathbf{x}) = \{i \mid x_i \neq 0\}$. The *weight* of $\mathbf{x}$ is defined as the number of elements of its support, which is denoted by $\mathrm{wt}(\mathbf{x})$. The *minimum weight* of a code C, denoted by $\mathrm{wt}(C)$, is defined as the minimal value of the weights of the nonzero codewords, that is

$$\mathrm{wt}(C) = \min\{\ \mathrm{wt}(\mathbf{c}) \mid \mathbf{c} \in C,\ \mathbf{c} \neq 0\ \},$$

in case there is a $\mathbf{c} \in C$ not equal to 0, and $n + 1$ otherwise.

Proposition 2.15 *The minimum distance of a linear code C is equal to its minimum weight.*

Proof. Since C is a linear code, we have that $0 \in C$ and for any $\mathbf{c}_1, \mathbf{c}_2 \in C$, $\mathbf{c}_1 - \mathbf{c}_2 \in C$. Then the conclusion follows from the fact that $\mathrm{wt}(\mathbf{c}) = d(0, \mathbf{c})$ and $d(\mathbf{c}_1, \mathbf{c}_2) = \mathrm{wt}(\mathbf{c}_1 - \mathbf{c}_2)$. ◇

Now let us see some examples of linear codes.

Example 2.16 The repetition code over $\mathbb{F}_q$ of length n consists of all words $\mathbf{c} = (c, c, \ldots, c)$ with $c \in \mathbb{F}_q$. This is a linear code of dimension 1 and minimum distance n.

Example 2.17 Let n be an integer with $n \geq 2$. The *even weight* code C of length n over $\mathbb{F}_q$ consists of all words in $\mathbb{F}_q^n$ of even weight. The minimum weight of C is by definition 2, the minimum distance of C is 2 if $q = 2$ and 1 otherwise. The code C linear if and only if $q = 2$.

Example 2.18 The *Hamming code* C of Example 2.2 consists of all the words $\mathbf{c} \in \mathbb{F}_2^7$ satisfying $H\mathbf{c}^T = \mathbf{0}$, where

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

This code is linear of dimension 4, since it is given by the solutions of three independent homogeneous linear equations. The minimum weight is 3 as shown in Example 2.10. So it is a $[7, 4, 3]$ code.

2.3 Generator matrix and systematic encoding

Let C be an $[n, k]$ linear code over $\mathbb{F}_q$. Since C is a k-dimensional linear subspace of $\mathbb{F}_q^n$, there exists a *basis* that consists of k linearly independent codewords, say $\mathbf{g}_1, \ldots, \mathbf{g}_k$. Suppose $\mathbf{g}_i = (g_{i1}, \ldots, g_{in})$ for $i = 1, \ldots, k$. Denote

$$G = \begin{pmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_k \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ g_{k1} & g_{k2} & \cdots & g_{kn} \end{pmatrix}.$$

Every codeword $\mathbf{c}$ can be written uniquely as a linear combination of the basis elements, so $\mathbf{c} = m_1\mathbf{g}_1 + \cdots + m_k\mathbf{g}_k$ where $m_1, \ldots, m_k \in \mathbb{F}_q$. Let $\mathbf{m} = (m_1, \ldots, m_k) \in \mathbb{F}_q^k$. Then $\mathbf{c} = \mathbf{m}G$. The *encoding*

$$\mathcal{E} : \mathbb{F}_q^k \longrightarrow \mathbb{F}_q^n,$$

from the message word $\mathbf{m} \in \mathbb{F}_q^k$ to the codeword $\mathbf{c} \in \mathbb{F}_q^n$ can be done efficiently by a matrix multiplication.

$$\mathbf{c} = \mathcal{E}(\mathbf{m}) := \mathbf{m}G.$$

Definition 2.19 A $k \times n$ matrix G with entries in $\mathbb{F}_q$ is called a *generator matrix* of an $\mathbb{F}_q$-linear code C if the rows of G are a basis of C.

A given $[n, k]$ code C can have more than one generator matrix, however every generator matrix of C is a $k \times n$ matrix of rank k. Conversely every $k \times n$ matrix of rank k is the generator matrix of an $\mathbb{F}_q$-linear $[n, k]$ code.

Example 2.20 The linear codes with parameters $[n, 0, n+1]$ and $[n, n, 1]$ are the *trivial codes* $\{0\}$ and $\mathbb{F}_q^n$, and they have the empty matrix and the $n \times n$ indentity matrix I_n as generator matrix, respectively.

Example 2.21 The repetition code of length n has generator matrix

$$G = (\ 1\ 1\ \cdots\ 1\ ).$$

Example 2.22 The binary even-weight code of length n has generator matrices

$$\begin{pmatrix} 1 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & 1 \\ 0 & 1 & \ldots & 0 & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 1 \\ 0 & 0 & \ldots & 0 & 1 & 1 \end{pmatrix}.$$

Example 2.23 The Hamming code C of Example 2.2 is a [7, 4] code. The message symbols m_i for $i = 1, \ldots, 4$ are free to choose. If we take $m_i = 1$ and the remaining $m_j = 0$ for $j \neq i$ we get the codeword $\mathbf{g}_i$. In this way we get the basis $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3, \mathbf{g}_4$. Therefore, C has the following generator matrix

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

2.4 Parity check matrix

There are two standard ways to describe a subspace, *explicitly* by giving a basis, or *implicitly* by the solution space of a set of homogeneous linear equations. Therefore there are two ways to describe a linear code. That is explicitly as we have seen by a generator matrix, or implicitly by a set of homogeneous linear equations that is by the null space of a matrix.

Let C be an $\mathbb{F}_q$-linear $[n, k]$ code. Suppose that H is an $m \times n$ matrix with entries in $\mathbb{F}_q$. Let C be the null space of H. So C is the set of all $\mathbf{c} \in \mathbb{F}_q^n$ such that $H\mathbf{c}^T = 0$. These m homogeneous linear equations are called *parity check equations*, or simply *parity checks*. The dimension k of C is at least $n - m$. If there are dependent rows in the matrix H, that is if $k > n - m$, then we can delete a few rows until we obtain an $(n - k) \times n$ matrix H' with independent rows and with the same null space as H. So H' has rank $n - k$.

Definition 2.24 An $(n - k) \times n$ matrix of rank $n - k$ is called a *parity check matrix* of an $[n, k]$ code C if C is the null space of this matrix.

Remark 2.25 The parity check matrix of a code can be used for *error detection*. This is useful in a communication channel where one asks for *retransmission* in case more than a certain number of errors occurred. Suppose that C is a linear code of minimum distance d and H is a parity check matrix of C. Suppose that the codeword $\mathbf{c}$ is transmitted and $\mathbf{r} = \mathbf{c} + \mathbf{e}$ is received. Then $\mathbf{e}$ is called the *error vector* and $\mathrm{wt}(\mathbf{e})$ the *number of errors*. Now $H\mathbf{r}^T = 0$ if there is no error and $H\mathbf{r}^T \neq 0$ for all $\mathbf{e}$ such that $0 < \mathrm{wt}(\mathbf{e}) < d$. Therefore we can detect any pattern of t errors with $t < d$. But not more, since if the error vector is equal to a nonzero codeword of minimal weight d, then the receiver would assume that no errors have been made. The vector $H\mathbf{r}^T$ is called the *syndrome* of the received word.

We show that every linear code has a parity check matrix and we give a method to obtain such a matrix in case we have a generator matrix G of the code.

Proposition 2.26 *Suppose C is an $[n, k]$ code. Let I_k be the $k \times k$ identity matrix. Let P be a $k \times (n - k)$ matrix. Then, $(I_k|P)$ is a generator matrix of C if and only if $(-P^T|I_{n-k})$ is a parity check matrix of C.*

Proof. Every codeword $\mathbf{c}$ is of the form $\mathbf{m}G$ with $\mathbf{m} \in \mathbb{F}_q^k$. Suppose that the generator matrix G is systematic at the first k positions. So $\mathbf{c} = (\mathbf{m}, \mathbf{r})$ with $\mathbf{r} \in \mathbb{F}_q^{n-k}$ and $\mathbf{r} = \mathbf{m}P$. Hence for a word of the form $\mathbf{c} = (\mathbf{m}, \mathbf{r})$ with $\mathbf{m} \in \mathbb{F}_q^k$ and $\mathbf{r} \in \mathbb{F}_q^{n-k}$ the following statements are equivalent:

$$\mathbf{c} \text{ is a codeword },$$

$$-\mathbf{m}P + \mathbf{r} = 0,$$
$$-P^T\mathbf{m}^T + \mathbf{r}^T = 0,$$
$$\left(-P^T|I_{n-k}\right)(\mathbf{m},\mathbf{r})^T = 0,$$
$$\left(-P^T|I_{n-k}\right)\mathbf{c}^T = 0.$$

Hence $\left(-P^T|I_{n-k}\right)$ is a parity check matrix of C. The converse is proved similarly. $\diamond$

Example 2.27 The trivial codes $\{0\}$ and $\mathbb{F}_q^n$ have I_n and the empty matrix as parity check matrix, respectively.

Example 2.28 As a consequence of Proposition 2.26 we see that a parity check matrix of the binary even weight code is equal to the generator matrix $(\ 1\ 1\ \cdots\ 1\ )$ of the repetition code, and the generator matrix G_2 of the binary even weight code of Example 2.22 is a parity check matrix of the repetition code.

Example 2.29 The generator matrix G of the Hamming code C in Example 2.23 is of the form $(I_4|P)$ and in Example 2.18 we see that the parity check matrix is equal to $(P^T|I_3)$.

2.5 Inner product and dual codes

Definition 2.30 The *inner product* on $\mathbb{F}_q^n$ is defined by

$$\mathbf{x}\cdot\mathbf{y} = x_1y_1 + \cdots + x_ny_n$$

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

This inner product is *bilinear*, *symmetric* and *nondegenerate*, but the notion of "positive definite" makes no sense over a finite field as it does over the real numbers. For instance for a binary word $\mathbf{x} \in \mathbb{F}_2^n$ we have that $\mathbf{x}\cdot\mathbf{x} = 0$ if and only if the weight of $\mathbf{x}$ is even.

Definition 2.31 For an $[n,k]$ code C we define the *dual* or *orthogonal code* $C^\perp$ as

$$C^\perp = \{\mathbf{x} \in \mathbb{F}_q^n \mid \mathbf{c}\cdot\mathbf{x} = 0 \ \text{ for all } \ \mathbf{c} \in C\}.$$

Proposition 2.32 *Let C be an $[n,k]$ code with generator matrix G. Then $C^\perp$ is an $[n,n-k]$ code with parity check matrix G.*

Proof. From the definition of dual codes, the following statements are equivalent:

$$\mathbf{x} \in C^\perp,$$
$$\mathbf{c}\cdot\mathbf{x} = 0 \text{ for all } \mathbf{c} \in C,$$
$$\mathbf{m}G\mathbf{x}^T = 0 \text{ for all } \mathbf{m} \in \mathbb{F}_q^k,$$
$$G\mathbf{x}^T = 0.$$

This means that $C^\perp$ is the null space of G. Because G is a $k \times n$ matrix of rank k, the linear space $C^\perp$ has dimension $n-k$ and G is a parity check matrix of $C^\perp$. $\diamond$

Example 2.33 The trivial codes $\{0\}$ and $\mathbb{F}_q^n$ are dual codes.

Example 2.34 The binary even weight code and the repetition code of the same length are dual codes.

A subspace C of a real vector space $\mathbb{R}^n$ has the property that $C \cap C^\perp = \{0\}$, since the standard inner product is positive definite. Over finite fields this is not always the case.

Definition 2.35 Two codes C_1 and C_2 in $\mathbb{F}_q^n$ are called *orthogonal* if $\mathbf{x} \cdot \mathbf{y} = 0$ for all $\mathbf{x} \in C_1$ and $\mathbf{y} \in C_2$, and they are called *dual* if $C_2 = C_1^\perp$.
If $C \subseteq C^\perp$, we call C *weakly self-dual* or *self-orthogonal*. If $C = C^\perp$, we call C *self-dual*.

Example 2.36 The binary repetition code of length n is self-orthogonal if and only if n is even. This code is self-dual if and only if $n = 2$.

Proposition 2.37 *Let C be an $[n, k]$ code. Then*
1) $(C^\perp)^\perp = C$.
2) C is self-dual if and only C is self-orthogonal and $n = 2k$.

Proof.
1) Let $\mathbf{c} \in C$. Then $\mathbf{c} \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in C^\perp$. So $C \subseteq (C^\perp)^\perp$. Moreover, applying Proposition 2.32 twice, we see that C and $(C^\perp)^\perp$ have the same finite dimension. Therefore equality holds.
2) Suppose C is self-orthogonal, then $C \subseteq C^\perp$. Now $C = C^\perp$ if and only if $k = n - k$, by Proposition 2.32. So C is self-dual if and only if $n = 2k$. ◇

Example 2.38 Consider

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

Let G be the generator matrix of the binary [8,4] code C. Notice that $GG^T = 0$. So $\mathbf{x} \cdot \mathbf{y} = 0$ for all $\mathbf{x}, \mathbf{y} \in C$. Hence C is self-orthogonal. Furthermore $n = 2k$. Therefore C is self-dual. Notice that all rows of G have even weight 4, therefore all codewords have even weight. Every nonzero codeword has weight at least 4, since it is even and it is at least 3 by looking at the first 7 columns that generate the $[7, 4, 3]$ Hamming code. Hence C has parameters [8,4,4].

Remark 2.39 Notice that $\mathbf{x} \cdot \mathbf{x} \equiv \mathrm{wt}(\mathbf{x}) \bmod 2$ if $\mathbf{x} \in \mathbb{F}_2^n$ and $\mathbf{x} \cdot \mathbf{x} \equiv \mathrm{wt}(\mathbf{x}) \bmod 3$ if $\mathbf{x} \in \mathbb{F}_3^n$. Therefore all weights are even for a binary self-orthogonal code and all weights are divisible by 3 for a ternary self-orthogonal code.

Corollary 2.40 *Let C be a linear code. Then,*
1) G is generator matrix of C if and only if G is a parity check matrix of $C^\perp$,
2) H is parity check matrix of C if and only if H is a generator matrix of $C^\perp$.

Proof. The first statement is Proposition 2.32 and the second statement is a consequence of the first applied to the code $C^\perp$ using Proposition 2.37(1). ◇

Proposition 2.41 *Let C be an $[n,k]$ code. Let G be a generator matrix of C and let H be an $(n-k)\times n$ matrix of rank $n-k$. Then H is a parity check matrix of C if and only if $GH^T = O$, where O is the $k\times(n-k)$ zero matrix.*

Proof. Suppose H is a parity check matrix. For any $\mathbf{m}\in\mathbb{F}_q^k$, $\mathbf{m}G$ is a codeword of C. So, $HG^T\mathbf{m}^T = H(\mathbf{m}G)^T = \mathbf{0}$. This implies that $\mathbf{m}GH^T = \mathbf{0}$. Since $\mathbf{m}$ can be any vector in $\mathbb{F}_q^k$. We have $GH^T = \mathbf{0}$.
Conversely, suppose $GH^T = \mathbf{0}$. We assumed that G is a $k\times n$ matrix of rank k and H is an $(n-k)\times n$ matrix of rank $n-k$. So H is the parity check matrix of an $[n,k]$ code C'. For any $\mathbf{c}\in C$, we have $\mathbf{c} = \mathbf{m}G$ for some $\mathbf{m}\in\mathbb{F}_q^k$. Now

$$H\mathbf{c}^T = (\mathbf{m}GH^T)^T = \mathbf{0}.$$

So $\mathbf{c}\in C'$. This implies that $C\subseteq C'$. Hence $C' = C$, since both C and C' have dimension k. Therefore H is a parity check matrix of C. ◇

Remark 2.42 A consequence of Proposition 2.41 is another proof of Proposition 2.26 Because, let $G = (I_k|P)$ be a generator matrix of C. Let $H = (-P^T|I_{n-k})$. Then G has rank k and H has rank $n-k$ and $GH^T = 0$. Therefore H is a parity check matrix of C.

2.6 The Hamming and simplex codes

The following proposition gives a method to determine the minimum distance of a code in terms of the number of dependent columns of the parity check matrix.

Proposition 2.43 *Let H be a parity check matrix of a code C. Then the minimum distance d of C is the smallest integer d such that d columns of H are linearly dependent.*

Proof. Let $\mathbf{h}_1,\ldots,\mathbf{h}_n$ be the columns of H. Let $\mathbf{c}$ be a nonzero codeword of weight w. Let $\mathrm{supp}(\mathbf{c}) = \{j_1,\ldots,j_w\}$ with $1\le j_1<\cdots<j_w\le n$. Then $H\mathbf{c}^T = 0$, so $c_{j_1}\mathbf{h}_{j_1}+\cdots+c_{j_w}\mathbf{h}_{j_w} = 0$ with $c_{j_i}\neq 0$ for all $i = 1,\ldots,w$. Therefore the columns $\mathbf{h}_{j_1},\ldots,\mathbf{h}_{j_w}$ are dependent. Conversely if $\mathbf{h}_{j_1},\ldots,\mathbf{h}_{j_w}$ are dependent, then there exist constants $a_1,\ldots,a_w$, not all zero, such that $a_1\mathbf{h}_{j_1}+\cdots+a_w\mathbf{h}_{j_w} = 0$. Let $\mathbf{c}$ be the word defined by $c_j = 0$ if $j\neq j_i$ for all i, and $c_j = a_i$ if $j = j_i$ for some i. Then $H\mathbf{c}^T = 0$. Hence $\mathbf{c}$ is a nonzero codeword of weight at most w. ◇

Remark 2.44 Let H be a parity check matrix of a code C. As a consequence of Proposition 2.43 we have the following special cases. The minimum distance of code is 1 if and only if H has a zero column. Now suppose that H has no zero column, then the minimum distance of C is at least 2. The minimum distance is equal to 2 if and only if H has two columns say $\mathbf{h}_{j_1},\mathbf{h}_{j_2}$ that are dependent. In the binary case that means $\mathbf{h}_{j_1} = \mathbf{h}_{j_2}$. In other words the minimum distance of a binary code is at least 3 if and only if H has no zero columns and all columns are mutually distinct. This is the case for the Hamming code of Example 2.18. For a given redundancy r the length of a binary linear code C of minimum distance 3 is at most 2^r-1, the number of all nonzero binary columns of length r. For arbitrary $\mathbb{F}_q$, the number of nonzero columns with entries in $\mathbb{F}_q$ is q^r-1. Two such columns are dependent if and only if one is a nonzero multiple of the other. Hence the length of an $\mathbb{F}_q$-linear code code C with $d(C)\ge 3$ and redundancy r is at most $(q^r-1)/(q-1)$.

Definition 2.45 Let $n = (q^r - 1)/(q - 1)$. Let $H_r(q)$ be a $r \times n$ matrix over $\mathbb{F}_q$ with nonzero columns, such that no two columns are dependent. The code $\mathcal{H}_r(q)$ with $H_r(q)$ as parity check matrix is called a q-ary *Hamming code*. The code with $H_r(q)$ as generator matrix is called a q-ary *simplex code* and is denoted by $\mathcal{S}_r(q)$.

Remark 2.46 The simplex code $\mathcal{S}_r(q)$ and the Hamming code $\mathcal{H}_r(q)$ are dual codes, and $H_r(q)$ is a parity check matrix of $\mathcal{H}_r(q)$ and a generator matrix of $\mathcal{S}_r(q)$

Proposition 2.47 *Let $r \geq 2$. Then the q-ary Hamming code $\mathcal{H}_r(q)$ has parameters $[(q^r - 1)/(q - 1), (q^r - 1)/(q - 1) - r, 3]$.*

Proof. The rank of the matrix $H_r(q)$ is r, since the r standard basis vectors of weight 1 are among the columns of the matrix. So indeed $H_r(q)$ is a parity check matrix of a code with redundancy r. Any 2 columns are independent by construction. And a column of weight 2 is a linear combination of two columns of weight 1, and such a triple of columns exists, since $r \geq 2$. Hence the minimum distance is 3 by Proposition 2.43. ◇

Example 2.48 Consider the following ternary Hamming $\mathcal{H}_3(3)$ code of redundancy 3 of length 13 with parity check matrix

$$H_3(3) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 2 & 2 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 0 & 1 \end{pmatrix}.$$

By Proposition 2.47 the code $\mathcal{H}_3(3)$ has parameters $[13, 10, 3]$. Notice that all rows of $H_3(3)$ have weight 9. In fact every linear combination $\mathbf{x}H_3(3)$ with $\mathbf{x} \in \mathbb{F}_3^3$ and $\mathbf{x} \neq 0$ has weight 9. So all nonzero codewords of the ternary simplex code of dimension 3 have weight 9. Hence $\mathcal{S}_3(3)$ is a *constant weight* code. This is a general fact of simplex codes as is stated in the following proposition.

Proposition 2.49 *The q-ary simplex code $\mathcal{S}_r(q)$ is a constant weight code with parameters $[(q^r - 1)/(q - 1), r, q^{r-1}]$.*

Proof. We have seen already in Proposition 2.47 that $H_r(q)$ has rank r, so it is indeed a generator matrix of a code of dimension r. Let $\mathbf{c}$ be a nonzero codeword of the simplex code. Then $\mathbf{c} = \mathbf{m}H_r(q)$ for some nonzero $\mathbf{m} \in \mathbb{F}_q^r$. Let $\mathbf{h}_j^T$ be the j-th column of $H_r(q)$. Then $c_j = 0$ if and only if $\mathbf{m} \cdot \mathbf{h}_j = 0$. Now $\mathbf{m} \cdot \mathbf{x} = 0$ is a nontrivial homogeneous linear equation. This equation has q^{r-1} solutions $\mathbf{x} \in \mathbb{F}_q^r$, it has $q^{r-1} - 1$ nonzero solutions. It has $(q^{r-1} - 1)/(q - 1)$ solutions $\mathbf{x}$ such that $\mathbf{x}^T$ is a column of $H_r(q)$, since for every nonzero $\mathbf{x} \in \mathbb{F}_q^r$ there is exactly one column in $H_r(q)$ that is a nonzero multiple of $\mathbf{x}^T$. So the number of zeros of $\mathbf{c}$ is $(q^{r-1} - 1)/(q - 1)$. Hence the weight of $\mathbf{c}$ is the number of nonzeros which is q^{r-1}. ◇

2.7 Singleton bound and MDS codes

The following bound gives us the maximal minimum distance of a code with a given length and dimension. This bound is called the *Singleton bound*.

