Chapter 1

Codes, arrangements and matroids

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

This chapter treats error-correcting codes and their weight enumerator as the center of several closely related topics such as arrangements of hyperplanes, graph theory, matroids, posets and geometric lattices and their characteristic, chromatic, Tutte, Möbius and coboundary polynomial, respectively. Their interrelations and many examples and counterexamples are given. It is concluded with a section with references to the literature for further reading and open questions.

AMS classification: 05B35, 05C31, 06A07, 14N20, 94B27, 94B70, 94C15

1.1. Introduction

A lot of mathematical objects are closely related to each other. While studying certain aspects of a mathematical object, one tries to find a way to "view" the object in a way that is most suitable for a specific problem. Or in other words, one tries to find the best way to model the problem. Many related fields of mathematics have evolved from one another this way. In practice, it is very useful to be able to transform your problem into other terminology: it gives a lot more available knowledge that can be helpful to solve a problem.

In this chapter we give a broad overview of closely related fields, starting from the weight enumerator of an error-correcting code. We explain the importance of this polynomial in coding theory. From various methods of determining the weight enumerator, we naturally run into other ways to

view an error-correcting code. We will introduce and link the following mathematical objects:

- linear codes and their weight enumerator (§1.2, §1.3, §1.5);
- arrangements and their characteristic polynomial (§1.4, §1.8);
- graphs and their chromatic polynomial (§1.6.2)
- matroids and their Tutte polynomial (§1.6);
- posets and their Möbius function (§1.7);
- geometric lattices and their coboundary polynomial (§1.7, §1.8).

A nice example to show the power of these connections are the MacWilliams identities, that relate the polynomials associated to an object and its dual. This will be treated in Section 1.6.5. Several examples and counterexamples are given in Section 1.8.6 and an overview is given to show which polynomials determine each other in Section 1.9.

These notes are based on the Master's thesis [1], ongoing research [2, 3] and the lecture notes [4] of the Soria Summer School in 2009. The whole chapter is sell-contained, and various references to further reading, background knowledge and open problems are given in Section 1.10.

1.2. Error-correcting codes

The basics of the theory of error-correcting codes one can find in [5–8].

1.2.1. *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 chapter. 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 the nowadays famous code named after him. Whereas the theory of Hamming [9] is about the actual construction, the encoding and decoding of codes and uses tools from *combinatorics* and *algebra*, the approach of Shannon [10] leads to *information theory* and his theorems tell us what is and what is not possible in a *probabilistic* sense.

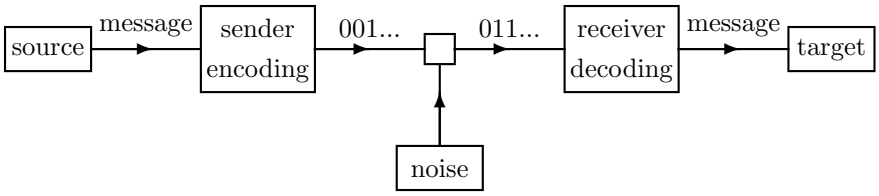


Fig. 1.1. Block diagram of a communication system

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 even tells 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 1.1. 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. We call this a *repetition code*. 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 1.2. 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 *Venn-diagram* of three intersecting circles as given in Figure 1.2. 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.

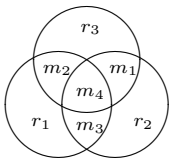


Fig. 1.2. Venn diagram of the Hamming code

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. Suppose that after transmission one symbol is flipped and $\mathbf{y} = (1, 0, 0, 1, 0, 0, 1)$ is

the received word as given in Figure 1.3.

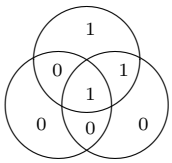


Fig. 1.3. Venn diagram of a received word for the Hamming code

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, if 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 2:

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

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 information 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.

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 a (n, M) code. If $M = q^k$, then C is called a $[n, k]$ code. For a (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 1.3. 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 $4/7$. These are the same values as we found in Examples 1.1 and 1.2

Example 1.4. Let C be the binary block code of length n consisting of all words with exactly two ones. This is a $(n, n(n-1)/2)$ code. In this example the number of codewords is not a power of the size of the alphabet.

Let C be a $[n, k]$ 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*. 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

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

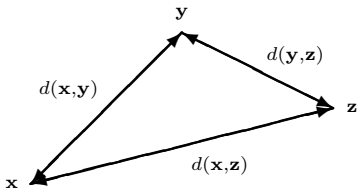


Fig. 1.4. Triangle inequality

Proposition 1.1. *The Hamming distance is a* metric *on Q^n, that means that it has the following properties for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in Q^n$:*

(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).

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

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

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

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 of 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 1.5. 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.

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

Figure 1.5 shows the ball in the Euclidean plane. This is misleading in some respects, but gives an indication what we should have in mind.
Figure 1.6 shows Q^2, where the alphabet Q consists of 5 elements. The ball $B_0(\mathbf{x})$ consists of the point 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 1.2. *Let Q be an alphabet of q elements and $\mathbf{x} \in Q^n$. Then*

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

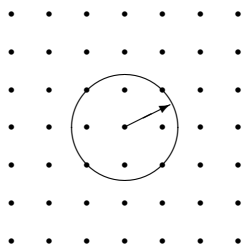


Fig. 1.