Theorem 2.50 (The Singleton Bound) *If C is an $[n, k, d]$ code, then*

$$d \leq n - k + 1.$$

Proof. Let H be a parity check matrix of C. This is an $(n-k) \times n$ matrix of row rank $n-k$. The minimum distance of C is the smallest integer d such that H has d linearly dependent columns, by Proposition 2.43. This means that every $d-1$ columns of H are linearly independent. Hence, the column rank of H is at least $d-1$. By the fact that the column rank of a matrix is equal to the row rank, we have $n-k \geq d-1$. This implies the Singleton bound. $\diamond$

Definition 2.51 Let C be an $[n,k,d]$ code. If $d = n-k+1$, then C is called a *maximum distance separable code* or an MDS code, for short.

Remark 2.52 From the Singleton bound, a maximum distance separable code achieves the maximum possible value for the minimum distance given the code length and dimension.

Example 2.53 The minimum distance of the the zero code of length n is $n+1$, by definition. Hence the zero code has parameters $[n,0,n+1]$ and is MDS. Its dual is the whole space $\mathbb{F}_q^n$ with parameters $[n,n,1]$ and is also MDS. The n-fold repetition code has parameters $[n,1,n]$ and its dual is an $[n,n-1,2]$ code and both are MDS.

Proposition 2.54 *For an $[n,k,d]$ code over $\mathbb{F}_q$, the following statements are equivalent:*
(1) C is an MDS code,
(2) every $n-k$ columns of a parity check matrix H are linearly independent,
(3) every k columns of a generator matrix G are linearly independent.

Proof. Let H be a parity check matrix of an $[n,k,d]$ code C. As the minimum distance of C is d any $d-1$ columns of H are linearly independent, by Proposition 2.43. Now $d \leq n-k+1$ by the Singleton bound. So $d = n-k+1$ if and only if every $n-k$ columns of H are independent. Hence (1) and (2) are equivalent.
Now let us assume (3). Let $\mathbf{c}$ be an element of C which is zero at k given coördinates. Let $\mathbf{c} = \mathbf{x}G$ for some $\mathbf{x} \in \mathbb{F}_q^k$. Let G' be the square matrix consisting of the k columns of G corresponding to the k given zero coördinates of $\mathbf{c}$. Then $\mathbf{x}G' = 0$. Hence $\mathbf{x} = 0$, since the k columns of G' are independent by assumption. So $\mathbf{c} = 0$. This implies that the minimum distance of C is at least $n-(k-1) = n-k+1$. Therefore C is an $[n,k,n-k+1]$ MDS code, by the Singleton bound.
Assume that C is MDS. Let G be a generator matrix of C. Let G' be the square matrix consisting of k chosen columns of G. Let $\mathbf{x} \in \mathbb{F}_q^k$ such that $\mathbf{x}G' = 0$. Then $\mathbf{c} = \mathbf{x}G$ is codeword and its weight is at most $n-k$. So $\mathbf{c} = 0$, since the minimum distance is $n-k+1$. Hence $\mathbf{x} = 0$, since the rank of G is k. Therefore the k columns are independent. $\diamond$

Proposition 2.55 *Let $n \leq q$. Let $\mathbf{a} = (a_1, \ldots, a_n)$ be an n-tuple of mutually distinct elements of $\mathbb{F}_q$. Let k be an integer such that $0 \leq k \leq n$. Define the matrices $G(\mathbf{a})$ and $G'(\mathbf{a})$ by*

$$G(\mathbf{a}) = \begin{pmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_n \\ \vdots & \ddots & \vdots \\ a_1^{k-1} & \cdots & a_n^{k-1} \end{pmatrix} \text{ and } G'(\mathbf{a}) = \begin{pmatrix} 1 & \cdots & 1 & 0 \\ a_1 & \cdots & a_n & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_1^{k-1} & \cdots & a_n^{k-1} & 1 \end{pmatrix}.$$

The codes with generator matrix $G(\mathbf{a})$ and $G'(\mathbf{a})$ are MDS.

Proof. This is left as an exercise. $\diamond$

2.8 Exercises

2.1 Let $\mathbf{x}$ and $\mathbf{y}$ be binary words of the same length. Show that

$$\mathrm{wt}(\mathbf{x}+\mathbf{y}) = \mathrm{wt}(\mathbf{x}) + \mathrm{wt}(\mathbf{y}) - 2|\mathrm{supp}(\mathbf{x}) \cap \mathrm{supp}(\mathbf{y})|.$$

2.2 Let C be an $\mathbb{F}_q$-linear code with generator matrix G. Let $q = 2$. Show that every codeword of C has even weight if and only if every row of a G has even weight. Show by means of a counter example that the above statement is not true if $q \neq 2$.

2.3 Consider the following matrix with entries in $\mathbb{F}_5$

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

Show that G is a generator matrix of a code with parameters $[5, 3, 3]$.

2.4 Give a parity check matrix of the C of Exercise 2.3. Show that C is self-dual.

2.5 Show that there exists a $[2k, k]$ self-dual code over $\mathbb{F}_q$ if and only if there is a $k \times k$ matrix P with entries in $\mathbb{F}_q$ such that $PP^T = -I_k$.

2.6 Show that a binary code is self-orthogonal if the weights of all codewords are divisible by 4. Hint: use Exercise 2.1.

2.7 Give a proof of Proposition 2.55.

3 Weight enumerators and error probability

3.1 Weight spectrum

The weight spectrum of a code is an important invariant, which provides useful information for both the code structure and practical applications of the code.

Definition 3.1 Let C be a code of length n. The *weight spectrum*, also called the *weight distribution* is the following set

$$\{(w, A_w) \mid w = 0, 1, \ldots, n\}$$

where A_w denotes the number of codewords in C of weight w.

The so-called weight enumerator is a convenient representation of the weight spectrum.

Definition 3.2 The *weight enumerator* of C is defined as the following polynomial

$$W_C(Z) = \sum_{w=0}^{n} A_w Z^w.$$

The *homogeneous weight enumerator* of C is defined as

$$W_C(X, Y) = \sum_{w=0}^{n} A_w X^{n-w} Y^w.$$

Remark 3.3 Note that $W_C(Z)$ and $W_C(X,Y)$ are equivalent in representing the weight spectrum. They determine each other uniquely by the following equations

$$W_C(Z) = W_C(1,Z)$$

and

$$W_C(X,Y) = X^n W_C(X^{-1}Y).$$

Given the weight enumerator or the homogeneous weight enumerator, the weight spectrum is determined completely by the coefficients.

Clearly, the weight enumerator and homogeneous weight enumerator can be written in another form, that is

$$W_C(Z) = \sum_{\mathbf{c}\in C} Z^{\mathrm{wt}(\mathbf{c})} \tag{1}$$

and

$$W_C(X,Y) = \sum_{\mathbf{c}\in C} X^{n-\mathrm{wt}(\mathbf{c})} Y^{\mathrm{wt}(\mathbf{c})}. \tag{2}$$

Example 3.4 The zero code has one codeword, and its weight is zero. Hence the homogeneous weight enumerator of this code is $W_{\{0\}}(X,Y) = X^n$. The number of words of weight w in the trivial code $\mathbb{F}_q^n$ is $A_w = \binom{n}{w}(q-1)^w$. So

$$W_{\mathbb{F}_q^n}(X,Y) = \sum_{w=0}^{n} \binom{n}{w}(q-1)^w X^{n-w}Y^w = (X+(q-1)Y)^n.$$

Example 3.5 The n-fold repetition code C has homogeneous weight enumerator

$$W_C(X,Y) = X^n + (q-1)Y^n.$$

In the binary case its dual is the even weight code. Hence it has homogeneous weight enumerator

$$W_{C^\perp}(X,Y) = \sum_{t=0}^{\lfloor n/2 \rfloor} \binom{n}{2t} X^{n-2t}Y^{2t} = \frac{1}{2}\left((X+Y)^n + (X-Y)^n\right).$$

Example 3.6 The nonzero entries of the weight distribution of the [7,4,3] binary Hamming code are given by $A_0 = 1$, $A_3 = 7$, $A_4 = 7$, $A_7 = 1$, as is seen by inspecting the weights of all 16 codewords. Hence its homogeneous weight enumerator is

$$X^7 + 7X^4Y^3 + 7X^3Y^4 + Y^7.$$

Example 3.7 The simplex code $\mathcal{S}_r(q)$ is a constant weight code by Proposition 2.49 with parameters $[(q^r-1)/(q-1), r, q^{r-1}]$. Hence its homogeneous weight enumerator is

$$W_{\mathcal{S}_r(q)}(X,Y) = X^n + (q^r-1)X^{n-q^{r-1}}Y^{q^{r-1}}.$$

Remark 3.8 Let C be a linear code. Then $A_0 = 1$ and the minimum distance $d(C)$ which is equal to the minimum weight, is determined by the weight enumerator as follows:

$$d(C) = \min\{\ i \mid A_i \neq 0, i > 0\ \}.$$

It also determines the dimension $k(C)$, since

$$W_C(1,1) = \sum_{w=0}^{n} A_w = q^{k(C)}.$$

Although there is no apparent relation between the minimum distances of a code and its dual, the weight enumerators satisfy the *MacWilliams identity*.

Theorem 3.9 (MacWilliams) *Let C be an $[n,k]$ code over $\mathbb{F}_q$. Then*

$$W_{C^\perp}(X,Y) = q^{-k} W_C(X+(q-1)Y, X-Y).$$

Proof. See [26, Ch.5. §2. Theorem 1] for a proof of binary codes. A general proof will be given via matroids in Theorem 9.2. ◇

The computation of the minimum distance and the weight enumerator of a code is NP-hard [3, 4, 43].

Example 3.10 The zero code C has homogeneous weight enumerator X^n and its dual $\mathbb{F}_q^n$ has homogeneous weight enumerator $(X+(q-1)Y)^n$, by Example 3.4, which is indeed equal to $q^0 W_C(X+(q-1)Y, X-Y)$ and confirms MacWilliams identity.

Example 3.11 The n-fold repetition code C has homogeneous weight enumerator $X^n+(q-1)Y^n$ and the homogeneous weight enumerator of its dual code in the binary case is $\frac{1}{2}\left((X+Y)^n+(X-Y)^n\right)$, by Example 3.5, which is equal to $2^{-1}W_C(X+Y, X-Y)$, confirming the MacWilliams identity for $q=2$. For arbitrary q we have

$$W_{C^\perp}(X,Y) = q^{-1} W_C(X+(q-1)Y, X-Y) =$$

$$q^{-1}((X+(q-1)Y)^n + (q-1)(X-Y)^n) =$$

$$\sum_{w=0}^{n} \binom{n}{w} \frac{(q-1)^w + (q-1)(-1)^w}{q} X^{n-w}Y^w.$$

3.2 The decoding problem

Definition 3.12 Let C be a linear code in $\mathbb{F}_q^n$ of minimum distance d. If $\mathbf{c}$ is a transmitted codeword and $\mathbf{r}$ is the received word, then $\{i | r_i \neq c_i\}$ is the set of *error positions* and the number of error positions is called the *number of errors* of the received word. Let $\mathbf{e} = \mathbf{r} - \mathbf{c}$. Then $\mathbf{e}$ is called the *error vector* and $\mathbf{r} = \mathbf{c} + \mathbf{e}$. Hence $\mathrm{supp}(\mathbf{e})$ is the set of error positions and $wt(\mathbf{e})$ the number of errors. The e_i's are called the *error values*.
If $t' = d(C, \mathbf{r})$ is the distance of $\mathbf{r}$ to the code C, then there exists a *nearest codeword* $\mathbf{c}'$ such that $t' = d(\mathbf{c}', \mathbf{r})$. So there exists an error vector $\mathbf{e}'$ such that $\mathbf{r} = \mathbf{c}' + \mathbf{e}'$ and $\mathrm{wt}(\mathbf{e}') = t'$. If the number of errors t is at most $(d-1)/2$, then we are sure that $\mathbf{c} = \mathbf{c}'$ and $\mathbf{e} = \mathbf{e}'$. In other words, the nearest codeword to $\mathbf{r}$ is unique when $\mathbf{r}$ has distance at most $(d-1)/2$ to C.

Definition 3.13 $e(C) = \lfloor (d(C)-1)/2 \rfloor$ is called the *error-correcting capacity* of the code C.

Definition 3.14 A *decoder* $\mathcal{D}$ for the code C is a map

$$\mathcal{D} : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^n \cup \{*\}$$

such that $\mathcal{D}(\mathbf{c}) = \mathbf{c}$ for all $\mathbf{c} \in C$.
If $\mathcal{E} : \mathbb{F}_q^k \to \mathbb{F}_q^n$ is an encoder of C and $\mathcal{D} : \mathbb{F}_q^n \to \mathbb{F}_q^k \cup \{*\}$ is a map such that $\mathcal{D}(\mathcal{E}(\mathbf{m})) = \mathbf{m}$ for all $\mathbf{m} \in \mathbb{F}_q^k$, then $\mathcal{D}$ is called a *decoder with respect to the encoder* $\mathcal{E}$. Then $\mathcal{E} \circ \mathcal{D}$ is a decoder of C.

Remark 3.15 It is allowed that the decoder gives as outcome the symbol $*$ in case it fails to find a codeword. This is called a *decoding failure*. If $\mathbf{c}$ is the codeword sent and $\mathbf{r}$ is the received word and $\mathcal{D}(\mathbf{r}) = \mathbf{c}' \neq \mathbf{c}$, then this is called a *decoding error*. If $\mathcal{D}(\mathbf{r}) = \mathbf{c}$, then $\mathbf{r}$ is *decoded correctly*. Notice that a decoding failure is noted on the receiving end, whereas there is no way that the decoder can detect a decoding error.

Definition 3.16 A *complete decoder* is a decoder that always gives a codeword in C as outcome. A *nearest neighbor decoder*, also called a *minimum distance decoder*, is a complete decoder with the property that $\mathcal{D}(\mathbf{r})$ is a nearest codeword. A decoder $\mathcal{D}$ for a code C is called a *t-bounded distance* decoder or a decoder that *corrects t errors* if $\mathcal{D}(\mathbf{r})$ is a nearest codeword for all received words $\mathbf{r}$ with $d(C, \mathbf{r}) \leq t$ errors. A decoder for a code C with error-correcting capacity $e(C)$ *decodes up to half the minimum distance* if it is an $e(C)$-bounded distance decoder, where $e(C) = \lfloor (d(C) - 1)/2 \rfloor$ is the error-correcting capacity of C.

Remark 3.17 If $\mathcal{D}$ is a t-bounded distance decoder, then it is not required that $\mathcal{D}$ gives a decoding failure as outcome for a received word $\mathbf{r}$ if the distance of $\mathbf{r}$ to the code is strictly larger than t. In other words: $\mathcal{D}$ is also a t'-bounded distance decoder for all $t' \leq t$.
The *covering radius* $\rho(C)$ is of a code C is the smallest ρ such that $d(C, \mathbf{y}) \leq \rho$ for all $\mathbf{y}$. A nearest neighbor decoder is a t-bounded distance decoder for all $t \leq \rho(C)$. A $\rho(C)$-bounded distance decoder is a nearest neighbor decoder, since $d(C, \mathbf{r}) \leq \rho(C)$ for all received words $\mathbf{r}$.

Definition 3.18 Let $\mathbf{r}$ be a received word with respect to a code C. A *coset leader* of $\mathbf{r} + C$ is a choice of an element of minimal weight in the coset $\mathbf{r} + C$. Let α_i be the number of cosets of C that are of weight i. Then $\alpha_C(X, Y)$, the *coset leader weight enumerator* of C is the polynomial defined by

$$\alpha_C(X, Y) = \sum_{i=0}^{n} \alpha_i X^{n-i} Y^i.$$

Remark 3.19 The choice of a coset leader of the coset $\mathbf{r} + C$ is unique if $d(C, \mathbf{r}) \leq (d-1)/2$, and $\alpha_i = \binom{n}{i}(q-1)^i$ for all $i \leq (d-1)/2$, where d is the minimum distance of C. Let $\rho(C)$ be the covering radius of the code, then there is at least one codeword $\mathbf{c}$ such that $d(\mathbf{c}, \mathbf{r}) \leq \rho(C)$. Hence the weight of a coset leader is at most $\rho(C)$ and $\alpha_i = 0$ for $i > \rho(C)$.

Definition 3.20 Let $\mathbf{r}$ be a received word. Let $\mathbf{e}$ be the chosen coset leader of the coset $\mathbf{r} + C$. The *coset leader decoder* gives $\mathbf{r} - \mathbf{e}$ as output.

Remark 3.21 The coset leader decoder is a nearest neighbor decoder.

Definition 3.22 A *list decoder* gives as output the collection of all nearest codewords.

Knowing the existence of a decoder is nice to know from a theoretical point of view, in practice the problem is to find an *efficient algorithm* that computes the outcome of the decoder. Whereas finding the closest vector of a given vector to a linear subspace in Euclidean n-space can be computed efficiently by an orthogonal projection to the subspace, the corresponding problem for linear codes is in general not such an easy task. In fact it is an NP-hard problem [3].

3.3 The q-ary symmetric channel

Definition 3.23 The *q-ary symmetric channel* (qSC) is a channel where q-ary words are sent with independent errors with the same *cross-over probability* p at each coordinate, with $0 \leq p \leq \frac{1}{2}$, such that all the $q-1$ wrong symbols occur with the same probability $p/(q-1)$. So a symbol is transmitted correctly with probability $1-p$. The special case $q = 2$ is called the *binary symmetric channel* (BSC).

Remark 3.24 Let $P(\mathbf{x})$ be the probability that the codeword $\mathbf{x}$ is sent. Then this probability is assumed to be the same for all codewords. Hence $P(\mathbf{x}) = \frac{1}{|C|}$ for all $\mathbf{x} \in C$. Let $P(\mathbf{y}|\mathbf{x})$ be the probability that $\mathbf{y}$ is received given that $\mathbf{x}$ is sent. Then

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

for a q-ary symmetric channel.

Definition 3.25 Let C be a code of minimum distance d. Consider the decoder that corrects up to t errors with $2t+1 \leq d$. Let $\mathbf{c}$ be the codeword that is sent. Let $\mathbf{r}$ be the received word. In case the distance of $\mathbf{r}$ to the code is at most t, then the decoder will produce a unique closest codeword $\mathbf{c}'$. If $\mathbf{c} = \mathbf{c}'$, then this is called *correct decoding* which is the case if $d(\mathbf{r}, \mathbf{c}) \leq t$. If $\mathbf{c} \neq \mathbf{c}'$ then it is called a *decoding error*. If $d(\mathbf{r}, C) > t$ the decoding algorithm fails to produce a codeword and such an instance is called a *decoding failure*.

Definition 3.26 For every decoding scheme and channel one defines three probabilities $P_{cd}(p)$, $P_{de}(p)$ and $P_{df}(p)$, that is the probability of correct decoding, decoding error and decoding failure, respectively. Then

$$P_{cd}(p) + P_{de}(p) + P_{df}(p) = 1 \quad \text{for all} \quad p.$$

So it suffices to find formulas for two of these three probabilities. The *error probability* is defined by $P_{err}(p) = 1 - P_{cd}(p)$. Hence $P_{err}(p) = P_{de}(p) + P_{df}(p)$.

Proposition 3.27 *The probability of correct decoding of a decoder that corrects up to t errors with $2t+1 \leq d$ of a code C of minimum distance d on a q-ary symmetric channel with cross-over probability p is given by*

$$P_{cd}(p) = \sum_{w=0}^{t} \binom{n}{w} p^w (1-p)^{n-w}.$$

Proof. Every codeword has the same probability of transmission. So

$$P_{cd}(p) = \sum_{\mathbf{x} \in C} P(\mathbf{x}) \sum_{d(\mathbf{x},\mathbf{y}) \leq t} P(\mathbf{y}|\mathbf{x}) = \frac{1}{|C|} \sum_{\mathbf{x} \in C} \sum_{d(\mathbf{x},\mathbf{y}) \leq t} P(\mathbf{y}|\mathbf{x}),$$

Hence

$$P_{cd}(p) = \sum_{w=0}^{t} \binom{n}{w} (q-1)^w \left(\frac{p}{q-1}\right)^w (1-p)^{n-w}$$

by Proposition 2.12 and Remark 3.24. Clearing the factor $(q-1)^w$ in the numerator and the denominator gives the desired result. ◇

In Proposition 3.36 a formula will be derived for the probability of decoding error for a decoding algorithm that corrects errors up to half the minimum distance.

Example 3.28 Consider the binary triple repetition code. Assume that $(0,0,0)$ is transmitted. In case the received word has weight 0 or 1, then it is correctly decoded to $(0,0,0)$. If the received word has weight 2 or 3, then it is decoded to $(1,1,1)$ which is a decoding error. Hence there are no decoding failures and

$$P_{cd}(p) = (1-p)^3 + 3p(1-p)^2 = 1 - 3p^2 + 2p^3 \text{ and } P_{err}(p) = P_{de}(p) = 3p^2 - 2p^3.$$

If the Hamming code is used, then there are no decoding failures and

$$P_{cd}(p) = (1-p)^7 + 7p(1-p)^6 \quad \text{and}$$

$$P_{err}(p) = P_{de}(p) = 21p^2 - 70p^3 + 105p^4 - 84p^5 + 35p^6 - 6p^7.$$

This shows that the error probabilities of the repetition code is smaller than the one for the Hamming code. This comparison is not fair, since only one bit of information is transmitted with the repetition code and 4 bits with the Hamming code. One could transmit 4 bits of information by using the repetition code four times. This would give the error probability

$$1 - (1 - 3p^2 + 2p^3)^4 = 12p^2 - 8p^3 - 54p^4 + 72p^5 + 84p^6 - 216p^7 + \cdots$$

Suppose that four bits of information are transmitted uncoded, by the Hamming code and the triple repetition code, respectively. Then the error probabilities are 0.04, 0.002 and 0.001, respectively if the cross-over probability is 0.01. The error probability for the repetition code is in fact smaller than that of the Hamming code for all $p \leq \frac{1}{2}$, but the transmission by the Hamming code is almost twice as fast as the repetition code.

Example 3.29 Consider the binary n-fold repetition code. Let $t = (n-1)/2$. Use the decoding algorithm correcting all patterns of t errors. Then

$$P_{err}(p) = \sum_{i=t+1}^{n} \binom{n}{i} p^i (1-p)^{n-i}.$$

Hence the error probability becomes arbitrarily small for increasing n. The price one has to pay is that the information rate $R = 1/n$ tends to 0. The remarkable result of Shannon states that for a fixed rate $R < C(p)$, where

$$C(p) = 1 + p\log_2(p) + (1-p)\log_2(1-p)$$

is the *capacity* of the binary symmetric channel, one can devise encoding and decoding schemes such that $P_{err}(p)$ becomes arbitrarily small.

The main problem of error-correcting codes from "Shannon's point view" is to construct efficient encoding and decoding algorithms of codes with the smallest error probability for a given information rate and cross-over probability.

3.4 Error probability

Definition 3.30 Consider the q-ary symmetric channel where the receiver checks whether the received word $\mathbf{r}$ is a codeword or not, for instance by computing wether $H\mathbf{r}^T$ is zero or not for a chosen parity check matrix H, and asks for *retransmission* in case $\mathbf{r}$ is not a codeword as explained in Remark 2.25. Now it may occur that $\mathbf{r}$ is again a codeword but not equal to the codeword that was sent. This is called an *undetected error* . See [24].

Proposition 3.31 *Let $W_C(X,Y)$ be the weight enumerator of C. Then the probability of undetected error on a q-ary symmetric channel with cross-over probability p is given by*

$$P_{ue}(p) = W_C\left(1-p, \frac{p}{q-1}\right) - (1-p)^n.$$

Proof. Every codeword has the same probability of transmission and the code is linear. So without loss of generality we may assume that the zero word is sent. Hence

$$P_{ue}(p) = \frac{1}{|C|}\sum_{\mathbf{x}\in C}\sum_{\mathbf{x}\neq\mathbf{y}\in C} P(\mathbf{y}|\mathbf{x}) = \sum_{0\neq\mathbf{y}\in C} P(\mathbf{y}|0).$$

If the received codeword $\mathbf{y}$ has weight w, then w symbols are changed and the remaining $n-w$ symbols remained the same. So $P(\mathbf{y}|0) = (1-p)^{n-w}\left(\frac{p}{q-1}\right)^w$ by Remark 3.24. So

$$P_{ue}(p) = \sum_{w=1}^{n} A_w (1-p)^{n-w}\left(\frac{p}{q-1}\right)^w.$$

Substituting $X = 1-p$ and $Y = p/(q-1)$ in $W_C(X,Y)$ gives the desired result, since $A_0 = 1$. ◇

Remark 3.32 Now $P_{retr}(p) = 1 - P_{ue}(p)$ is the probability of *retransmission*.

Example 3.33 Let C be the binary triple repetition code. Then $P_{ue}(p) = p^3$, since $W_C(X,Y) = X^3 + Y^3$ by Example 3.5.

Example 3.34 Let C be the $[7,4,3]$ Hamming code. Then

$$P_{ue}(p) = 7(1-p)^4p^3 + 7(1-p)^3p^4 + p^7 = 7p^3 - 21p^4 + 21p^5 - 7p^6 + p^7$$

by Example 3.6.

Proposition 3.35 *Let $N(v,w,s)$ be the number of error patterns in $\mathbb{F}_q^n$ of weight w that are at distance s from a given word of weight v. Then*

$$N(v,w,s) = \sum_{0\leq i,j\leq n; i+2j+w=s+v} \binom{n-v}{j+w-v}\binom{v}{i}\binom{v-i}{j}(q-1)^{j+w-v}(q-2)^i.$$

Proof. Consider a given word $\mathbf{x}$ of weight v. Let $\mathbf{y}$ be a word of weight w and distance s to $\mathbf{x}$. Suppose that $\mathbf{y}$ has k nonzero coordinates in the complement of the support of $\mathbf{x}$, j zero coordinates in the support of $\mathbf{x}$, and i nonzero coordinates in the support of $\mathbf{x}$ that are distinct form the coordinates of $\mathbf{x}$. Then $s = d(\mathbf{x},\mathbf{y}) = i+j+k$ and $\mathrm{wt}(\mathbf{y}) = w = v+k-j$. There are $\binom{n-v}{k}$ possible subsets of k elements in the complement of the support of $\mathbf{x}$ and there are $(q-1)^k$ possible choices for the nonzero symbols at the corresponding coordinates. There are $\binom{v}{i}$ possible subsets of i elements in the support of $\mathbf{x}$ and there are $(q-2)^i$ possible choices of the symbols at those positions that are distinct form the coordinates of $\mathbf{x}$. There are $\binom{v-i}{j}$ possible subsets of j elements in the support of $\mathbf{x}$ that are zero at those positions. Hence

$$N(v,w,s) = \sum_{i+j+k=s, v+k-j=w} \left[\binom{n-v}{k}(q-1)^k\right]\left[\binom{v}{i}(q-2)^i\right]\binom{v-i}{j}.$$

Rewriting this formula using $k = j+w-v$ gives the desired result. ◇

Proposition 3.36 *The probability of decoding error of a decoder that corrects up to t errors with $2t+1 \leq d$ of a code C of minimum distance d on a q-ary symmetric channel with cross-over probability p is given by*

$$P_{de}(p) = \sum_{w=0}^{n} \left(\frac{p}{q-1}\right)^{w} (1-p)^{n-w} \sum_{s=0}^{t} \sum_{v=1}^{n} A_v N(v,w,s).$$

Proof. This is left as an exercise. ◇

Proposition 3.37 *The probability of correct decoding of the coset leader decoder on a q-ary symmetric channel with cross-over probability p is given by*

$$P_{cd}(p) = \alpha_C \left(1-p, \frac{p}{q-1}\right).$$

Proof. This is left as an exercise or see [26]. ◇

Another application of the coset leader weight enumerator is given in steganography, where messages are secretly hidden in a content such as a picture. See [27].

3.5 Exercises

3.1 Give a proof of Proposition 3.36.

3.2 Give a proof of Proposition 3.37.

4 Codes, projective systems and arrangements

Let $\mathbb{F}$ be a field. A *projective system* $\mathcal{P} = (P_1, \ldots, P_n)$ in $\mathbb{P}^r(\mathbb{F})$, the projective space over $\mathbb{F}$ of dimension r is an n-tuple of points P_j in this projective space, such that not all these points lie in a hyperplane. See [36, §1.1.2]. Let P_j be given by the homogeneous coordinates $(p_{0j} : p_{1j} : \ldots : p_{rj})$. Let $G_{\mathcal{P}}$ be the $(r+1) \times n$ matrix with $(p_{0j}, p_{1j}, \ldots, p_{rj})^T$ as j-th column. Then $G_{\mathcal{P}}$ has rank $r+1$, since not all points lie in a hyperplane. If $\mathbb{F}$ is a finite field, then $G_{\mathcal{P}}$ is the generator matrix of a nondegenerate code over $\mathbb{F}$ of length n and dimension $r+1$. Conversely, let G be a generator matrix of a nondegenerate code C of dimension k over $\mathbb{F}_q$. Then G has no zero columns. Take the columns of G as homogeneous coordinates of points in $\mathbb{P}^{k-1}(\mathbb{F}_q)$. This gives the projective system $\mathcal{P}_G$ over $\mathbb{F}_q$ of G.

Proposition 4.1 *Let C be a nondegenerate code over $\mathbb{F}_q$ of length n and dimension k with generator matrix G. Let $\mathcal{P}_G$ be the projective system of G. The code has minimum distance d if and only if $n-d$ is the maximal number of points of $\mathcal{P}_G$ in a hyperplane of $\mathbb{P}^{k-1}(\mathbb{F})$.*

Proof. See [37]. ◇

An n-tuple $(H_1, \ldots, H_n)$ of hyperplanes in $\mathbb{F}^k$ is called an *arrangement* in $\mathbb{F}^k$. The arrangement is called *central* if all the hyperplanes contain $\{0\}$. A central arrangement is called *essential* if the intersection of all its hyperplanes is equal to $\{0\}$. In case of an essential arrangement one considers the hyperplanes in $\mathbb{P}^{k-1}(\mathbb{F})$. Note that projective systems and arrangements are dual notions and that there is a one-to-one correspondence between generalized equivalence classes of non-degenerate $[n,k,d]$ codes over $\mathbb{F}_q$,

equivalence classes of projective systems over $\mathbb{F}_q$ of n points in $\mathbb{P}^{k-1}(\mathbb{F}_q)$ and equivalence classes of essential arrangements of n hyperplanes in $\mathbb{P}^{k-1}(\mathbb{F}_q)$.

The translation for an arrangement of Proposition 4.1 gives for the minimum distance d.

Proposition 4.2 *Let C be a nondegenerate code over $\mathbb{F}_q$ with generator matrix G. Let $\mathbf{c}$ be a codeword $\mathbf{c} = \mathbf{x}G$ for some $\mathbf{x} \in \mathbb{F}_q^k$. Then $n - \mathrm{wt}(\mathbf{c})$ is equal to the number of hyperplanes in $\mathcal{A}_G$ through $\mathbf{x}$.*

Proof. [21, 36]. ◇

A code C is called *projective* if $d(C^\perp) \geq 3$. Let G be generator matrix C. Then C is projective if and only if C is nondegenerate and any two columns of G are independent. So C is projective if and only if C is nondegenerate and the hyperplanes of $\mathcal{A}_G$ are mutually distinct.

5 The extended and generalized weight enumerator

The number A_w of codewords of weight w equals the number of points that are on exactly $n - w$ of the hyperplanes in $\mathcal{A}_G$, by Proposition 4.2. In particular A_n is equal to the number of points that is in the complement of the union of these hyperplanes in $\mathbb{F}_q^k$. This number can be computed by the *principle of inclusion/exclusion*:

$$A_n = q^k - |H_1 \cup \cdots \cup H_n| =$$

$$q^k + \sum_{w=1}^{n} (-1)^w \sum_{i_1 < \cdots < i_w} |H_{i_1} \cap \cdots \cap H_{i_w}|.$$

The following notations are introduced to find a formalism as above for the computation of the weight enumerator. This method is based on Katsman and Tsfasman [21]. Later we will encounter two more methods: by geometric lattices and the characteristic polynomial in Section 6 and by matroids and the Tutte polynomial in Section 7.

Definition 5.1 For a subset J of $[n] := \{1, 2, \ldots, n\}$ define

$$\begin{aligned} C(J) &= \{\mathbf{c} \in C \mid c_j = 0 \text{ for all } j \in J\} \\ l(J) &= \dim C(J) \\ B_J &= q^{l(J)} - 1 \\ B_t &= \sum_{|J|=t} B_J. \end{aligned}$$

The encoding map $\mathbf{x} \mapsto \mathbf{x}G = \mathbf{c}$ from vectors $\mathbf{x} \in \mathbb{F}_q^k$ to codewords gives the following isomorphism of vector spaces

$$\bigcap_{j \in J} H_j \cong C(J)$$

by Proposition 4.2. Furthermore B_J is equal to the number of nonzero codewords $\mathbf{c}$ that are zero at al j in J, and this is equal to the number of nonzero elements of the intersection $\bigcap_{j \in J} H_j$.

Proposition 5.2 *We have the following connection between the B_t and the weight distribution of a code:*

$$B_t = \sum_{w=0}^{n} \binom{n-w}{t} A_w \ .$$

Proof. Count in two ways the number of elements of the set

$$\{(J, \mathbf{c}) : J \subseteq [n], |J| = t, \mathbf{c} \in C, \mathbf{c} \neq 0\}.$$

◇

We will generalize this idea to the determine the generalized weight enumerators.

5.1 Generalized weight enumerators

We first generalize the weight distribution in the following way, see [22, 44]. Instead of looking at words of C, we consider all the subcodes of C of a certain dimension r. We say that the *weight of a subcode* (also called the *effective length* or *support weight*) is equal to n minus the number of coordinates which are zero for every word in the subcode. The smallest weight for which a subcode of dimension r exists, is called the *r-th generalized Hamming weight* of C. To summarize:

Definition 5.3 Let D be an r-dimensional subcode of the $[n, k]$ code C. Then we define

$$\mathrm{supp}(D) = \{i \in [n] : \text{ there is an } \mathbf{x} \in D : x_i \neq 0\},$$
$$\mathrm{wt}(D) = |\mathrm{supp}(D)| \quad \text{and} \quad d_r = \min\{\mathrm{wt}(D) : D \subseteq C \text{ subcode}, \dim D = r\}.$$

Note that $d_0 = 0$ and $d_1 = d$, the minimum distance of the code. The number of subcodes with a given weight w and dimension r is denoted by A_w^r. Together they form the *r-th generalized weight distribution* of the code. Just as with the ordinary weight distribution, we can make a polynomial with the distribution as coefficients: the *generalized weight enumerator*.

Definition 5.4 The r-th generalized weight enumerator is given by

$$W_C^r(X, Y) = \sum_{w=0}^{n} A_w^r X^{n-w} Y^w,$$

where $A_w^r = |\{D \subseteq C : \dim D = r, \mathrm{wt}(D) = w\}|$.

We can see from this definition that $A_0^0 = 1$ and $A_0^r = 0$ for all $0 < r \leq k$. Furthermore, every 1-dimensional subspace of C contains $q - 1$ non-zero codewords, so $(q-1)A_w^1 = A_w$ for $0 < w \leq n$. This means we can find back the original weight enumerator by using $W_C(X, Y) = W_C^0(X, Y) + (q-1)W_C^1(X, Y)$.

We will give a way to determine the generalized weight enumerator of a linear $[n, k]$ code C over $\mathbb{F}_q$. We give two lemmas about the determination of $l(J)$, which will become useful later.

Lemma 5.5 *Let C be a linear code with generator matrix G. Let $J \subseteq [n]$ and $|J| = t$. Let G_J be the $k \times t$ submatrix of G existing of the columns of G indexed by J, and let $r(J)$ be the rank of G_J. Then the dimension $l(J)$ is equal to $k - r(J)$.*

Proof. Let C_J be the code generated by G_J. Consider C_J as a subcode of C, so a word of C_J has zeros on the coordinates not indexed by J. Then we have $C_J \cong C/C(J)$. It follows that $\dim C_J = \dim C - \dim C(J)$ so $l(J) = k - r(J)$. $\diamond$

Lemma 5.6 *Let d and $d^\perp$ be the minimum distance of C and $C^\perp$, respectively. Let $J \subseteq [n]$ and $|J| = t$. Then we have*

$$l(J) = \begin{cases} k-t & \text{for all } t < d^\perp \\ 0 & \text{for all } t > n-d \end{cases}$$

Proof. Let $|J| = t$, $t > n-d$ and let $\mathbf{c} \in C(J)$. Then J is contained in the complement of $\mathrm{supp}(\mathbf{c})$, so $t \leq n - \mathrm{wt}(\mathbf{c})$. It follows that $\mathrm{wt}(\mathbf{c}) \leq n - t < d$, so $\mathbf{c}$ is the zero word and therefore $l(J)$=0.
Let G be a generator matrix for C, then G is also a parity check matrix for $C^\perp$. We saw in lemma 5.5 that $l(J) = k - r(J)$, where $r(J)$ is the rank of the matrix formed by the columns of G indexed by J. Let $t < d^\perp$, then every t-tuple of columns of G is linearly independent by Proposition 2.43, so $r(J) = t$ and $l(J) = k - t$. $\diamond$

Note that by the Singleton bound, we have $d^\perp \leq n-(n-k)+1 = k+1$ and $n-d \geq k-1$, so for $t = k$ both of the above cases apply. This is no problem, because if $t = k$ then $k - t = 0$.

Definition 5.7 We introduce the following notations:

$$[m,r]_q = \prod_{i=0}^{r-1} (q^m - q^i)$$

$$\langle r \rangle_q = [r,r]_q$$

$$\begin{bmatrix} k \\ r \end{bmatrix}_q = \frac{[k,r]_q}{\langle r \rangle_q}.$$

Remark 5.8 The first number is equal to the number of $m \times r$ matrices of rank r over $\mathbb{F}_q$. The second is the number of bases of $\mathbb{F}_q^r$. The third number is the Gaussian binomial, and it represents the number of r-dimensional subspaces of $\mathbb{F}_q^k$.

Definition 5.9 For $J \subseteq [n]$ and $r \geq 0$ an integer we define:

$$\begin{aligned} B_J^r &= |\{D \subseteq C(J) : D \text{ subspace of dimension } r\}| \\ B_t^r &= \sum_{|J|=t} B_J^r \end{aligned}$$

Note that $B_J^r = \begin{bmatrix} l(J) \\ r \end{bmatrix}_q$. For $r = 0$ this gives $B_t^0 = \binom{n}{t}$. So we see that in general $l(J) = 0$ does not imply $B_J^r = 0$, because $\begin{bmatrix} 0 \\ 0 \end{bmatrix}_q = 1$. But if $r \neq 0$, we do have that $l(J) = 0$ implies $B_J^r = 0$ and $B_t^r = 0$.

Proposition 5.10 *Let d_r be the r-th generalized Hamming weight of C, and $d^\perp$ the minimum distance of the dual code $C^\perp$. Then we have*

$$B_t^r = \begin{cases} \binom{n}{t} \begin{bmatrix} k-t \\ r \end{bmatrix}_q & \text{for all } t < d^\perp \\ 0 & \text{for all } t > n - d_r \end{cases}$$

Proof. The first case is is a direct corollary of lemma 5.6, since there are $\binom{n}{t}$ subsets $J \subseteq [n]$ with $|J| = t$. The proof of the second case goes analogous to the proof of the same lemma: let $|J| = t$, $t > n - d_r$ and suppose there is a subspace $D \subseteq C(J)$ of dimension r. Then J is contained in the complement of $\mathrm{supp}(D)$, so $t \leq n - \mathrm{wt}(D)$. It follows that $\mathrm{wt}(D) \leq n - t < d_r$, which is impossible, so such a D does not exist. So $B_J^r = 0$ for all J with $|J| = t$ and $t > n - d_r$, and therefore $B_t^r = 0$ for $t > n - d_r$. $\diamond$

We can check that the formula is well-defined: if $t < d^\perp$ then $l(J) = k - t$. If also $t > n - d_r$, we have $t > n - d_r \geq k - r$ by the generalized Singleton bound. This implies $r > k - t = l(J)$, so $\begin{bmatrix} k-t \\ r \end{bmatrix}_q = 0$.
The relation between B_t^r and A_w^r becomes clear in the next proposition.

Proposition 5.11 *The following formula holds:*

$$B_t^r = \sum_{w=0}^{n} \binom{n-w}{t} A_w^r.$$

Proof. We will count the elements of the set

$$\mathcal{B}_t^r = \{(D, J) : J \subseteq [n], |J| = t, D \subseteq C(J) \text{ subspace of dimension } r\}$$

in two different ways. For each J with $|J| = t$ there are B_J^r pairs (D, J) in $\mathcal{B}_t^r$, so the total number of elements in this set is $\sum_{|J|=t} B_J^r = B_t^r$. On the other hand, let D be an r-dimensional subcode of C with $\mathrm{wt}(D) = w$. There are A_w^r possibilities for such a D. If we want to find a J such that $D \subseteq C(J)$, we have to pick t coordinates from the $n - w$ all-zero coordinates of D. Summation over all w proves the given formula. $\diamond$

Note that because $A_w^r = 0$ for all $w < d_r$, we can start summation at $w = d_r$. We can end summation at $w = n - t$ because for $t > n - w$ we have $\binom{n-w}{t} = 0$. So the formula can be rewritten as

$$B_t^r = \sum_{w=d_r}^{n-t} \binom{n-w}{t} A_w^r.$$

In practice, we will often prefer the summation given in the proposition.

Theorem 5.12 *The generalized weight enumerator is given by the following formula:*

$$W_C^r(X, Y) = \sum_{t=0}^{n} B_t^r (X - Y)^t Y^{n-t}.$$

Proof. By using the previous proposition, changing the order of summation and using the binomial expansion of $X^{n-w} = ((X - Y) + Y)^{n-w}$ we have

$$\begin{aligned}
\sum_{t=0}^{n} B_t^r (X-Y)^t Y^{n-t} &= \sum_{t=0}^{n} \sum_{w=0}^{n} \binom{n-w}{t} A_w^r (X-Y)^t Y^{n-t} \\
&= \sum_{w=0}^{n} A_w^r \left(\sum_{t=0}^{n-w} \binom{n-w}{t} (X-Y)^t Y^{n-w-t} \right) Y^w \\
&= \sum_{w=0}^{n} A_w^r X^{n-w} Y^w \\
&= W_C^r(X, Y).
\end{aligned}$$

In the second step, we can let the summation over t run to $n-w$ instead of n because $\binom{n-w}{t} = 0$ for $t > n-w$. ◇

It is possible to determine the A_w^r directly from the B_t^r, by using the next proposition.

Proposition 5.13 *The following formula holds:*

$$A_w^r = \sum_{t=n-w}^{n} (-1)^{n+w+t} \binom{t}{n-w} B_t^r.$$

There are several ways to prove this proposition. One is to reverse the argument from Theorem 5.12, which we will not use here. Instead, we first prove the following general lemma:

Lemma 5.14 *Let V be a vector space of dimension $n+1$ and let $\mathbf{a} = (a_0, \ldots, a_n)$ and $\mathbf{b} = (b_0, \ldots, b_n)$ be vectors in V. Then the following formulas are equivalent:*

$$a_j = \sum_{i=0}^{n} \binom{i}{j} b_i, \qquad b_j = \sum_{i=j}^{n} (-1)^{i+j} \binom{i}{j} a_i.$$

Proof. We can view the relations between $\mathbf{a}$ and $\mathbf{b}$ as linear transformations, given by the matrices $\left(\binom{i}{j}\right)_{i,j=0,\ldots,n}$ and $\left((-1)^{i+j}\binom{i}{j}\right)_{i,j=0,\ldots,n}$. So it is sufficient to prove that these matrices are each other's inverse. We calculate the entry on the i-th row and j-th column. Note that we can start the summation at $l = j$, because for $l < j$ we have $\binom{l}{j} = 0$.

$$\begin{aligned}
\sum_{l=j}^{i} (-1)^{j+l} \binom{i}{l}\binom{l}{j} &= \sum_{l=j}^{i} (-1)^{l-j} \binom{i}{j}\binom{i-j}{l-j} \\
&= \sum_{l=0}^{i-j} (-1)^{l} \binom{i}{j}\binom{i-j}{l} \\
&= \binom{i}{j}(1-1)^{i-j} \\
&= \delta_{ij}.
\end{aligned}$$

Here δ_{ij} is the Kronecker-delta. So the product matrix is exactly the $(n+1) \times (n+1)$ identity matrix, and therefore the matrices are each other's inverse. ◇

Proof. (Proposition 5.13) The proposition is now a direct consequence of Proposition 5.11 and Lemma 5.14. ◇

5.2 Extended weight enumerator

Let C be an $[n,k]$ code over $\mathbb{F}_q$ with generator matrix G. Then we can form the $[n,k]$ code $C \otimes \mathbb{F}_{q^m}$ over $\mathbb{F}_{q^m}$ by taking all $\mathbb{F}_{q^m}$-linear combinations of the codewords in C. We call this the *extension code* of C over $\mathbb{F}_{q^m}$. We denote the number of codewords in $C \otimes \mathbb{F}$ of weight w by $A_{C\otimes\mathbb{F}_q,w}$. We can determine the weight enumerator of such an extension code by using only the code C.
By embedding its entries in $\mathbb{F}_{q^m}$, we find that G is also a generator matrix for the

extension code $C \otimes \mathbb{F}_{q^m}$. In Lemma 5.5 we saw that $l(J) = k - r(J)$. Because $r(J)$ is independent of the extension field $\mathbb{F}_{q^m}$, we have $\dim_{\mathbb{F}_q} C(J) = \dim_{\mathbb{F}_{q^m}}(C \otimes \mathbb{F}_{q^m})(J)$. This motivates the usage of T as a variable for q^m in the next definition, which is an extention of Definition 5.1.

Definition 5.15 Let C be a linear code over $\mathbb{F}_q$. Then we define

$$B_J(T) = T^{l(J)} - 1$$

$$B_t(T) = \sum_{|J|=t} B_J(T)$$

The *extended weight enumerator* is given by

$$W_C(X, Y, T) = X^n + \sum_{t=0}^{n} B_t(T)(X - Y)^t Y^{n-t}.$$

Note that $B_J(q^m)$ is the number of nonzero codewords in $(C \otimes \mathbb{F}_{q^m})(J)$.

Proposition 5.16 *Let d and $d^\perp$ be the minimum distance of C and $C^\perp$ respectively. Then we have*

$$B_t(T) = \begin{cases} \binom{n}{t}(T^{k-t} - 1) & \text{for all } t < d^\perp \\ 0 & \text{for all } t > n - d \end{cases}$$

Proof. This is a direct consequence of Lemma 5.6. For $t < d^\perp$ we have $l(J) = k - t$, so $B_J(T) = T^{k-t} - 1$ and $B_t(T) = \binom{n}{t}(T^{k-t} - 1)$. For $t > n - d$ we have $l(J) = 0$, so $B_J(T) = 0$ and $B_t(T) = 0$. ◇

Theorem 5.17 *The following holds:*

$$W_C(X, Y, T) = \sum_{w=0}^{n} A_w(T) X^{n-w} Y^w$$

with $A_w(T) \in \mathbb{Z}[T]$ given by $A_0(T) = 1$ and

$$A_w(T) = \sum_{t=n-w}^{n} (-1)^{n+w+t} \binom{t}{n-w} B_t(T)$$

for $0 < w \leq n$.

Proof. Note that $A_w(T) = 0$ for $0 < w < d$ because the summation is empty. By substituting $w = n - t + j$ and reversing the order of summation, we have

$$\begin{aligned} W_C(X, Y, T) &= X^n + \sum_{t=0}^{n} B_t(T)(X - Y)^t Y^{n-t} \\ &= X^n + \sum_{t=0}^{n} B_t(T) \left(\sum_{j=0}^{t} \binom{t}{j} (-1)^j X^{t-j} Y^j \right) Y^{n-t} \end{aligned}$$

$$
\begin{aligned}
&= X^n + \sum_{t=0}^{n} \sum_{j=0}^{t} (-1)^j \binom{t}{j} B_t(T) X^{t-j} Y^{n-t+j} \\
&= X^n + \sum_{t=0}^{n} \sum_{w=n-t}^{n} (-1)^{t-n+w} \binom{t}{t-n+w} B_t(T) X^{n-w} Y^w \\
&= X^n + \sum_{w=0}^{n} \sum_{t=n-w}^{n} (-1)^{n+w+t} \binom{t}{n-w} B_t(T) X^{n-w} Y^w
\end{aligned}
$$

Hence $W_C(X,Y,T)$ is of the form $\sum_{w=0}^{n} A_w(T) X^{n-w} Y^w$ with $A_w(T)$ of the form given in the theorem. ◇

Note that in the definition of $A_w(T)$ we can let the summation over t run to $n-d$ instead of n, because $B_t(T) = 0$ for $t > n-d$.

Proposition 5.18 *The following formula holds:*

$$B_t(T) = \sum_{w=d}^{n-t} \binom{n-w}{t} A_w(T).$$

Proof. The statement is a direct consequence of Lemma 5.14 and Theorem 5.17. ◇

As we said before, the motivation for looking at the extended weight enumerator comes from the extension codes. In the next proposition we show that the extended weight enumerator for $T = q^m$ is indeed the weight enumerator of the extensioncode $C \otimes \mathbb{F}_{q^m}$.

Proposition 5.19 *Let C be a linear $[n,k]$ code over $\mathbb{F}_q$. Then we have*

$$W_C(X,Y,q^m) = W_{C \otimes \mathbb{F}_{q^m}}(X,Y).$$

Proof. For $w = 0$ it is clear that $A_0(q^m) = A_0(C \otimes \mathbb{F}_{q^m}) = 1$, so assume $w \neq 0$. It is enough to show that $A_w(q^m) = (q^m - 1) A_w^1(C \otimes \mathbb{F}_{q^m})$. First we have

$$
\begin{aligned}
B_t(q^m) &= \sum_{|J|=t} B_J(q^m) \\
&= \sum_{|J|=t} |\{\mathbf{c} \in (C \otimes \mathbb{F}_{q^m})(J) : \mathbf{c} \neq 0\}| \\
&= (q^m - 1) \sum_{|J|=t} |\{D \subseteq (C \otimes \mathbb{F}_{q^m})(J) : \dim D = 1\} \\
&= (q^m - 1) B_t^1(C \otimes \mathbb{F}_{q^m}).
\end{aligned}
$$

We also know that $A_w(T)$ and $B_t(T)$ are related the same way as A_w^1 and B_t^1. Combining this proves the statement. ◇

Therefore we can view $W_C(X,Y,T)$ as the weight enumerator of the extension code over the algebraic closure of $\mathbb{F}_q$. This means we can find a relation with the two variable zeta-function of a code, see Duursma [14]. The notion of the extended weight enumerator was first introduced by Helleseth, Kløve and Mykkeltveit [17, 22] and later studied by by [44]. This notion has applications in the wire-tap channel II [30] and trellis complexity [15].

For further applications, the next way of writing the extended weight enumerator will be useful:

Proposition 5.20 *The extended weight enumerator of a linear code C can be written as*

$$W_C(X,Y,T) = \sum_{t=0}^{n} \sum_{|J|=t} T^{l(J)}(X-Y)^t Y^{n-t}.$$

Proof. By rewriting and using the binomial expansion of $((X-Y)+Y)^n$, we get

$$\begin{aligned}
& \sum_{t=0}^{n} \sum_{|J|=t} T^{l(J)}(X-Y)^t Y^{n-t} \\
= & \sum_{t=0}^{n} (X-Y)^t Y^{n-t} \sum_{|J|=t} \big((T^{l(J)}-1)+1\big) \\
= & \sum_{t=0}^{n} (X-Y)^t Y^{n-t} \left(\sum_{|J|=t} (T^{l(J)}-1) + \binom{n}{t} \right) \\
= & \sum_{t=0}^{n} B_t(T)(X-Y)^t Y^{n-t} + \sum_{t=0}^{n} \binom{n}{t} (X-Y)^t Y^{n-t} \\
= & \sum_{t=0}^{n} B_t(T)(X-Y)^t Y^{n-t} + X^n \\
= & \; W_C(X,Y,T)
\end{aligned}$$

◇

5.3 Connections

There is a connection between the extended weight enumerator and the generalized weight enumerators. We first proof the next proposition.

Proposition 5.21 *Let C be a linear $[n,k]$ code over $\mathbb{F}_q$, and let C^m be the linear subspace consisting of the $m \times n$ matrices over $\mathbb{F}_q$ whose rows are in C. Then there is an isomorphism of $\mathbb{F}_q$-vector spaces between $C \otimes \mathbb{F}_{q^m}$ and C^m.*

Proof. Let α be a primitive m-th root of unity in $\mathbb{F}_{q^m}$. Then we can write an element of $\mathbb{F}_{q^m}$ in an unique way on the basis $(1, \alpha, \alpha^2, \ldots, \alpha^{m-1})$ with coefficients in $\mathbb{F}_q$. If we do this for all the coordinates of a word in $C \otimes \mathbb{F}_{q^m}$, we get an $m \times n$ matrix over $\mathbb{F}_q$. The rows of this matrix are words of C, because C and $C \otimes \mathbb{F}_{q^m}$ have the same generator matrix. This map is clearly injective. There are $(q^m)^k = q^{km}$ words in $C \otimes \mathbb{F}_{q^m}$, and the number of elements of C^m is $(q^k)^m = q^{km}$, so our map is a bijection. It is given by

$$\left(\sum_{i=0}^{m-1} c_{i1}\alpha^i, \sum_{i=0}^{m-1} c_{i2}\alpha^i, \ldots, \sum_{i=0}^{m-1} c_{in}\alpha^i \right) \mapsto$$

$$\begin{pmatrix}
c_{01} & c_{02} & c_{03} & \cdots & c_{0n} \\
c_{11} & c_{12} & c_{13} & \cdots & c_{1n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
c_{(m-1)1} & c_{(m-1)2} & c_{(m-1)3} & \cdots & c_{(m-1)n}
\end{pmatrix}.$$

We see that the map is $\mathbb{F}_q$-linear, so it gives an isomorphism $C \otimes \mathbb{F}_{q^m} \to C^m$. $\diamond$

Note that this isomorphism depends on the choice of a primitive element α. The use of this isomorphism for the proof of Theorem 5.24 was suggested in [34] by Simonis. We also need the next subresult.

Lemma 5.22 *Let $\mathbf{c} \in C \otimes \mathbb{F}_{q^m}$ and $M \in C^m$ the corresponding $m \times n$ matrix under a given isomorphism. Let $D \subseteq C$ be the subcode generated by the rows of M. Then $wt(\mathbf{c}) = wt(D)$.*

Proof. If the j-th coordinate c_j of $\mathbf{c}$ is zero, then the j-th column of M consists of only zero's, because the representation of c_j on the basis $(1, \alpha, \alpha^2, \ldots, \alpha^{m-1})$ is unique. On the other hand, if the j-th column of M consists of all zeros, then c_j is also zero. Therefore $\mathrm{wt}(\mathbf{c}) = \mathrm{wt}(D)$. $\diamond$

Proposition 5.23 *Let C be a linear code over $\mathbb{F}_q$. Then the weight numerator of an extension code and the generalized weight enumerators are connected via*

$$A_w(q^m) = \sum_{r=0}^{m} [m, r]_q A_w^r.$$

Proof. We count the number of words in $C \otimes \mathbb{F}_{q^m}$ of weight w in two ways, using the bijection of Proposition 5.21. The first way is just by substituting $T = q^m$ in $A_w(T)$: this gives the left side of the equation. For the second way, note that every $M \in C^m$ generates a subcode of C whose weight is equal to the weight of the corresponding word in $C \otimes \mathbb{F}_{q^m}$. Fix this weight w and a dimension r: there are A_w^r subcodes of C of dimension r and weight w. Every such subcode is generated by an $r \times n$ matrix whose rows are words of C. Left multiplication by an $m \times r$ matrix of rank r gives an element of C^m which generates the same subcode of C, and all such elements of C^m are obtained this way. The number of $m \times r$ matrices of rank r is $[m, r]_q$, so summation over all dimensions r gives

$$A_w(q^m) = \sum_{r=0}^{k} [m, r]_q A_w^r.$$

We can let the summation run to m, because $A_w^r = 0$ for $r > k$ and $[m, r]_q = 0$ for $r > m$. This proves the given formula. $\diamond$

This result first appears in [17, Theorem 3.2], although the term "generalized weight enumerator" was yet to be invented. In general, we have the following theorem.

Theorem 5.24 *Let C be a linear code over $\mathbb{F}_q$. Then the extended weight numerator and the generalized weight enumerator are connected via*

$$W_C(X, Y, T) = \sum_{r=0}^{k} \left(\prod_{j=0}^{r-1} (T - q^j) \right) W_C^r(X, Y).$$

Proof. If we know A_w^r for all r, we can determine $A_w(q^m)$ for every m. If we have $k + 1$ values of m for which $A_w(q^m)$ is known, we can use Lagrange interpolation to find $A_w(T)$, for this is a polynomial in T of degree at most k. In fact, we have

$$A_w(T) = \sum_{r=0}^{k} \left(\prod_{j=0}^{r-1} (T - q^j) \right) A_w^r.$$

This formula has the right degree and is correct for $T = q^m$ for all integer values $m \geq 0$, so we know it must be the correct polynomial. Therefore the theorem follows. $\diamond$

The converse of the theorem is also true: we can write the generalized weight enumerator in terms of the extendend weight enumerator.

Theorem 5.25 *Let C be a linear code over $\mathbb{F}_q$. Then the generalized weight enumerator and the extended weight enumerator are connected via*

$$W_C^r(X,Y) = \frac{1}{\langle r \rangle_q} \sum_{j=0}^{r} \begin{bmatrix} r \\ j \end{bmatrix}_q (-1)^{r-j} q^{\binom{r-j}{2}} \; W_C(X,Y,q^j).$$

Proof. We consider the generalized weight enumerator in terms of Proposition 5.20. Then rewriting gives the following:

$$\begin{aligned}
W_C^r(X,Y) &= \sum_{t=0}^{n} B_t^r (X-Y)^t Y^{n-t} \\
&= \sum_{t=0}^{n} \sum_{|J|=t} \begin{bmatrix} l(J) \\ r \end{bmatrix}_q (X-Y)^t Y^{n-t} \\
&= \sum_{t=0}^{n} \sum_{|J|=t} \left(\prod_{j=0}^{r-1} \frac{q^{l(J)} - q^j}{q^r - q^j} \right) (X-Y)^t Y^{n-t} \\
&= \frac{1}{\prod_{v=0}^{r-1}(q^r - q^v)} \sum_{t=0}^{n} \sum_{|J|=t} \left(\prod_{j=0}^{r-1} (q^{l(J)} - q^j) \right) (X-Y)^t Y^{n-t} \\
&= \frac{1}{\langle r \rangle_q} \sum_{t=0}^{n} \sum_{|J|=t} \sum_{j=0}^{r} \begin{bmatrix} r \\ j \end{bmatrix}_q (-1)^{r-j} q^{\binom{r-j}{2}} q^{j \cdot l(J)} (X-Y)^t Y^{n-t} \\
&= \frac{1}{\langle r \rangle_q} \sum_{j=0}^{r} \begin{bmatrix} r \\ j \end{bmatrix}_q (-1)^{r-j} q^{\binom{r-j}{2}} \sum_{t=0}^{n} \sum_{|J|=t} (q^j)^{l(J)} (X-Y)^t Y^{n-t} \\
&= \frac{1}{\langle r \rangle_q} \sum_{j=0}^{r} \begin{bmatrix} r \\ j \end{bmatrix}_q (-1)^{r-j} q^{\binom{r-j}{2}} \; W_C(X,Y,q^j)
\end{aligned}$$

In the fourth step, we use the following identity (see [22]), which can be proven by induction:

$$\prod_{j=0}^{r-1} (Z - q^j) = \sum_{j=0}^{r} \begin{bmatrix} r \\ j \end{bmatrix}_q (-1)^{r-j} q^{\binom{r-j}{2}} Z^j.$$

$\diamond$

5.4 MDS-codes

We can use the theory in the second chapter to calculate the weight distribution, generalized weight distribution, and extended weight distribution of a linear $[n,k]$ code C. This is done by determining the values $l(J)$ for each $J \subseteq [n]$. In general, we have to look at the 2^n different subcodes of C to find the $l(J)$, but for the special case of MDS codes we can find the weight distributions much faster.

Proposition 5.26 *Let C be a linear $[n,k]$ MDS code, and let $J \subseteq [n]$. Then we have*

$$l(J) = \begin{cases} 0 & \text{for } t > k \\ k-t & \text{for } t \leq k \end{cases}$$

so for a given t the value of $l(J)$ is independent of the choice of J.

Proof. We know that the dual of a MDS code is also MDS, so $d^\perp = k+1$. Now use $d = n-k+1$ in lemma 5.6. ◇

Now that we know all the $l(J)$ for an MDS code, it is easy to find the weight distribution. We will give the construction for the generalized weight enumerator here: the case of the extended weight enumerator goes similar.

Theorem 5.27 *Let C be a MDS code with parameters $[n,k]$. Then the generalized weight distribution is given by*

$$A_w^r = \binom{n}{w} \sum_{j=0}^{w-d} (-1)^j \binom{w}{j} \begin{bmatrix} w-d+1-j \\ r \end{bmatrix}_q .$$

Proof. We know from the proposition that for a MDS code, B_t^r depends only on the size of J, so $B_t^r = \binom{n}{t} \begin{bmatrix} k-t \\ r \end{bmatrix}_q$. Using this in the formula for A_w^r and substituting $j = t-n+w$, we have

$$\begin{aligned} A_w^r &= \sum_{t=n-w}^{n-d_r} (-1)^{n+w+t} \binom{t}{n-w} B_t^r \\ &= \sum_{t=n-w}^{n-d_r} (-1)^{t-n+w} \binom{t}{n-w} \binom{n}{t} \begin{bmatrix} k-t \\ r \end{bmatrix}_q \\ &= \sum_{j=0}^{w-d_r} (-1)^j \binom{n}{w} \binom{w}{j} \begin{bmatrix} k+w-n-j \\ r \end{bmatrix}_q \\ &= \binom{n}{w} \sum_{j=0}^{w-d_r} (-1)^j \binom{w}{j} \begin{bmatrix} w-d+1-j \\ r \end{bmatrix}_q . \end{aligned}$$

In the second step, we are using the binomial equivalence

$$\binom{n}{t}\binom{t}{n-w} = \binom{n}{n-w}\binom{n-(n-w)}{t-(n-w)} = \binom{n}{w}\binom{w}{n-t}.$$

◇

So, for all MDS-codes with given parameters $[n,k]$ the extended and generalized weight distributions are the same. But not all such codes are equivalent. We can conclude from this, that the generalized extended weight enumerator is not enough to distinguish between codes with the same parameters.

6 Lattices and the characteristic polynomial

In this section we consider the characteristic polynomial of codes and arrangemetns of hyperplanes using the the theory of posets and lattices and the Möbius function. See [12, 28, 33].

6.1 Posets, the Möbius function and lattices

Definition 6.1 Let L be a set and $\leq$ a relation on L that is *reflexive*, *anti-symmetric* and *transitive*. Then the pair $(L, \leq)$ or just L is called a *poset* with *partial order* $\leq$ on the set L. Define $x < y$ if $x \leq y$ and $x \neq y$. The elements x and y in L are *comparable* if $x \leq y$ or $y \leq x$. A poset L is called a *linear order* if every two elements are comparable. Define $L_x = \{y \in L | x \leq y\}$ and $L^x = \{y \in L | y \leq x\}$ and the *the interval between* x and y by $[x, y] = \{z \in L | x \leq z \leq y\}$. Notice that $[x, y] = L_x \cap L^y$.

Definition 6.2 Let $(L, \leq)$ be a poset. A *chain of length* r *from* x *to* y *in* L is a sequence of elements $x_0, x_1, \ldots, x_r$ in L such that

$$x = x_0 < x_1 < \cdots < x_r = y.$$

Let r be a number. Let x, y in L. Then $c_r(x, y)$ denotes the number of chains of length r from x to y. Now $c_r(x, y)$ is finite if L is finite. The poset is called *locally finite* if $c_r(x, y)$ is finite for all $x, y \in L$ and every number r.

Proposition 6.3 *Let L be a locally finite poset. Let x, y in L and $x \leq y$. Then*

(C.1) $c_0(x, x) = 1$ *and* $c_0(x, y) = 0$ *if* $x < y$.

(C.2) $c_{r+1}(x, y) = \sum_{x \leq z < y} c_r(x, z) = \sum_{x < z \leq y} c_r(z, y)$.

Proof. The statement (C.1) is trivial. Let $z < y$ and $x = x_0 < x_1 < \cdots < x_r = z$ a chain of length r from x to z, then $x = x_0 < x_1 < \cdots < x_r < x_{r+1} = y$ is a chain of length $r + 1$ from x to y, and every chain of length $r + 1$ from x to y is obtained uniquely in this way. Hence $c_{r+1}(x, y) = \sum_{x \leq z < y} c_r(x, z)$. The last equality is proved similarly. ◇

Definition 6.4 The *Möbius function* of L, denoted by μ_L or μ is defined by

$$\mu(x, y) = \sum_{r=0}^{\infty} (-1)^r c_r(x, y).$$

Proposition 6.5 *Let L be a locally finite poset. Then for all x, y in L:*

(M.1) $\mu(x, x) = 1$.

(M.2) If $x < y$, then $\sum_{x \leq z \leq y} \mu(x, z) = \sum_{x \leq z \leq y} \mu(z, y) = 0$.

Proof. This is left as an exercise. ◇

Remark 6.6 Proposition 6.5 can be used as an alternative to compute $\mu(x, y)$ by induction:

$$\mu(x, y) = -\sum_{x \leq z < y} \mu(x, z).$$

Definition 6.7 Let L be a poset. If L has an element 0_L such that 0_L is the unique minimal element of L, then 0_L is called the *minimum* of L. Similarly 1_L is called the *maximum* of L if 1_L is the unique maximal element of L. If x, y in L and $x \leq y$, then the interval $[x, y]$ has x as minimum and y as maximum. Suppose that L has 0_L and 1_L as minimum and maximum also denoted by 0 and 1, respectively. Then $0 \leq x \leq 1$ for all $x \in L$. Define $\mu(x) = \mu(0, x)$ and $\mu(L) = \mu(0, 1)$ if L is finite.

Definition 6.8 Let L be a locally finite poset with minimum element 0. Let A be an abelian group and $f : L \to A$ a map from L to A. The *sum function* $\tilde{f}$ of f is defined by

$$\tilde{f}(x) = \sum_{y \leq x} f(y).$$

Proposition 6.9 *Let L be a locally finite poset with minimum element 0. Then the* Möbius inversion formula *holds:*

$$f(x) = \sum_{y \leq x} \mu(y, x) \cdot \tilde{f}(y).$$

Proof. This is left as an exercise. ◇

Example 6.10 Let $f(x) = 1$ if $x = 0$ and $f(x) = 0$ otherwise. Then the sum function $\tilde{f}(x) = \sum_{y \leq x} f(y)$ is constant 1 for all x. The Möbius inversion formula gives that

$$\sum_{y \leq x} \mu(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{if } x > 0, \end{cases}$$

which is a special case of Proposition 6.5.

Remark 6.11 Let $(L, \leq)$ be a poset. Let $\leq_R$ be the *reverse relation* on L defined by $x \leq_R y$ if and only if $y \leq x$. Then $(L, \leq_R)$ is a poset. Let $(L, \leq)$ be locally finite with Möbius function μ. Then the number of chains of length r from x to y in $(L, \leq_R)$ is the same as the number of chains of length r from y to x in $(L, \leq)$. Hence $(L, \leq_R)$ is locally finite with Möbius function μ_R such that $\mu_R(x, y) = \mu(y, x)$.

Definition 6.12 Let L be a poset. Let $x, y \in L$. Then y is called a *cover* of x if $x < y$, and there is no z such that $x < z < y$. The *Hasse diagram* of L is a directed graph that has the elements of L as vertices, and there is a directed edge from y to x if and only if y is a cover of x.

Example 6.13 Let $L = \mathbb{Z}$ be the set of integers with the usual linear order. Let $x, y \in L$ and $x \leq y$. Then $c_0(x, x) = 1$, $c_0(x, y) = 0$ if $x < y$, and $c_r(x, y) = \binom{y-x-1}{r-1}$ for all $r \geq 1$. So L infinite and locally finite. Furthermore $\mu(x, x) = 1$, $\mu(x, x + 1) = -1$ and $\mu(x, y) = 0$ if $y > x + 1$.

Definition 6.14 Let L be a poset. Let x, y in L. Then x and y have a *least upper bound* if there is a $z \in L$ such that $x \leq z$ and $y \leq z$, and if $x \leq w$ and $y \leq w$, then $z \leq w$ for all $w \in L$. If x and y have a least upper bound, then such an element is unique and it is called the *join* of x and y and denoted by $x \vee y$. Similarly the *greatest lower bound* of x and y is defined. If it exists, then it is unique and it is called the *meet* of x and y and denoted by $x \wedge y$. A poset L is called a *lattice* if $x \vee y$ and $x \wedge y$ exist for all x, y in L.

Remark 6.15 Let $(L, \leq)$ be a finite poset with maximum 1 such that $x \wedge y$ exists for all $x, y \in L$. The collection $\{z | x \leq z, y \leq z\}$ is finite and not empty, since it contains 1. Hence the meet of all the elements in this collection is well defined and is equal to $x \vee y$. Hence L is a lattice. Similarly L is a lattice if L is a finite poset with minimum 0 such that $x \vee y$ exists for all $x, y \in L$, since $x \wedge y = \bigvee \{ z \mid z \leq x, z \leq y \}$.

Example 6.16 Let L be the collection of all finite subsets of a given set $\mathcal{X}$. Let $\leq$ be defined by the inclusion, that means $x \leq y$ if and only if $x \subseteq y$. Then $0_L = \emptyset$, and L has a maximum if and only if $\mathcal{X}$ is finite in which case $1_L = \mathcal{X}$. Let $x, y \in L$ and $x \leq y$. Then $|x| \leq |y| < \infty$. Let $m = |y| - |x|$. Then

$$c_r(x,y) = \sum_{m_1 < m_2 < \cdots < m_{r-1} < m} \binom{m_2}{m_1} \cdots \binom{m}{m_{r-1}}.$$

Hence L is locally finite. L is finite if and only if $\mathcal{X}$ is finite. Furthermore $x \vee y = x \cup y$ and $x \wedge y = x \cap y$. So L is a lattice. Using Remark 6.6 we see that $\mu(x,y) = (-1)^{|y|-|x|}$ if $x \leq y$. This is much easier than computing $\mu(x,y)$ by means of Definition 6.4.

Example 6.17 Now suppose that $\mathcal{X} = \{1, \ldots, n\}$. Let L be the poset of subsets of $\mathcal{X}$. Let $A_1, \ldots, A_n$ be a collection of subsets of a finite set A. Define for a subset $x \in L$

$$A_x = \bigcap_{j \in x} A_j \quad \text{and} \quad f(x) = |A_x \setminus \left(\bigcup_{y<x} A_y \right)|.$$

Then A_x is the disjoint union of the subsets $A_y \setminus (\bigcup_{z<y} A_z)$ for all $y \leq x$. Hence the sum function

$$\tilde{f}(x) = \sum_{y \leq x} f(y) = \sum_{y \leq x} |A_y \setminus \left(\bigcup_{z<y} A_z \right)| = |A_x|.$$

Möbius inversion gives that

$$|A_x \setminus \left(\bigcup_{y<x} A_y \right)| = \sum_{y \leq x} (-1)^{|x|-|y|} |A_y|$$

which is called the *principle of inclusion/exclusion*.

Example 6.18 A variant of the principle of inclusion/exclusion is given as follows. Let $A_1, \ldots, A_n$ be a collection of subsets of a finite set A. Let L be the poset of all intersections of the A_j with the inclusion as partial order. Let $x \in L$. Define

$$f(x) = |x \setminus \left(\bigcup_{y<x} y \right)|.$$

Then

$$\tilde{f}(x) = \sum_{y \leq x} f(y) = \sum_{y \leq x} |y \setminus \left(\bigcup_{z<y} z \right)| = |x|.$$

Hence

$$|x \setminus \left(\bigcup_{y<x} y \right)| = \sum_{y \leq x} \mu(y) |y|.$$

Example 6.19 Let $L = \mathbb{N}$ be the set of positive integers with the divisibility relation as partial order. Then $0_L = 1$ is the minimum of L, it is locally finite and it has no

maximum. Now $m \vee n = \operatorname{lcm}(m,n)$ and $m \wedge n = \gcd(m,n)$. Hence L is a lattice. By Remark 6.6 we see that

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^r & \text{if } n \text{ is the product of } r \text{ mutually distinct primes,} \\ 0 & \text{if } n \text{ is divisible by the square of a prime.} \end{cases}$$

Hence $\mu(n)$ is the classical Möbius function. Furthermore $\mu(d,n) = \mu(\frac{n}{d})$ if $d|n$.
Let $\varphi(n) = |\{i \in \mathbb{N} | \gcd(i,n) = 1\}|$ be *Euler's φ function*. Let $[n] = \{1, \ldots, n\}$. Then $[n]$ is the disjoint union of the subsets $V_d = \{i \in [n] | \gcd(i,n) = n/d\}$ for all $d|n$, and $\{i \in [d] | \gcd(i,d) = 1\} \cdot \frac{n}{d} = V_d$, so $|V_d| = \varphi(d)$. Hence the sum function of $\varphi(n)$ is given by

$$\tilde{\varphi}(n) = \sum_{d|n} \varphi(d) = n.$$

Therefore

$$\varphi(n) = \sum_{d|n} \mu(d) \frac{n}{d},$$

by Möbius inversion.

Definition 6.20 Let $(L_1, \leq_1)$ and $(L_2, \leq_2)$ be posets. A map $\varphi : L_1 \to L_2$ is called *monotone* if $\varphi(x) \leq_2 \varphi(y)$ for all $x \leq_1 y$ in L_1. The map φ is called *strictly monotone* if $\varphi(x) <_2 \varphi(y)$ for all $x <_1 y$ in L_1. The map is called an *isomorphism of posets* if it is strictly monotone and there exists a strictly monotone map $\psi : L_2 \to L_1$ that is the inverse of φ. The posets are called *isomorphic* if there is an isomorphism of posets between them.

Remark 6.21 If $(L_1, \leq_1)$ and $(L_2, \leq_2)$ are isomorphic posets and L_1 is a lattice, then L_2 is also a lattice.

Example 6.22 Let n be a positive integer that is the product of r mutually distinct primes $p_1, \ldots, p_r$. Let L_1 be the set of all positive integers that divide n with divisibility as partial order $\leq_1$ as in Example 6.19. Let L_2 be the collection of all subsets of $\{1, \ldots, r\}$ with the inclusion as partial order $\leq_2$ as in Example 6.16. Define the maps $\varphi : L_1 \to L_2$ and $\psi : L_2 \to L_1$ by $\varphi(d) = \{i | p_i \text{ divides } n\}$ and $\psi(x) = \prod_{i \in x} p_i$. Then φ and ψ are strictly monotone and they are inverses of each other. Hence L_1 and L_2 are isomorphic lattices.

Remark 6.23 Let $(L, \leq)$ be a lattice without infinite chains. Then L has a minimum and a maximum.

Definition 6.24 Let L be a lattice with minimum 0. An *atom* is an element $a \in L$ that is a cover of 0. A lattice is called *atomic* if every $x > 0$ in L there exist atoms $a_1, \ldots, a_r$ such that $x = a_1 \vee \cdots \vee a_r$, and the minimum possible r is called the *rank* of x and is denoted by $r_L(x)$ or $r(x)$ for short. A lattice is called *semimodular* if for all mutually distinct x, y in L, $x \vee y$ covers x and y if there exists a z such that x and y cover z. A lattice is called *modular* if $x \vee (y \wedge z) = (x \vee y) \wedge z$ for all x, y and z in L such that $x \leq z$. A lattice L is called a *geometric lattice* if it is atomic and semimodular and has no infinite chains. If L is a geometric lattice L, then it has a minimum and a maximum and $r(1)$ is called the *rank* of L and is denoted by $r(L)$.

Example 6.25 Let L be the collection of all subsets of a given finite set $\mathcal{X}$ as in Example 6.16. The atoms are the singleton sets, that is subsets consisting of exactly one element of $\mathcal{X}$. Every $x \in L$ is the finite union of its singleton subsets. So L is atomic and $r(x) = |x|$. Now y covers x if and only if there is an element Q not in x such that $y = x \cup \{Q\}$. If $x \neq y$ and x and y both cover z, then there is an element P not in z such that $x = z \cup \{P\}$, and there is an element Q not in z such that $y = z \cup \{Q\}$. Now $P \neq Q$, since $x \neq y$. Hence $x \vee y = z \cup \{P, Q\}$ covers x and y. Hence L is semimodular. In fact L is modular. L is locally finite. L is a geometric lattice if and only if $\mathcal{X}$ is finite.

Example 6.26 Let L be the set of positive integers with the divisibility relation as in Example 6.19. The atoms of L are the primes. But L is not atomic, since a square is not the join of finitely many elements. L is semimodular. The interval $[1, n]$ in L is a geometric lattice if and only if n is square free. If n is square free and $m \leq n$, then $r(m) = r$ if and only if m is the product of r mutually distinct primes.

Proposition 6.27 *Let L be a geometric lattice. Then for all $x, y \in L$:*

(GL.1) If $x < y$, then $r(x) < r(y)$. (strictly monotone)

(GL.2) $r(x \vee y) + r(x \wedge y) \leq r(x) + r(y)$. (semimodular inequality)

(GL.3) All maximal chains from 0 to x have the same length $r(x)$.

Proof. This is left as an exercise. ◇

Remark 6.28 Let L be an atomic lattice. Then L is semimodular if and only if the semimodular inequality (GL.2) holds for all $x, y \in L$. And L is modular if and only if the *modular equality*: $r(x \vee y) + r(x \wedge y) = r(x) + r(y)$ holds for all $x, y \in L$.

Remark 6.29 Let L be a geometric lattice. Let $x, y \in L$ and $x \leq y$. Then every chain from x to y can be completed by a maximal chain with the same end points, and all such maximal chains have the same length $r(y) - r(x)$. This is called the *Jordan-Hölder property*.

Remark 6.30 Let L be a geometric lattice. Let $L_j = \{x \in L | r(x) = j\}$. Then L_j is called the *level* of L. Then the *Hasse diagram* of L is a graph that has the elements of L as vertices. If $x, y \in L$, $x < y$ and $r(y) = r(x) + 1$, then x and y are connected by an edge. So only elements between two consecutive levels L_j and L_{j+1} are connected by an edge. The Hasse diagram of L considered as a poset as in Definition 6.12 is the directed graph with an arrow from y to x if $x, y \in L$, $x < y$ and $r(y) = r(x) + 1$.

Remark 6.31 Let L be a geometric lattice. Then L_x is a geometric lattice of rank $r_L(1) - r_L(x)$, and $\mu_{L_x}(y) = \mu(x, y)$ and $r_{L_x}(y) = r_L(y) - r_L(x)$ for all $x \in L$ and $y \in L_x$. Similar remarks hold for L^x and $[x, y]$.

Example 6.32 Let L be the collection of all linear subspaces of a given finite dimensional vector space V over a field $\mathbb{F}$ with the inclusion as partial order. Then $0_L = \{0\}$ is the minimum and $1_L = V$ is the maximum of L. The partial order L is locally finite if and only if L is finite if and only if the field $\mathbb{F}$ is finite. Now $x \vee y = x + y$ and $x \wedge y = x \cap y$. So L is a lattice. The atoms are the one dimensional linear subspaces. Let x be a subspace of dimension r over $\mathbb{F}$. So x is generated by a basis $\mathbf{g}_1, \ldots, \mathbf{g}_r$. Let a_i be the one dimensional subspace generated by $\mathbf{g}_i$. Then $x = a_1 \vee \cdots \vee a_r$. Hence L is atomic and $r(x) = \dim(x)$. Moreover L is modular, since $\dim(x \cap y) + \dim(x + y) = \dim(x) + \dim(y)$ for all $x, y \in L$. Furthermore L has no infinite chains, since V is finite dimensional. Therefore L is a modular geometric lattice. $\mathbb{F}$ is finite if and only if L is finite if and only if L is locally finite.

Example 6.33 Let $\mathbb{F}$ be a field. Let $\mathcal{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ be an n-tuple of nonzero vectors in $\mathbb{F}^k$. Let $L = L(\mathcal{V})$ be the collection of all linear subspaces of $\mathbb{F}^k$ that are generated by subsets of $\mathcal{V}$ with inclusion as partial order. So L is finite and a fortiori locally finite. By definition $\{0\}$ is the linear subspace space generated by the empty set. Then $0_L = \{0\}$ and 1_L is the subspace generated by all $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Furthermore L is a lattice with $x \vee y = x + y$ and

$$x \wedge y = \bigvee \{\ z \mid z \leq x, z \leq y\}$$

by Remark 6.15. Let a_j be the linear subspace generated by $\mathbf{v}_j$. Then $a_1, \ldots, a_n$ are the atoms of L. Let x be the subspace generated by $\{\mathbf{v}_j | j \in J\}$. Then $x = \bigvee_{j \in J} a_j$. If x has dimension r, then there exists a subset I of J such that $|I| = r$ and $x = \bigvee_{i \in I} a_i$. Hence L is atomic and $r(x) = \dim(x)$. Now $x \wedge y \subseteq x \cap y$, so

$$r(x \vee y) + r(x \wedge y) \leq \dim(x + y) + \dim(x \cap y) = r(x) + r(y).$$

Hence the semimodular inequality holds and L is a geometric lattice. In most cases L is not modular.

Example 6.34 Let $\mathbb{F}$ be a field. Let $\mathcal{A} = (H_1, \ldots, H_n)$ be an arrangement over $\mathbb{F}$ of hyperplanes in the vector space $V = \mathbb{F}^k$. Let $L = L(\mathcal{A})$ be the collection of all nonempty intersections of elements of $\mathcal{A}$. By definition $\mathbb{F}^k$ is the empty intersection. Define the partial order $\leq$ by

$$x \leq y \ \text{ if and only if } \ y \subseteq x.$$

Then V is the minimum denoted by 0, and $\{0\}$ is the maximum denoted by 1, and

$$x \vee y = x \cap y \ \text{ if } x \cap y \neq \emptyset, \ \text{ and } \ x \wedge y = \bigcap \{\ z \mid x \cup y \subseteq z\ \}.$$

Suppose that $\mathcal{A}$ is a central arrangement. Then $x \cap y$ is nonempty for all x, y in L. So $x \vee y$ and $x \wedge y$ exist for all x, y in L, and L is a lattice. Let $\mathbf{v}_j = (v_{1j}, \ldots, v_{kj})$ be a nonzero vector such that $\sum_{i=1}^{k} v_{ij} X_i = 0$ is a homogeneous equation of H_j. Let $\mathcal{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. Consider the map $\varphi : L(\mathcal{V}) \to L(\mathcal{A})$ defined by

$$\varphi(x) = \bigcap_{j \in J} H_j \text{ if } x \text{ is the subspace generated by } \{\mathbf{v}_j | j \in J\}.$$

Now $x \subset y$ if and only if $\varphi(y) \subset \varphi(x)$ for all $x, y \in L(\mathcal{V})$. So φ is a strictly monotone map. Furthermore φ is a bijection and its inverse map is also strictly monotone. Hence $L(\mathcal{V})$ and $L(\mathcal{A})$ are isomorphic lattices. Therefore $L(\mathcal{A})$ is also a geometric lattice.

6.2 The characteristic polynomial of a geometric lattice

Definition 6.35 Let L be a finite geometric lattice. The *characteristic polynomial* $\chi_L(T)$ and the *Poincaré polynomial* π_L of L are defined by:

$$\chi_L(T) = \sum_{x \in L} \mu_L(x) T^{r(L) - r(x)}, \ \text{ and } \ \pi_L(T) = \sum_{x \in L} \mu_L(x) (-T)^{r(x)}.$$

Remark 6.36 So $\mu(L) = \chi_L(0)$, and $\chi_L(1) = 0$ if and only if L consists of one element $0 = 1$. Furthermore $\chi_L(T) = T^{r(L)} \pi_L(-T^{-1})$.

Remark 6.37 The *Whitney polynomial* ω_L in the two variables S and T is defined by

$$\omega_L(S,T) = \sum_{x \leq y} \mu(x,y) S^{r(x)} T^{r(L)-r(y)}.$$

It is also called the *Möbius polynomial* by [48, Section 1] and [49, Section 2]. Remember that $L_x = \{y \in L | x \leq y\}$ from Definition 6.1. The following relation holds for the Whitney polynomial in terms of characteristic polynomials

$$\omega_L(S,T) = \sum_{x \in L} S^{r(x)} \chi_{L_x}(T),$$

by Remark 6.31. Hence $\omega_L(0,T) = \chi_L(T)$. Define

$$\chi_{L,i}(T) = \sum_{x \in L_i} \chi_{L_x}(T).$$

Abbreviate $\chi_{L,i}(T)$ by $\chi_i(T)$. Then $\omega_L(S,T) = \sum_{i=0}^{r(L)} S^i \chi_i(T)$.

Example 6.38 Let L be the lattice of all subsets of a given finite set of r elements as in Example 6.16. Then $r(x) = |x|$ and $\mu(x,y) = (-1)^{|y|-|x|}$ if $x \leq y$. Hence

$$\chi_L(T) = \sum_{j=0}^{r} \binom{r}{j} (-1)^j T^{r-j} = (T-1)^r \quad \text{and } \chi_i(T) = \binom{r}{i} (T-1)^{r-i}.$$

Therefore $\omega_L(S,T) = (S+T-1)^r$.

Example 6.39 Let L be the lattice of all linear subspaces of a given vector space of dimension r over the finite field $\mathbb{F}_q$ as in Example 6.32. Then $r(x)$ is the dimension of x over $\mathbb{F}_q$. The number of subspaces of dimension i is counted in Remark 5.1. It is left as an exercise to show that $\mu(x,y) = (-1)^i q^{(j-i)(j-i-1)/2}$ if $r(x) = i$, $r(y) = j$ and $x \leq y$, and

$$\chi_L(T) = \sum_{i=0}^{r} \begin{bmatrix} r \\ i \end{bmatrix}_q (-1)^i q^{\binom{i}{2}} T^{r-i} = (T-1)(T-q)\cdots(T-q^{r-1}) \quad \text{and}$$

$$\chi_i(T) = \begin{bmatrix} r \\ i \end{bmatrix}_q (T-1)(T-q)\cdots(T-q^{r-i-1}).$$

See [22].

Remark 6.40 Every polynomial in one variable with coefficients in a field $\mathbb{F}$ factorizes in linear factors over the algebraic closure $\mathbb{A}$. In Examples 6.38 and 6.39 we see that $\chi_L(T)$ factorizes in linear factors over $\mathbb{Z}$. This is always the case for so called *super solvable* geometric lattices and lattices from *free* central arrangements. See [28].

6.3 The characteristic polynomial of an arrangement

An arrangement $\mathcal{A}$ gives rise to a geometric lattice $L(\mathcal{A})$ and characteristic polynomial $\chi_{L(\mathcal{A})}$, that will be denoted by $\chi_{\mathcal{A}}$. Similarly $\pi_{\mathcal{A}}$ denotes the Poincaré polynomial of $\mathcal{A}$. If $\mathcal{A}$ is an arrangement over the real numbers, then $\pi_{\mathcal{A}}(1)$ counts the number of connected components of the complement of the arrangement. See [48].

Proposition 6.41 *Let q be a prime power, and let $\mathcal{A} = (H_1, \ldots, H_n)$ be an arrangement in $\mathbb{F}_q^k$. Then*

$$\chi_{\mathcal{A}}(q^m) = |\mathbb{F}_{q^m}^k \setminus (H_1 \cup \cdots \cup H_n)|.$$

Proof. See [13, Sect. 16] [28, Theorem 2.69], [1, Theorem 2.2], [5, Proposition 3.2]. Let $A = \mathbb{F}_{q^m}^k$ and $A_j = H_j(\mathbb{F}_q^m)$. Let L be the poset of all intersections of the A_j. The principle of inclusion/exclusion as formulated in Example 6.18 gives that

$$|\mathbb{F}_{q^m}^k \setminus (H_1 \cup \cdots \cup H_n)| = \sum_{x \in L} \mu(x)|x| = \sum_{x \in L} \mu(x) q^{m \dim(x)}.$$

The expression on the right hand side is equal to $\chi_{\mathcal{A}}(q^m)$, since L is isomorphic with the reverse of the geometric lattice $L(\mathcal{A})$ of the arrangement $\mathcal{A} = (H_1, \ldots, H_n)$, so $\dim(x) = \mu_{L(\mathcal{A})} - \mu_{L(\mathcal{A})}(x)$ and $\mu_L(x) = \mu_{L(\mathcal{A})}(x)$ by Remark 6.11. $\diamond$

A nondegenerate code C over $\mathbb{F}_q$ in $\mathbb{F}_q^n$ with generator matrix G gives rise to the arrangement $\mathcal{A}_G$, and this to the characteristic polynomial $\chi_{\mathcal{A}_G}$, that does not depend on the chosen generator matrix G for C. So $\chi_{\mathcal{A}_G}$ will be denoted by χ_C.

Proposition 6.42 *Let C be a nondegenerate $\mathbb{F}_q$-linear code. Then*

$$A_n(T) = \chi_C(T).$$

Proof. The elements in $\mathbb{F}_{q^m}^k \setminus (H_1 \cup \cdots \cup H_n)$ correspond one-to-one to codewords of weight n in $C \otimes \mathbb{F}_{q^m}$ by Proposition 4.2. So $A_n(q^m) = \chi_C(q^m)$ for all positive integers m by Proposition 6.41. Now $A_n(T)$ and $\chi_C(T)$ are both polynomials that have the same value at q^m for all positive integers m. Hence $A_n(T) = \chi_C(T)$. $\diamond$

Definition 6.43 Let $\mathcal{A} = (H_1, \ldots, H_n)$ be an arrangement in $\mathbb{F}^k$ over the field $\mathbb{F}$. Let $H = H_i$. Then the *deletion* $\mathcal{A} \setminus H$ is the arrangement in $\mathbb{F}^k$ obtained from $(H_1, \ldots, H_n)$ by deleting all the H_j such that $H_j = H$. Let $x = \cap_{i \in I} H_i$ be an intersection of hyperplanes of $\mathcal{A}$. Let l be the dimension of x. The *restriction* $\mathcal{A}_x$ is the arrangement in $\mathbb{F}^l$ of all hyperplanes $x \cap H_j$ in x such that $x \cap H_j \neq \emptyset$ and $x \cap H_j \neq x$, for a chosen isomorphism of x with $\mathbb{F}^l$.

Proposition 6.44 *Let $\mathcal{A} = (H_1, \ldots, H_n)$ be an arrangement in $\mathbb{F}^k$ over the field $\mathbb{F}$. Let $H = H_i$. Then the following* deletion-restriction formula *holds:*

$$\chi_{\mathcal{A}}(T) = \chi_{\mathcal{A} \setminus H}(T) - \chi_{\mathcal{A}_H}(T).$$

Proof. A proof for an arbitrary field can be found in [28, Theorem 2.56]. Here the special case of a central arrangement over the finite field $\mathbb{F}_q$ will be treated. Without loss of generality we may assume that $H = H_1$ and $\mathcal{A}$ is simple. Denote $H_j(\mathbb{F}_{q^m})$ by H_j and $\mathbb{F}_{q^m}^k$ by V. Then

$$V \setminus (H_2 \cup \cdots \cup H_n) = (V \setminus (H_1 \cup H_2 \cup \cdots \cup H_n)) \cup (H_1 \setminus (H_2 \cup \cdots \cup H_n)).$$

The number of elements of the left hand side is equal to $\chi_{\mathcal{A}\setminus H}(q^m)$, and the number of elements of the two sets on the right hand side are equal to $\chi_{\mathcal{A}}(q^m)$ and $\chi_{\mathcal{A}_H}(q^m)$, respectively by Proposition 6.41. Hence

$$\chi_{\mathcal{A}\setminus H}(q^m) = \chi_{\mathcal{A}}(q^m) + \chi_{\mathcal{A}_H}(q^m)$$

for all positive integers m, since the union is disjoint. Therefore the identity of the polynomial holds. ◇

Definition 6.45 Let $\mathcal{A} = (H_1, \ldots, H_n)$ be an essential arrangement over $\mathbb{F}_q$ in $\mathbb{F}_q^k$. Consider the *stratification* of the affine space $\mathbb{A}^k$ of dimension k by:

$$\mathcal{Y}_k \subset \mathcal{Y}_1 \subset \cdots \subset \mathcal{Y}_1 \subset \mathcal{Y}_0,$$

where $\mathcal{Y}_0 = \mathbb{A}^k$, $\mathcal{Y}_1 = \cup_{j=1}^n H_j$ and $\mathcal{Y}_k = \{0\}$, and more generally

$$\mathcal{Y}_t = \bigcup_{r(\cap_{i=1}^t H_{j_i})=t} H_{j_1} \cap \cdots \cap H_{j_t}.$$

Then $\mathcal{Y}_t$ is a union of linear subspaces of $\mathbb{A}^k$ of dimension $k-t$. Define $\mathcal{X}_i = (\mathcal{Y}_i \setminus \mathcal{Y}_{i+1})$ for all $0 \leq i < k$.

Proposition 6.46 *Let $\mathcal{A}$ be an essential arrangement. Let $L = L(\mathcal{A})$ be the geometric lattice of $\mathcal{A}$. Then*

$$\chi_i(q^m) = |\mathcal{X}_i(\mathbb{F}_{q^m})|.$$

Proof. Remember that $\chi_{L,i}(T) = \sum_{r(x)=i} \chi_{L_x}(T)$ as defined in Remark 6.37. Let $L = L(\mathcal{A})$ and $x \in L$. Then $L(\mathcal{A}_x) = L_x$. Let $\cup\mathcal{A}_x$ be the union of the hyperplanes of $\mathcal{A}_x$. Then $|(x \setminus (\cup\mathcal{A}_x))(\mathbb{F}_{q^m})| = \chi_{L_x}(q^m)$ by Proposition 6.41. Now $\mathcal{X}_i$ is the disjoint union of complements of the arrangements $\mathcal{A}_x$ for all $x \in L$ such that $r(x) = i$. Hence

$$|\mathcal{X}_i(\mathbb{F}_{q^m})| = \sum_{x\in L, r(x)=i} |(x \setminus (\cup\mathcal{A}_x))(\mathbb{F}_{q^m})| = \sum_{x\in L, r(x)=i} \chi_{L_x}(q^m).$$

◇

6.4 Arrangement of lines in the projective plane

Let C be a nondegenerate code of length n and dimension 3 over $\mathbb{F}_q$ with generator matrix G. Let $A_w(q^m)$ be the number of codewords in $C \otimes \mathbb{F}_{q^m}$ of weight w. Then $A_w(q) = A_w$. The arrangement $\mathcal{A}_G = (H_1, \ldots, H_n)$ of planes in $\mathbb{F}_q^3$ is essential, and the corresponding arrangement of lines in $\mathbb{P}^2(\mathbb{F}_q)$ is also denoted by $\mathcal{A}_G$. Define

$$M_i(\mathbb{F}_{q^m}) = \{P \in \mathbb{P}^2(\mathbb{F}_{q^m}) \mid P \text{ is in exactly } i \text{ lines of } \mathcal{A}_G\}$$

and $\mu_i(q^m) = |M_i(\mathbb{F}_{q^m})|$.

Proposition 6.47 *If $0 < w \leq n$, then $A_w(q^m) = (q^m - 1)\mu_{n-w}(q^m)$.*

Proof. Every $P \in \mathbb{P}^2(\mathbb{F}_{q^m})$ corresponds one-to-one to $q^m - 1$ codewords of $C \otimes \mathbb{F}_{q^m}$. If P is in exactly i lines of $\mathcal{A}_G$, then the codewords that correspond to P have weight $n - i$ by Proposition 4.2. ◇

Remark 6.48 Notice that $M_i(\mathbb{F}_{q^m}) = M_i(\mathbb{F}_q)$ for all nonnegative integers m and $i \geq 2$ if the code is projective. Abbreviate $\mu_i(q^m) = \mu_i$ for $i \geq 2$ in case the code is projective.

Consider the following *stratification* of the affine space $\mathbb{A}^3$ by:

$$\mathcal{Y}_3 \subset \mathcal{Y}_2 \subset \mathcal{Y}_1 \subset \mathcal{Y}_0,$$

where $\mathcal{Y}_0 = \mathbb{A}^3$, $\mathcal{Y}_1 = H_1 \cup \cdots \cup H_n$,

$$\mathcal{Y}_2 = \bigcup \{\ H_i \cap H_j \mid 1 \leq i < j \leq n, H_i \neq H_j\ \},$$

and $\mathcal{Y}_3 = \{0\}$. Now $\mathcal{X}_i = (\mathcal{Y}_i \setminus \mathcal{Y}_{i+1})$ for $0 \leq i < 3$.

Definition 6.49 Let $\mathcal{X}(\mathbb{F})$ denote the set of $\mathbb{F}$-*rational points* of $\mathcal{X}$, that is the set of points of $\mathcal{X}$ that have coordinates in $\mathbb{F}$.

Proposition 6.50 *If the code is projective, then*

$$A_n(q^m) = |\mathcal{X}_0(\mathbb{F}_{q^m})| \quad \textit{and} \quad A_{n-1}(q^m) = |\mathcal{X}_1(\mathbb{F}_{q^m})|.$$

Proof. The H_j are mutually distinct, since the code is projective. Proposition 6.47 implies that $A_n(q^m) = (q^m - 1)\mu_0(q^m)$ and $A_{n-1}(q^m) = (q^m - 1)\mu_1(q^m)$. So codewords of weight n over $\mathbb{F}_{q^m}$ correspond one to one to points in $\mathbb{F}_{q^m}^3$ that are not in any of the H_j, that is in $\mathcal{X}_0(\mathbb{F}_{q^m})$. And codewords of weight $n-1$ over $\mathbb{F}_{q^m}$ correspond one to one to $\mathbb{F}_{q^m}$-rational points that are on exactly one H_j, that is in $\mathcal{X}_1(\mathbb{F}_{q^m})$. $\diamond$

Proposition 6.51 *Let C be a projective code of length n and dimension 3 over $\mathbb{F}_q$. Then*

$$\begin{cases} |\mathcal{X}_0(\mathbb{F}_{q^m})| &= (q^m-1)\left(q^{2m} - (n-1)q^m + \sum_{i\geq 2}(i-1)\mu_i - n + 1\right), \\ |\mathcal{X}_1(\mathbb{F}_{q^m})| &= (q^m-1)\left(nq^m + n - \sum_{i\geq 2} i\mu_i\right), \\ |\mathcal{X}_2(\mathbb{F}_{q^m})| &= (q^m-1)\left(\sum_{i\geq 2}\mu_i\right). \end{cases}$$

Proof. Let $\bar{P}$ be the corresponding point in $\mathbb{P}^2(\mathbb{F}_{q^m})$ for $P \in \mathbb{F}_{q^m}^3$ and $P \neq 0$. Abbreviate $\mathcal{X}_i(\mathbb{F}_{q^m})$ by $\mathcal{X}_i$. Define $\bar{\mathcal{X}}_i = \{\bar{P} \mid P \in \mathcal{X}_i\}$. Then $|\mathcal{X}_i| = (q^m - 1)|\bar{\mathcal{X}}_i|$ for all $i < 3$.
(1) If $\bar{P} \in \bar{\mathcal{X}}_2$, then $\bar{P} \in H_j \cap H_k$ for some $j \neq k$. Hence $\bar{P} \in M_i(\mathbb{F}_q)$ for some $i \geq 2$, since the code is projective. Therefore $\bar{\mathcal{X}}_2$ is the disjoint union of the $M_i(\mathbb{F}_q)$, $i \geq 2$, and $|\bar{\mathcal{X}}_2| = \sum_{i\geq 2}\mu_i$.
(2) $\bar{P} \in \bar{\mathcal{X}}_1$ if and only if P is on exactly one line H_j. There are n lines, and every line has $q^m + 1$ points that are defined over $\mathbb{F}_{q^m}$. If $i \geq 2$, then every $\bar{P} \in M_i(\mathbb{F}_q)$ is on i lines H_j. Hence $|\bar{\mathcal{X}}_1| = n(q^m + 1) - \sum_{i\geq 2} i\mu_i$.
(3) $\mathbb{P}^2$ is the disjoint union of $\bar{\mathcal{X}}_1$, $\bar{\mathcal{X}}_2$ and $\bar{\mathcal{X}}_0$. The numbers $|\bar{\mathcal{X}}_2|$ and $|\bar{\mathcal{X}}_1|$ are computed in (1) and (2), and $|\mathbb{P}^2(\mathbb{F}_{q^m})| = q^{2m} + q^m + 1$. From this we derive the number of elements of $\bar{\mathcal{X}}_0$. $\diamond$

Remark 6.52 The polynomials $\chi_i(T)$ are given by

$$\begin{cases} \chi_0(T) &= (T-1)\left(T^2 - (n-1)T + \sum_{i\geq 2}(i-1)\mu_i - n + 1\right), \\ \chi_1(T) &= (T-1)\left(nT + n - \sum_{i\geq 2} i\mu_i\right), \\ \chi_2(T) &= (T-1)\left(\sum_{i\geq 2}\mu_i\right), \end{cases}$$

since $\chi_i(q^m)$ is the number of elements of $\mathcal{X}_i(\mathbb{F}_{q^m})$ by Propositions 6.46 and 6.51.

Theorem 6.53 *Let C be a projective code of length n and dimension 3 over $\mathbb{F}_q$. Then $A_0(T) = 1$ and $A_w(T) = (T-1)\mu_{n-w}$ for $0 < w < n-1$, and*

$$A_{n-1}(T) = (T-1)\left(nT + n - \sum_{i\geq 2} i\mu_i\right) \quad \textit{and}$$

$$A_n(T) = (T-1)\left(T^2 - (n-1)T + \sum_{i\geq 2}(i-1)\mu_i - n + 1\right).$$

Proof. This is a consequence of Propositions 6.47, 6.50 and 6.51. ◇

Remark 6.54 The polynomials $A_i(T)$ and $\chi_i(T)$ are divisible by $T-1$. Hence there are polynomials $\bar{A}_i(T)$ and $\bar{\chi}_i(T)$ such that $A_i(T) = (T-1)\bar{A}_i(T)$ for all $i > 0$, and $\chi_i(T) = (T-1)\bar{\chi}_i(T)$ for all $i < 3$.

Example 6.55 Consider the matrices G and P given by

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

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

Let C be the code over $\mathbb{F}_q$ with generator matrix G. The columns of G represent also the coefficients of the lines of $\mathcal{A}_G$. The j-th column of P represents the homogenous coordinates of the points P_j in the projective plane that occur as intersections of two lines of $\mathcal{A}_G$. In case q is even, the points P_7, P_8 and P_9 coincide.
If q is even, then $\mu_3 = 7$. If q is odd, then $\mu_2 = 3$.

	i	1	2	3	4	5	6	7
	μ_i		0	7	0	0	0	0
q even	$\bar{A}_i$	0	0	0	7	0	$7T-14$	T^2-6T+8
	$\bar{\chi}_{3-i}$	7	$7T-14$	T^2-6T+8				
	μ_i		3	6	0	0	0	0
q odd	$\bar{A}_i$	0	0	0	6	3	$7T-17$	T^2-6T+9
	$\bar{\chi}_{3-i}$	9	$7T-17$	T^2-6T+9				

Notice that there is a codeword of weight 7 in case q is even and $q > 4$ or q is odd and $q > 3$, since $\bar{A}_7(T) = (T-2)(T-4)$ or $\bar{A}_7(T) = (T-3)^2$, respectively.

Example 6.56 Let G be a $3 \times n$ generator matrix of an MDS code. The lines of the arrangement $\mathcal{A}_G$ are in general position. That means that every two distinct lines meet in one point, and every three mutually distinct lines have an empty intersection. So $\mu_2 = \binom{n}{2}$ and $\mu_i = 0$ for all $i > 2$. Hence $\bar{A}_{n-2}(T) = \bar{\chi}_2(T) = \binom{n}{2}$ and $\bar{A}_{n-1}(T) = \bar{\chi}_1(T) = nT + 2n - n^2$ and $\bar{A}_n(T) = \bar{\chi}_0(T) = T^2 - (n-1)T + \binom{n-1}{2}$, by Proposition 6.46 and Theorem 6.53 which is in agreement with Theorem 5.27.

Example 6.57 Let a and b positive integers such that $2 < a < b$. Let $n = a + b$. Let G be a $3 \times n$ generator matrix of a nondegenerate code. Suppose that there are two points P and Q in the projective plane over $\mathbb{F}_q$ such that the lines of the projective arrangement of $\mathcal{A}_G$ consists of a distinct lines incident with P and b distinct lines incident with Q. Then $\mu_2 = ab$, $\mu_a = 1$ and $\mu_b = 1$. Hence $\bar{\chi}_1(T) = ab + 2$ and $\bar{A}_a = \bar{A}_b = 1$ and $\bar{A}_{n-2} = ab$. Furthermore

$$\bar{A}_{n-1}(T) = \bar{\chi}_1(T) = (a+b)T - 2ab,$$

$$\bar{A}_n(T) = \bar{\chi}_0(T) = T^2 - (a+b-1)T + ab - 1$$

and $\bar{A}_i(T) = 0$ for all $i \neq a, b, n-2, n-1, n$.

Remark 6.58 In the Appendix A a table is given of the weight enumerator and characteristic polynomials $\chi_i(C)$ of all codes C of dimension 3 and length at most 6. One sees that the codes 5.1c and 5.2a have the same characteristic polynomials χ_i for all i, but their weight enumerators are distinct. But these codes are not projective.

6.5 Graphs, arrangments and codes

Definition 6.59 A *graph* Γ is a pair (V, E) where V is a non-empty set and E is a set disjoint from V. The elements of V are called *vertices*, and members of E are called *edges*. Edges are *incident* to one or two vertices, which are called the *ends* of the edge. If an edge is incident with exactly one vertex, then it is called a *loop*. If u and v are vertices that are incident with an edge, then they are called *neighbors* or *adjacent*, Two edges are called *parallel* if they are incident with the same vertices. The graph is called *simple* if it has no loops and no parallel edges. A graph $\Gamma' = (V', E')$ is called a *subgraph* of Γ if $V' \subseteq V$ and $E' \subseteq E$. Two vertices u to v are *connected* by *path* from u to v if there is a t-tuple of mutually distinct vertices $(v_1, \ldots, v_t)$ with $u = v_1$ and $v = v_t$, and an $t-1$-tuple of mutually distinct edges $(e_1, \ldots, e_{t-1})$ such that e_i is incident with v_i and v_{i+1} for all $1 \leq i < t$. If moreover e_t is an edge that is incident with u and v and distinct from e_i for all $i < t$, then $(e_1, \ldots, e_{t-1}, e_t)$ is called a *cycle*. The length of the smallest cycle is called the *girth* of the graph and is denoted by $\gamma(\Gamma)$.

Definition 6.60 The graph is called *connected* if every two vertices are connected by a path. A maximal connected subgraph of Γ is called a *connected component* of Γ. The vertex set V of Γ is a disjoint union of subsets V_i such that $\Gamma_i = (V_i, E_i)$ is a connected component of Γ. The number of connected components of Γ is denoted by $c(\Gamma)$.

Definition 6.61 Let $\Gamma = (V, E)$ be a finite graph. Suppose that V consists of m elements enumerated by $v_1, \ldots, v_m$. Suppose that E consists of n elements enumerated by $e_1, \ldots, e_n$. The *incidence matrix* $I(\Gamma)$ is an $m \times n$ matrix with entries a_{ij} defined by

$$a_{ij} = \begin{cases} 1 & \text{if } e_j \text{ is incident with } v_i \text{ and } v_k \text{ for some } i < k, \\ -1 & \text{if } e_j \text{ is incident with } v_i \text{ and } v_k \text{ for some } i > k, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose moreover that Γ is simple. Then $\mathcal{A}(\Gamma)$ is the arrangement $(H_1, \ldots, H)n)$ of hyperplanes where $H_J = X_i - X_k$ if e_j is incident with v_i and v_k for i and j such that $i < k$. An arrangement $\mathcal{A}$ is called *graphic* if $\mathcal{A}$ is isomorphic with $\mathcal{A}(\Gamma)$ for some graph Γ.

Definition 6.62 Let $\Gamma = (V, E)$ be a graph. Let K be a finite set and $k = |K|$. The elements of K are called *colors*. A *k-coloring* of Γ is an a map $\gamma : \Gamma \to K$ such that $\gamma(u) \neq \gamma(v)$ for all adjacent vertices u and v in V. So vertex u has color $\gamma(u)$ and adjacent vertices have distinct colors. Let $P_\Gamma(k)$ be the number of k-colorings of Γ. Then P_Γ is called the *chromatic polynomial* of Γ.

Remark 6.63 The number of coloring of graphs was studied by Whitney [46, 45] and Tutte [38, 39, 40, 41, 42]. That P_Γ is indeed a polynomial will be a consequence of Proposition 6.66. A tool in order to proof this is by deletion-contraction of graphs, similar to deletion-restriction of arrangements. Notice that the number of k-colorings of Γ does not change by deleting loops and a parallel edge.

Definition 6.64 Let $\Gamma = (V, E)$ be a graph. Let e be an edge that is incident to the vertices u and v. Then the *deletion* $\Gamma \setminus e$ is the graph with vertices V and edges $E \setminus \{e\}$. The *contraction* Γ/e is the graph obtained by identifying u and v. Formally this is defined as follows. Let $\bar{u} = \bar{v} = \{u, v\}$, and $\bar{w} = \{w\}$ if $w \neq u$ and $w \neq v$. Let $\bar{V} = \{\bar{w} | w \in V\}$. Then $\Gamma/e = (\bar{V}, E)$, where e is incident with $\bar{w}$ in Γ/e if $e \in E$ is incident with w in Γ.

Proposition 6.65 *Let $\Gamma = (V, E)$ be a graph. Let e be an edge of Γ. Then the following* deletion-contraction formula *holds:*

$$P_\Gamma(k) = P_{\Gamma \setminus e}(k) - P_{\Gamma/e}(k).$$

for all positive integers k.

Proof. See [28, Proposition 2.84] ◇

Proposition 6.66 *Let $\Gamma = (V, E)$ be a finite simple graph. Let χ_Γ be the characteristic polynomial of the geometric lattice $L(M(\Gamma))$. Then $P_\Gamma(k) = \chi_\Gamma(k)$ for all positive integers k.*

Proof. See [28, Theorem 2.88] ◇

Definition 6.67 The *graph code* of Γ over $\mathbb{F}_q$ is the $\mathbb{F}_q$-linear code that is generated by the rows of $I(\Gamma)$. The *cycle code* $C(\Gamma)$ of Γ is the dual of the graph code of C.

Remark 6.68 Let Γ be a finite graph without loops. Then the arrangement $\mathcal{A}(\Gamma)$ is isomorphic with $\mathcal{A}_{C(\Gamma)}$.

Proposition 6.69 *Let Γ be a finite graph. Then $C(\Gamma)$ is a code with parameters $[n, k, d]$, where $n = |E|$, $k = |E| - |V| + c(\Gamma)$ and $d = \gamma(\Gamma)$.*

Proof. This is left as exercise. ◇

Sparse graph codes, *Gallager* or *Low-density parity check* codes and *Tanner graph* codes play an important role in the research of coding theory at this moment. See [25, 32].

6.6 The zeta function of an arrangement

Definition 6.70 Let $\mathcal{X}$ be an *affine variety* in $\mathbb{A}^k$ defined over $\mathbb{F}_q$, that is the zeroset of a collection of polynomials in $\mathbb{F}_q[X_1, \ldots, X_k]$. The *zeta function* $Z_\mathcal{X}(T)$ of $\mathcal{X}$ is the formal power series in T defined by

$$Z_\mathcal{X}(T) = \exp\left(\sum_{m=1}^{\infty} \frac{|\mathcal{X}(\mathbb{F}_{q^m})|}{r} T^r\right).$$

Theorem 6.71 *Let $\mathcal{A}$ be an arrangement in $\mathbb{F}_q^k$. Let $\chi_{\mathcal{A}}(T) = \sum_{j=0}^{k} c_j T^j$ be the characteristic polynomial of $\mathcal{A}$. Let $\mathcal{M} = \mathbb{A}^k \setminus (H_1 \cup \cdots \cup H_n)$ be the complement of the arrangement. Then the zeta function of $\mathcal{M}$ is given by:*

$$Z_{\mathcal{M}}(T) = \prod_{j=0}^{k} (1 - q^j T)^{-c_j}.$$

Proof. See [5, Theorem 3.6]. ◇

6.7 Exercises

6.1 Show that a poset L is locally finite if and only if $[x, y]$ is finite for all $x \leq y$ in L.

6.2 Give a proof of the formulas for $c_r(x, y)$ and $\mu(x, y)$ in Example 6.16.

6.3 Give a proof of the formula for $\mu(x)$ in Example 6.19.

6.4 Give a proof of the statements in Example 6.26.

6.5 Give an example of an atomic finite lattice with minimum 0 and maximum 1 that is not semimodular.

6.6 Give a proof of the statements in Remark 6.28.

6.7 Let L be a geometric lattice. Show that the rank $r(x)$ is the length of a maximal chain from 0 to x for all x in L.

6.8 Let L be a geometric lattice. Let a be an atom of L and $x \in L$. Show that $r(x \vee a) \leq r(x) + 1$ and $r(x \vee a) = r(x)$ if and only if $a \leq x$.

6.9 Give a proof of Remark 6.31.

6.10 Give an example of a central arrangement $\mathcal{A}$ such that the lattice $L(\mathcal{A})$ is not modular.

6.11 Give a proof of the formulas for $\mu(x, y)$, $\chi_L(T)$ and $\omega_L(S, T)$ in Example 6.39.

6.12 Show that the Whitney polynomial determined by the extended weight enumerator of a code? Give two projective codes over a fixed finite field of the same length and dimension and with the same characteristic polynomials χ_i for all i, but with distinct extended weight enumerators.

6.13 Give a proof of Remark 6.48.

6.14 Give a proof of Proposition 6.69.

7 Matroids and the Tutte polynomial

The notion of a matroid is almost equivalent to a geometric lattice. See [10, 12, 29, 35]. Matroids were introduced by Whitney [46, 47] in axiomatizing the concept of independence. In the theory of arrangements one uses the notion of a geometric lattice. In graph and coding theory one refers more to matroids.

7.1 Matroids

Definition 7.1 A *matroid* is a pair $(M, \mathcal{I})$ consisting of a finite set M and a collection $\mathcal{I}$ of subsets of M such that the following three conditions hold.

(M.0) $\emptyset \in \mathcal{I}$.

(M.1) If $J \subseteq I$ and $I \in \mathcal{I}$, then $J \in \mathcal{I}$.

(M.2) If $I, J \in \mathcal{I}$ and $|I| < |J|$, then there exists an $j \in (J \setminus I)$ such that $I \cup \{j\}$ in $\mathcal{I}$.

A subset I of M is called *independent* if $I \in \mathcal{I}$, otherwise it is called *dependent*. Condition (M.2) is called the *independence augmentation axiom*.

Remark 7.2 If J is a subset of M, then J has a *maximal independent subset*, that is there exists an $I \in \mathcal{I}$ such that $I \subseteq J$ and I is maximal with respect to this property and the inclusion. If I_1 and I_2 are maximal independent subsets of J, then $|I_1| = |I_2|$. The *rank* or *dimension* of a subset J of M is the number of elements of a maximal independent subset of J. An independent set of rank $r(M)$ is called a *basis*. The collection of all bases of M is denoted by $\mathcal{B}$.

Example 7.3 Let n and k be non-negative integers such that $k \leq n$. Let $U_{n,k}$ be a set consisting of n elements and $\mathcal{I}_{n,k} = \{I \subseteq U_{n,k} || I| \leq k\}$. Then $(U_{n,k}, \mathcal{I}_{n,k})$ is a matroid and called the *uniform matroid* of rank k on an n-element set. A subset B of $U_{n,k}$ is a basis if and only if $|B| = k$. The matroids $U_{n,n}$ have no dependent sets and are called *free*.

Proposition 7.4 *Let L be a finite geometric lattice. Let $M(L)$ be the set of all atoms of L. Let $\mathcal{I}(L)$ be the collection of all subsets I of $M(L)$ such that $r(a_1 \vee \cdots \vee a_r) = r$ if $I = \{a_1, \ldots, a_r\}$ is a collection of r atoms of L. Then $(M(L), \mathcal{I}(L))$ is a matroid.*

Proof. The proof is left as an exercise. ◇

Proposition 7.5 *Let L be a finite geometric lattice. Let $M(L)$ be the matroid associated with L. Then*

$$\chi_L(T) = \sum_{I \subseteq M(L)} (-1)^{|I|} T^{r(L)-r(I)}.$$

Proof. The reader is referred to the literature. ◇

Definition 7.6 Let $(M, \mathcal{I})$ be a matroid. An element x in M is called a *loop* if $\{x\}$ is a dependent set. Let x and y in M be two distinct elements that are not loops. Then x and y are called *parallel* if $r(\{x, y\}) = 1$. The matroid is called *simple* if it has no loops and no parallel elements.

Remark 7.7 Let G be a $k \times n$ matrix with entries in a field $\mathbb{F}$. Let M_G be the set $\{1, \ldots, n\}$ indexing the columns of G and $\mathcal{I}_G$ be the collection of all subsets I of M_G such that the submatrix G_I consisting of the columns of G at the positions of I are independent. Then $(M_G, \mathcal{I}_G)$ is a matroid. Suppose that $\mathbb{F}$ is a finite field and G_1 and G_2 are generator matrices of a code C, then $(M_{G_1}, \mathcal{I}_{G_1}) = (M_{G_2}, \mathcal{I}_{G_2})$. So the matroid $(M_C, \mathcal{I}_C)$ of a code C is well defined by $(M_G, \mathcal{I}_G)$ for some generator matrix G of C. If C is degenerate, then there are positions i such that $c_i = 0$ for every codeword $\mathbf{c} \in C$ and these positions correspond one-to-one with loops of M_C. Let C be nondegenerate. Then M_C has no loops, and the positions i and j with $i \neq j$ are parallel in M_C if and only if the i-th column of G is a scalar multiple of the j-th column. The code C is projective if and only if the arrangement $\mathcal{A}_G$ is simple if and only if the matroid M_C is simple. An $[n, k]$ code C is MDS if and only if the matroid M_C is uniform.

Definition 7.8 Let $(M_1, \mathcal{I}_1)$ and $(M_2, \mathcal{I}_2)$ be matroids. A map $\varphi : M_1 \to M_2$ is called a *morphism of matroids* if $\varphi(I) \in \mathcal{I}_2$ for all $I \in \mathcal{I}_1$. The map is called an *isomorphism of matroids* if it is a morphism of matroids and there exists a map $\psi : M_2 \to M_1$ such that it is a morphism of matroids and it is the inverse of φ. The matroids are called *isomorphic* if there is an isomorphism of matroids between them.

Remark 7.9 Let C be a projective code with generator matrix G. Then $\mathcal{A}_G$ is an essential simple arrangement with geometric lattice $L(\mathcal{A}_G)$. Furthermore the matroids $M(L(\mathcal{A}_G))$ and M_C are isomorphic.

Remark 7.10 A matroid M is called *realizable* over the field $\mathbb{F}$ if there exists a matrix G with entries in $\mathbb{F}$ such that M is isomorphic with M_G.

Definition 7.11 Let $(M, \mathcal{I})$ be a matroid. A *k-flat* of M is a maximal subset of M of rank k. Let $L(M)$ be the collection of all flats of M, it is called the *lattice of flats* of M. The *closure* $\bar{J}$ of a subset J of M is the intersection of all flats that contain J.

Remark 7.12 M is a k-flat with $k = r(M)$. If F_1 and F_2 are flats, then $F_1 \cap F_2$ is also a flat. Consider $L(M)$ with the inclusion as partial order. Then M is the maximum of $L(M)$. And $F_1 \cap F_2 = F_1 \wedge F_2$ for all F_1 and F_2 in $L(M)$. Hence $L(M)$ is indeed a lattice by Remark 6.15. Let J be a subset of M, then $\bar{J}$ is a flat, since it is a nonempty, finite intersection of flats. So $\bar{\emptyset}$ is the minimum of $L(M)$.

Remark 7.13 An element x in M is a loop if and only if $\bar{x} = \bar{\emptyset}$. If $x, y \in M$ are no loops, then x and y are parallel if and only if $\bar{x} = \bar{y}$. Let $\bar{M} = \{\bar{x} | x \in M, \bar{x} \neq \bar{\emptyset}\}$ and $\bar{I} = \{\bar{x} | x \in I\}$ for a subset I of M. Let $\bar{\mathcal{I}} = \{\bar{I} | I \in \mathcal{I}, \bar{\emptyset} \notin \bar{I}\}$. Then $(\bar{M}, \bar{\mathcal{I}})$ is a simple matroid.
Let G a generator matrix of a code C. The *reduced matrix* $\bar{G}$ is the matrix obtained from G by deleting all zero columns from G and all columns that are a scalar multiple of a previous column. The *reduced code* $\bar{C}$ of C is the code with generator matrix $\bar{G}$. The matroids $\bar{M}_G$ and $M_{\bar{G}}$ are isomorphic.

Proposition 7.14 *Let $(M, \mathcal{I})$ be a matroid. Then $L(M)$ with the inclusion as partial order is a geometric lattice and $L(M)$ is isomorphic with $L(\bar{M})$.*

Proof. This is left as an exercise. ◇

Definition 7.15 Let $(M, \mathcal{I})$ be a matroid. Let $\mathcal{B}$ be the collection of all bases of M. Define $B^\perp = (M \setminus B)$ for $B \in \mathcal{B}$, and $\mathcal{B}^\perp = \{B^\perp | B \in \mathcal{B}\}$. Define $M^\perp = M$ and $\mathcal{I}^\perp = \{I \subseteq M | I \subseteq B \text{ for some } B \in \mathcal{B}^\perp\}$. Then $(M^\perp, \mathcal{I}^\perp)$ is called the *dual matroid* of $(M, \mathcal{I})$.

Remark 7.16 The dual matroid is indeed a matroid. Let C be a code over a finite field. Then $(M_C)^\perp$ is isomorphic with $M_{C^\perp}$ as matroids.

Proposition 7.17 *Let $(M, \mathcal{I})$ be a matroid with rank function r. Then the dual matroid has rank function*

$$r^\perp(J) = |J| - r(M) + r(M \setminus J).$$

Proof. The proof is based on the observarion that $r(J) = \max_{B\in\mathcal{B}} |B \cap J|$ and $B \setminus J = B \cap (M \setminus J)$.

$$\begin{aligned} r^{\perp}(J) &= \max_{B\in\mathcal{B}} |(M \setminus B) \cap J| \\ &= \max_{B\in\mathcal{B}} |J \setminus B| \\ &= |J| - \min_{B\in\mathcal{B}} |J \cap B| \\ &= |J| - (|B| - \max_{B\in\mathcal{B}} |B \setminus J|) \\ &= |J| - r(M) + \max_{B\in\mathcal{B}} |B \cap (M \setminus J)| \\ &= |J| - r(M) + r(S \setminus J). \end{aligned}$$

◇

7.2 Graphs and matroids

Definition 7.18 Let $(M, \mathcal{I})$ be a matroid. A subset C of M is called a *circuit* if it is dependent and all its proper subsets are independent.

Proposition 7.19 *Let $\mathcal{C}$ be the collection of circuits of a matroid. Then*

(C.0) $\emptyset \notin \mathcal{C}$.

(C.1) If $C_1, C_2 \in \mathcal{C}$ *and* $C_1 \subseteq C_2$, *then* $C_1 = C_2$.

(C.2) If $C_1, C_2 \in \mathcal{C}$ *and* $C_1 \neq C_2$ *and* $x \in C_1 \cap C_2$, *then there exists a* $C_3 \in \mathcal{C}$ *such that* $C_3 \subseteq (C_1 \cup C_2) \setminus \{x\}$.

Proof. See [29, Lemma 1.1.3]. ◇

Condition (C.2) is called the *circuit elimination axiom*. The converse of Proposition 7.19 holds.

Proposition 7.20 *Let $\mathcal{C}$ be a collection of subsets of a finite set M that satisfies the conditions (C.0), (C.1) and (C.2). Let $\mathcal{I}$ be the collection of all subsets of M that contain no member of $\mathcal{C}$. Then $(M, \mathcal{I})$ is a matroid with $\mathcal{C}$ as its collection of circuits.*

Proof. See [29, Theorem 1.1.4]. ◇

Proposition 7.21 *Let $\Gamma = (V, E)$ be a finite graph. Let $\mathcal{C}$ the collection of all subsets $\{e_1, \ldots, e_t\}$ such that $(e_1, \ldots, e_t)$ is a cycle in Γ. Then $\mathcal{C}$ is the collection of circuits of a matroid $M(\Gamma)$ on E. This matroid is called the* cycle matroid *of Γ.*

Proof. See [29, Proposition 1.1.7]. ◇

Remark 7.22 A matroid M is called *graphic* if M is isomorphic with $M(\Gamma)$ for some graph Γ. Loops in Γ correspond one-to-one to loops in $M(\Gamma)$. Two edges that are no loops, are parallel in Γ if and only if they are parallel in $M(\Gamma)$. So Γ is simple if and only if $M(\Gamma)$ is simple.

Remark 7.23 Let Γ be a finite graph. Then $M(\Gamma)$ is isomorphic with $M_{C(\Gamma)}$.

7.3 The Tutte and Whitney polynomials of a matroid

See [1, 2, 6, 7, 8, 10, 16, 18] for references of this section.

Definition 7.24 Let $(M, \mathcal{I})$ be a matroid. Then the *Whitney rank generating function* $R_M(X, Y)$ is defined by

$$R_M(X,Y) = \sum_{I \subseteq M} X^{r(M)-r(I)} Y^{|I|-r(I)}$$

and the *Tutte polynomial* by

$$t_M(X,Y) = \sum_{I \subseteq M} (X-1)^{r(M)-r(I)} (Y-1)^{|I|-r(I)} \ .$$

In other words

$$t_M(X,Y) = R_M(X-1, Y-1).$$

Remark 7.25 Let L be a finite geometric lattice with associated matroid $L(M)$. Then

$$\chi_L(T) = \sum_{I \subseteq M(L)} (-1)^{|I|} T^{r(L)-r(I)},$$

by Proposition 7.5. Hence

$$\chi_L(T) = (-1)^{r(L)} t_{M(L)}(1-X, 0).$$

describes the characteristic polynomial of L in terms of the Tutte polynomial of $M(L)$.

7.4 Weight enumerator and Tutte polynomial

As we have seen, we can interpret a linear $[n, k]$ code C over $\mathbb{F}_q$ as a matroid via the columns of a generator matrix G. Using Lemma 5.5 we can rewrite the Tutte polynomial associated to a code:

Proposition 7.26 *Let C be a $[n, k]$ code over $\mathbb{F}_q$ with generator matrix G. Then the Tutte polynomial associated with the code C is*

$$t_G(X,Y) = \sum_{t=0}^{n} \sum_{|J|=t} (X-1)^{l(J)} (Y-1)^{l(J)-(k-t)} \ .$$

This formula and Proposition 5.20 already suggest the next connection between the weight enumerator and the Tutte polynomial.

Theorem 7.27 *Let C be a $[n, k]$ code over $\mathbb{F}_q$ with generator matrix G. Then the following holds for the Tutte polynomial and the extended weight enumerator:*

$$W_C(X,Y,T) = (X-Y)^k Y^{n-k} \ t_G\left(\frac{X+(T-1)Y}{X-Y}, \frac{X}{Y}\right) \ .$$

Proof. By using the previous proposition about the Tutte polynomial, rewriting, and Proposition 5.20 we get

$$\begin{aligned}
& (X-Y)^k Y^{n-k}\ t_G\left(\frac{X+(T-1)Y}{X-Y}, \frac{X}{Y}\right) \\
= \ & (X-Y)^k Y^{n-k} \sum_{t=0}^{n} \sum_{|J|=t} \left(\frac{TY}{X-Y}\right)^{l(J)} \left(\frac{X-Y}{Y}\right)^{l(J)-(k-t)} \\
= \ & (X-Y)^k Y^{n-k} \sum_{t=0}^{n} \sum_{|J|=t} T^{l(J)} Y^{k-t} (X-Y)^{-(k-t)} \\
= \ & \sum_{t=0}^{n} \sum_{|J|=t} T^{l(J)} (X-Y)^t Y^{n-t} \\
= \ & W_C(X,Y,T).
\end{aligned}$$

◇

We use the extended weight enumerator here, because extending a code does not change the generator matrix and therefore not the matroid G. The converse of this theorem is also true: the Tutte polynomial is completely defined by the extended weight enumerator.

Theorem 7.28 *Let C be a $[n,k]$ code over $\mathbb{F}_q$ with generator matrix G. Then the following holds for the extended weight enumerator and the Tutte polynomial:*

$$t_G(X,Y) = Y^n (Y-1)^{-k} W_C(1, Y^{-1}, (X-1)(Y-1)) \ .$$

Proof. The proof of this theorem goes analogous to the proof of the previous theorem.

$$\begin{aligned}
& Y^n (Y-1)^{-k} W_C(1, Y^{-1}, (X-1)(Y-1)) \\
= \ & Y^n (Y-1)^{-k} \sum_{t=0}^{n} \sum_{|J|=t} ((X-1)(Y-1))^{l(J)} (1-Y^{-1})^t Y^{-(n-t)} \\
= \ & \sum_{t=0}^{n} \sum_{|J|=t} (X-1)^{l(J)} (Y-1)^{l(J)} Y^{-t} (Y-1)^t Y^{-(n-k)} Y^n (Y-1)^{-k} \\
= \ & \sum_{t=0}^{n} \sum_{|J|=t} (X-1)^{l(J)} (Y-1)^{l(J)-(k-t)} \\
= \ & t_G(X,Y).
\end{aligned}$$

◇

We see that the Tutte polynomial depends on two variables, while the extended weight enumerator depends on three variables. This is no problem, because the weight enumerator is given in its homogeneous form here: we can view the extended weight enumerator as a polynomial in two variables via $W_C(Z,T) = W_C(1,Z,T)$.
Greene [16] already showed that the Tutte polynomial determines the weight enumerator, but not the other way round. By using the extended weight enumerator, we get a two-way equivalence and the proof reduces to rewriting.

We can also give expressions for the generalized weight enumerator in terms of the Tutte polynomial, and the other way round. The first formula was found by Britz [8] and independently by Jurrius [18].

Theorem 7.29 *For the generalized weight enumerator of a $[n,k]$ code C and the associated Tutte polynomial we have that*

$$W_C^r(X,Y) = \frac{1}{\langle r \rangle_q} \sum_{j=0}^{r} \begin{bmatrix} r \\ j \end{bmatrix}_q (-1)^{r-j} q^{\binom{r-j}{2}} (X-Y)^k Y^{n-k} \; t_G\left(\frac{X+(q^j-1)Y}{X-Y}, \frac{X}{Y} \right);$$

and, conversely,

$$t_G(X,Y) = Y^n (Y-1)^{-k} \sum_{r=0}^{k} \left(\prod_{j=0}^{r-1} ((X-1)(Y-1)-q^j) \right) W_C^r(1,Y^{-1}) \; .$$

Proof. For the first formula, use Theorems 5.25 and 7.27. Use Theorems 5.24 and 7.28 for the second formula. ◇

7.5 Exercises

7.1 Give a proof of the statements in Remark 7.2.

7.2 Let L be a finite geometric lattice. Show that $(M(L), \mathcal{I}(L))$ is a matroid as stated in Proposition 7.4. Show moreover that this matroid is simple.

7.3 Give a proof of the statements in Remark 7.7.

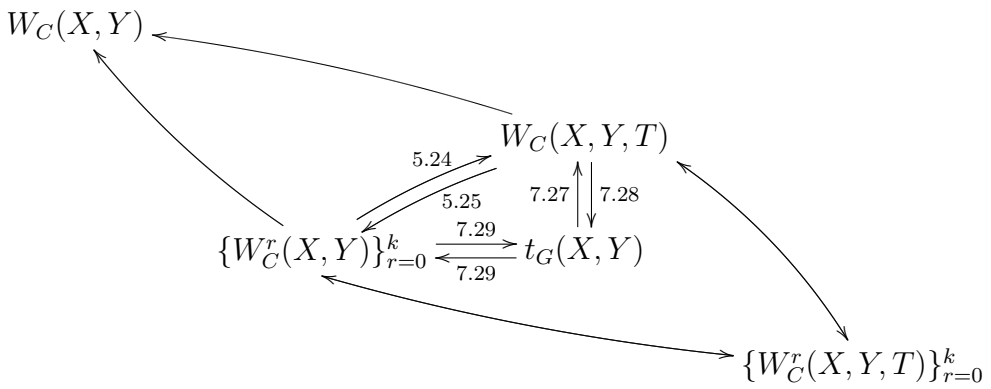
7.4 Give a proof of the statements in Remark 7.12.

7.5 Give a proof of Proposition 7.14.

7.6 Show that all matroids on at most 3 elements are graphic. Give an example of a matroid that is not graphic.

8 Overview

We have established relations between the generalized weight enumerators for $0 \leq r \leq k$, the extended weight enumerator and the Tutte polynomial. We summarize this in the following diagram:



We see that the Tutte polynomial, the extended weight enumerator and the collection of generalized weight enumerators all contain the same amount of information about

a code, because they completely define each other. The original weight enumerator $W_C(X,Y)$ contains less information and therefore does not determine $W_C(X,Y,T)$ or $\{W_C^r(X,Y)\}_{r=0}^k$. See Simonis [34].
One may wonder if the method of generalizing and extending the weight enumerator can be continued, creating the generalized extended weight enumerator, in order to get a stronger invariant. The answer is no: the generalized extended weight enumerator can be defined, but does not contain more information then the three underlying polynomials.
It was shown by Gray [9] that the matroid of a code is a stronger invariant than its Tutte polynomial.

8.1 Exercises

8.1 Show that the extended weight enumerator $W_C(X,Y,T)$ is not determined by the characteristic polynomial $\chi_L(T)$ of the code. Hint: Because $\chi_L(T) = A_n(T)$, it is not difficult to find a counterexample of two codes that have the same $A_n(T)$, but not the same extended weight distribution.

8.2 Investigate whether the Whitney polynomial and the extended weight enumerator determine each other.

8.3 Is the Möbius function or Whitney polynomial of L as defined in Remark 6.37 described in terms of the Tutte polynomial of $M(L)$?

9 McWilliams type property for duality

For both codes and matroids we defined the dual structure. These objects obviousely completely define there dual. But how about the various polynomials associated to a code and a matroid? We know that for example the weight enumerator is a less strong invariant for a code then the code itself: this means there are non-equivalent codes with the same weight enumerator. So it is a priori not clear that the weight enumerator of a code completely defines the weight enumerator of its dual code. We already saw that there is in fact such a relation, namely the MacWilliams identity in Theorem 3.9. We will give a proof of this relation, and consider the similar question for the extended weight enumerator and the extended coset leader weight enumerator, as well as for the Tutte polynomial of a matroid.

9.1 Using the Tutte polynomial

In this section, we will prove the MacWilliams identities using the Tutte polynomial. We do this because of the following very useful relation between the Tutte polynomial of a matroid and its dual:

Theorem 9.1 *Let $t_M(X,Y)$ be the Tutte polynomial of a matroid M, and let $M^\perp$ be the dual matroid. Then $t_M(X,Y) = t_{M^\perp}(Y,X)$.*

Proof. In Proposition 7.17 we proved $r^\perp(J) = |J| - r(M) + r(M \setminus J)$. In particular, we have $r^\perp(M) + r(M) = |M|$. Substituting the last relation into the definition of the Tutte polynomial for the dual code, gives

$$t_{M^\perp}(X,Y) \;=\; \sum_{J \subseteq M^\perp} (X-1)^{r^\perp(M^\perp)-r^\perp(J)} (Y-1)^{|J|-r^\perp(J)}$$

$$
\begin{aligned}
&= \sum_{J\subseteq M} (X-1)^{r^\perp(M^\perp)-|J|-r(M\setminus J)+r(M)}(Y-1)^{r(M)-r(M\setminus J)} \\
&= \sum_{J\subseteq M} (X-1)^{|M\setminus J|-r(M\setminus J)}(Y-1)^{r(M)-r(M\setminus J)} \\
&= t_M(Y,X)
\end{aligned}
$$

In the last step, we use that the summation over all $J \subseteq M$ is the same as a summation over all $M \setminus J \subseteq M$. This proves the theorem. ◇

If we consider a code as a matroid, then the dual matroid is the dual code. Therefore we can use the above theorem to prove the MacWilliams relations. Greene [16] was the first to use this idea, see also Brylawsky and Oxley [11].

Theorem 9.2 (MacWilliams) *Let C be a code and let $C^\perp$ be its dual. Then the extended weight enumerator of C completely determines the extended weight enumerator of $C^\perp$ and vice versa, via the following formula:*

$$W_{C^\perp}(X,Y,T) = T^{-k} W_C(X+(T-1)Y, X-Y, T).$$

Proof. Let G be te matroid associated to the code. Using the previous theorem and the relation between the weight enumerator and the Tutte polynomial, we find

$$
\begin{aligned}
& T^{-k} W_C(X+(T-1)Y, X-Y, T) \\
= \; & T^{-k}(TY)^k (X-Y)^{n-k}\, t_G\left(\frac{X}{Y}, \frac{X+(T-1)Y}{X-Y}\right) \\
= \; & Y^k (X-Y)^{n-k}\, t_{G^\perp}\left(\frac{X+(T-1)Y}{X-Y}, \frac{X}{Y}\right) \\
= \; & W_{C^\perp}(X,Y,T).
\end{aligned}
$$

Notice in the last step that $\dim C^\perp = n-k$, and $n-(n-k) = k$. ◇

9.2 Generalized MacWilliams identities

We can use the relations in Theorems 5.24 and 5.25 to prove the MacWilliams identities for the generalized weight enumerator.

Theorem 9.3 *Let C be a code and let $C^\perp$ be its dual. Then the generalized weight enumerators of C completely determine the generalized weight enumerators of $C^\perp$ and vice versa, via the following formula:*

$$W^r_{C^\perp}(X,Y) = \sum_{j=0}^{r}\sum_{l=0}^{j} (-1)^{r-j} \frac{q^{\binom{r-j}{2}-j(r-j)-l(j-l)-jk}}{\langle r-j\rangle_q \langle j-l\rangle_q} W^l_C(X+(q^j-1)Y, X-Y).$$

Proof. We write the generalized weight enumerator in terms of the extended weight enumerator, use the MacWilliams identities for the extended weight enumerator, and convert back to the generalized weight enumerator.

$$W^r_{C^\perp} = \frac{1}{\langle r\rangle_q} \sum_{j=0}^{r} \begin{bmatrix} r \\ j \end{bmatrix}_q (-1)^{r-j} q^{\binom{r-j}{2}}\, W_{C^\perp}(X,Y,q^i)$$

$$
\begin{aligned}
&= \sum_{j=0}^{r}(-1)^{r-j}\frac{q^{\binom{r-j}{2}-j(r-j)}}{\langle j\rangle_q\langle r-j\rangle_q}q^{-jk}W_c(X+(q^j-1)Y,X-Y,q^j)\\
&= \sum_{j=0}^{r}(-1)^{r-j}\frac{q^{\binom{r-j}{2}-j(r-j)-jk}}{\langle j\rangle_q\langle r-j\rangle_q}\\
&\quad\times\sum_{l=0}^{j}\frac{\langle j\rangle_q}{q^{l(j-l)}\langle j-l\rangle_q}W_C^l(X+(q^j-1,X-Y)\\
&= \sum_{j=0}^{r}\sum_{l=0}^{j}(-1)^{r-j}\frac{q^{\binom{r-j}{2}-j(r-j)-l(j-l)-jk}}{\langle r-j\rangle_q\langle j-l\rangle_q}\\
&\quad\times W_C^l(X+(q^j-1,X-Y).
\end{aligned}
$$

◇

This theorem was proved by Kløve in [23], although the proof uses only half of the relations between the generalized weight enumerator and the extended weight enumerator. Using both makes the proof much shorter.

References

[1] C.A. Anthanasiadis. Characteristic polynomials of subspace arrangements and finite fields. *Advances in Mathematics*, 122:193–233, 1996.

[2] A. Barg. The matroid of supports of a linear code. *AAECC*, 8:165–172, 1997.

[3] A. Barg. Complexity issues in coding theory. In V.S. Pless and W.C. Huffman, editors, *Handbook of coding theory, vol. 1*, pages 649–754. North-Holland, Amsterdam, 1998.

[4] E.R. Berlekamp, R.J. McEliece, and H.C.A. van Tilborg. On the inherent intractability of certain coding problems. *IEEE Transactions on Information*, 24:384–386, 1978.

[5] A. Björner and T. Ekedahl. Subarrangments over finite fields: Chomological and enumerative aspects. *Adv. Math.*, 129:159–187, 1997.

[6] T. Britz. MacWilliams identities and matroid polynomials. *The Electronic Journal of Combinatorics*, 9:R19, 2002.

[7] T. Britz. *Relations, matroids and codes*. PhD thesis, Univ. Aarhus, 2002.

[8] T. Britz. Higher support matroids. *Discrete Mathematics*, 307:2300–2308, 2007.

[9] T. Britz and C.G. Rutherford. Covering radii are not matroid invariants. *Discrete Mathematics*, 296:117–120, 2005.

[10] T.H. Brylawski and J.G. Oxley. The tutte polynomial and its applications. In N. White, editor, *Matroid Applications*. Cambridge University Press, Cambridge, 1992.

[11] Tom Brylawski and James Oxley. Several identities for the characteristic polynomial of a combinatorial geometry. *Discrete Mathematics*, 31(2):161–170, 1980.

[12] P. Cartier. Les arrangements d'hyperplans: un chapitre de géométrie combinatoire. *Seminaire N. Bourbaki*, 561:1–22, 1981.

[13] H. Crapo and G.-C. Rota. *On the foundations of combinatorial theory: Combinatorial geometries*. MIT Press, Cambridge MA, 1970.

[14] Iwan M. Duursma. Combinatorics of the two-variable zeta function. In Gary L. Mullen, Alain Poli, and Henning Stichtenoth, editors, *International Conference on Finite Fields and Applications*, volume 2948 of *Lecture Notes in Computer Science*, pages 109–136. Springer, 2003.

[15] G.D. Forney. Dimension/length profiles and trellis complexity of linear block codes. *IEEE Trans. Inform. Theory*, 40:1741–1752, 1994.

[16] C. Greene. Weight enumeration and the geometry of linear codes. *Studies in Applied Mathematics*, 55:119–128, 1976.

[17] T. Helleseth, T. Kløve, and J. Mykkeltveit. The weight distribution of irreducible cyclic codes with block lengths $n_1((q^l-1)/n)$. *Discrete Mathematics*, 18:179–211, 1977.

[18] R.P.M.J. Jurrius. Classifying polynomials of linear codes. Master's thesis, Leiden University, 2008.

[19] R.P.M.J. Jurrius and R. Pellikaan. Extended and generalized weight enumerators. In T. Helleset and Ø Ytrehus, editors, *Proc. Int. Workshop on Coding and Cryptography WCC-2009*, pages ...–... Selmer Center, Bergen, 2009.

[20] R.P.M.J. Jurrius and R. Pellikaan. The extended coset leader weight enumerator. In F. Willems and T. Tjalkens, editors, *Proc. 30th Symposium 2009 on Information Theory in the Benelux*, pages 217–224. WIC, Eindhoven, 2009.

[21] G.L. Katsman and M.A. Tsfasman. Spectra of algebraic-geometric codes. *Problemy Peredachi Informatsii*, 23:19–34, 1987.

[22] T. Kløve. The weight distribution of linear codes over $\mathrm{GF}(q^l)$ having generator matrix over $\mathrm{GF}(q)$. *Discrete Mathematics*, 23:159–168, 1978.

[23] T. Kløve. Support weight distribution of linear codes. *Discrete Matematics*, 106/107:311–316, 1992.

[24] T. Kløve. *Codes for error detection*. Series on Coding Theory and Cryptology vol. 2. World Scientific Publishing Co. Pte. Ltd., Hackensack, 2007.

[25] David MacKay. *Information theory, inference and learning algorithms*. Cambridge University Press, Cambridge, 2003.

[26] F.J. MacWilliams and N.J.A. Sloane. *The theory of error-correcting Codes*. North-Holland Mathematical Library, Amsterdam, 1977.

[27] M. Munuera. Steganography and error-correcting codes. *Signal Processing*, 87:1528–1533, 2007.

[28] P. Orlik and H. Terao. *Arrangements of hyperplanes*, volume 300. Springer-Verlag, Berlin, 1992.

[29] J.G. Oxley. *Matroid theory*. Oxford University Press, Oxford, 1992.

[30] L.H. Ozarev and A.D. Wyner. Wire-tap channel II. *AT&T Bell labs Techn. J.*, 63:2135–2157, 1984.

[31] R. Pellikaan, X.-W. Wu, and S. Bulygin. Codes and cryptography on algebraic curves. Book in preparation for Cambridge University Press.

[32] T. Richardson and R. Urbanke. *Modern coding theory.* Cambridge University Press, Cambridge, 2008.

[33] G.-C. Rota. On the foundations of combinatorial theory I: Theory of möbius functions. *Zeit. für Wahrsch.*, 2:340–368, 1964.

[34] J. Simonis. The effective length of subcodes. *AAECC*, 5:371–377, 1993.

[35] R.P. Stanley. An introduction to hyperplane arrangements. In *Geometric combinatorics, IAS/Park City Math. Ser., 13*, pages 389–496. Amer. Math. Soc., Providence, RI, 2007.

[36] M.A. Tsfasman and S.G. Vlădut. *Algebraic-geometric codes.* Kluwer Academic Publishers, Dordrecht, 1991.

[37] M.A. Tsfasman and S.G. Vlădut. Geometric approach to higher weights. *IEEE Transactions on Information Theory*, 41:1564–1588, 1995.

[38] W.T. Tutte. A contribution to the theory of chromatic polynomials. *Canad.J. Math.*, 6:80–91, 1954.

[39] W.T. Tutte. On the algebraic theory of graph coloring. *J. Comb. Theory*, 1:15–50, 1966.

[40] W.T. Tutte. On dichromatic polynomials. *J. Comb. Theory*, 2:301–320, 1967.

[41] W.T. Tutte. Cochromatic graphs. *J. Comb. Theory*, 16:168–174, 1974.

[42] W.T. Tutte. Graphs-polynomials. *Advences in Applied Mathematics*, 32:5–9, 2004.

[43] A. Vardy. The intractability of computing the minimum distance of a code. *IEEE Trans. Inform. Theory*, 43:1757–1766, 1997.

[44] V.K. Wei. Generalized Hamming weights for linear codes. *IEEE Transactions on Information Theory*, 37:1412–1418, 1991.

[45] H. Whitney. Colorings of graphs. *Ann. Math.*, 33:688–718, 1932.

[46] H. Whitney. A logical expansion in mathematics. *Bull. Amer. Math. Soc.*, 38:572–579, 1932.

[47] H. Whitney. On the abstract properties of linear dependence. *Amer. J. Math.*, 57:509, 1935.

[48] T. Zaslavsky. *Facing up to arrangements: Face-count fomulas for partitions of space by hyperplanes.* Mem. Amer. Math. Soc. vol. 1, No. 154, Amer. Math. Soc., 1975.

[49] T. Zaslavsky. Signed graph colouring. *Discrete. Math.*, 39:215–228, 1982.

A Codes of dimension 3 and length at most 6

In this Appendix a table is given of the weight enumerator and characteristic polynomials χ_i of all codes C of dimension 3 and length at most 6 with the methods of Section 6.4. In this classification of the codes are denoted by $n.d$, where n denotes the length and d the minimum distance.

no.	G	constraints	$\mu_i \neq 0$
3.1	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$		$\mu_1 = 3T-3$ $\mu_2 = 3$
4.1a	$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$		$\mu_1 = 2T-2$ $\mu_2 = T+1$ $\mu_3 = 2$
4.1b	$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$		$\mu_2 = 3$ $\mu_3 = 1$
4.2	$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$		$\mu_2 = 6$
5.1a	$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$		$\mu_1 = 2T-2$ $\mu_2 = 1$ $\mu_3 = T-1$ $\mu_4 = 2$
5.1b	$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$		$\mu_1 = T-1$ $\mu_2 = 2T-2$ $\mu_3 = 2$ $\mu_4 = 1$
5.1c	$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$		$\mu_1 = 3T-4$ $\mu_2 = T+1$ $\mu_3 = 1$ $\mu_4 = 1$
5.2a	$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$		$\mu_1 = 3T-3$ $\mu_2 = T-2$ $\mu_3 = 4$
5.2b	$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$		$\mu_1 = 3T-6$ $\mu_2 = T+1$ $\mu_3 = 3$
5.2c	$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$		$\mu_1 = 5T-9$ $\mu_2 = 4$ $\mu_3 = 3$
5.2d	$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & a \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$	$a \in \mathbb{F}_q \setminus \{0,1\}, ...$	$\mu_1 = 5T-12$ $\mu_2 = 7$ $\mu_3 = 1$
5.3	$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & a \\ 0 & 0 & 1 & 1 & b \end{pmatrix}$	$a, b \in \mathbb{F}_q \setminus \{0,1\}, ...$	$\mu_1 = 5T-15$ $\mu_2 = 10$

no.	G	constraints	$\mu_i \neq 0$
6.1	$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & a & b \end{pmatrix}$	$a, b \in \mathbb{F}_q \setminus \{0, 1\},\ a \neq b$	$\mu_2 = 5$ $\mu_5 = 1$
6.2a	$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & a \end{pmatrix}$	$a \in \mathbb{F}_q \setminus \{0, 1\}, ...$	$\mu_2 = 6$ $\mu_3 = 1$ $\mu_4 = 1$
6.2b	$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & a & b \end{pmatrix}$	$a, b \in \mathbb{F}_q \setminus \{0, 1\},\ a \neq b$	$\mu_2 = 9$ $\mu_4 = 1$
6.3a	$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$		$\mu_2 = 3$ $\mu_3 = 4$
6.3b	$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & a \end{pmatrix}$	$a \in \mathbb{F}_q \setminus \{0, 1\}, ...$	$\mu_2 = 6$ $\mu_3 = 3$
6.3c	$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & a & b \end{pmatrix}$	$a, b \in \mathbb{F}_q \setminus \{0, 1\},\ a \neq b, ...$	$\mu_2 = 9$ $\mu_3 = 2$
6.3d	$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & b \\ 0 & 0 & 1 & 1 & a & c \end{pmatrix}$	$a, b, c \in \mathbb{F}_q \setminus \{0, 1\},\ ...$	$\mu_2 = 12$ $\mu_3 = 1$
6.4	$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & a & b \\ 0 & 0 & 1 & 1 & c & d \end{pmatrix}$	$a, b, c, d \in \mathbb{F}_q \setminus \{0, 1\},\ ...$	$\mu_2 = 15$

no.	i	1	2	3	4	5
3.1	μ_i	$3T-3$	3			
	$\bar{A}_i$	3	$3T-3$	T^2-2T+1		
	$\bar{\chi}_{3-i}$	3	$3T-3$	T^2-2T+1		
4.1a	μ_i	$2T-2$	T	2		
	$\bar{A}_i$	2	T	$2T-2$	T^2-2T+1	
	$\bar{\chi}_{3-i}$	3	$3T-3$	T^2-2T+1		
4.1b	μ_i	3	1	0		
	$\bar{A}_i$	1	3	$4T-5$	T^2-3T+2	
	$\bar{\chi}_{3-i}$	4	$4T-5$	T^2-3T+2		
4.2	μ_i	0	6	0		
	$\bar{A}_i$	0	6	$4T-8$	T^2-3T+3	
	$\bar{\chi}_{3-i}$	6	$4T-8$	T^2-3T+3		
5.1a	μ_i	$2T-2$	1	$T-1$	2	
	$\bar{A}_i$	2	$T-1$	1	$2T-2$	T^2-2T+1
	$\bar{\chi}_{3-i}$	3	$3T-3$	T^2-2T+1		
5.1b	μ_i	$T-1$	$2T-2$	2	1	
	$\bar{A}_i$	1	2	$2T-2$	$T-1$	T^2-2T+1
	$\bar{\chi}_{3-i}$	3	$3T-3$	T^2-2T+1		
5.1c	μ_i	$3T-4$	$T+1$	1	1	
	$\bar{A}_i$	1	1	$T+1$	$3T-4$	T^2-3T+2
	$\bar{\chi}_{3-i}$	4	$4T-5$	T^2-3T+2		
5.2a	μ_i	$3T-3$	$T-2$	4	0	
	$\bar{A}_i$	0	4	$T-2$	$3T-3$	T^2-3T+2
	$\bar{\chi}_{3-i}$	4	$4T-5$	T^2-3T+2		
5.2b	μ_i	$3T-6$	$T+1$	3	0	
	$\bar{A}_i$	0	3	$T+1$	$3T-6$	T^2-3T+3
	$\bar{\chi}_{3-i}$	6	$4T-8$	T^2-3T+3		
5.2c	μ_i	$5T-9$	4	2	0	
	$\bar{A}_i$	0	2	4	$5T-9$	T^2-4T+4
	$\bar{\chi}_{3-i}$	6	$5T-9$	T^2-4T+4		
5.2d	μ_i	$5T-12$	7	1	0	
	$\bar{A}_i$	0	1	7	$5T-12$	T^2-4T+5
	$\bar{\chi}_{3-i}$	8	$5T-12$	T^2-4T+5		
5.3	μ_i	$5T-15$	10	0	0	
	$\bar{A}_i$	0	0	10	$5T-15$	T^2-4T+6
	$\bar{\chi}_{3-i}$	10	$5T-15$	T^2-4T+6		

no.	i	1	2	3	4	5	6
6.1	μ_i		5	0	0	1	
	$\bar{A}_i$	1	0	0	5	$6T-9$	T^2-5T+4
	$\bar{\chi}_{3-i}$	6	$6T-9$	T^2-5T+4			
6.2a	μ_i		6	1	1	0	0
	$\bar{A}_i$	0	1	1	6	$6T-13$	T^2-5T+6
	$\bar{\chi}_{3-i}$	8	$6T-13$	T^2-5T+6			
6.2b	μ_i		9	0	1	0	0
	$\bar{A}_i$	0	1	0	9	$6T-16$	T^2-5T+7
	$\bar{\chi}_{3-i}$	10	$6T-16$	T^2-5T+7			
6.3a	μ_i		3	4	0	0	0
	$\bar{A}_i$	0	0	4	3	$6T-12$	T^2-5T+6
	$\bar{\chi}_{3-i}$	7	$6T-12$	T^2-5T+6			
6.3b	μ_i		6	3	0	0	0
	$\bar{A}_i$	0	0	3	6	$6T-15$	T^2-5T+7
	$\bar{\chi}_{3-i}$	9	$6T-15$	T^2-5T+7			
6.3c	μ_i		9	2	0	0	0
	$\bar{A}_i$	0	0	2	9	$6T-18$	T^2-5T+8
	$\bar{\chi}_{3-i}$	11	$6T-18$	T^2-5T+8			
6.3d	μ_i		12	1	0	0	0
	$\bar{A}_i$	0	0	1	12	$6T-21$	T^2-5T+9
	$\bar{\chi}_{3-i}$	11	$6T-21$	T^2-5T+9			
6.4	μ_i		15	0	0	0	0
	$\bar{A}_i$	0	0	0	15	$6T-24$	$T^2-5T+10$
	$\bar{\chi}_{3-i}$	15	$6T-24$	$T^2-5T+10$			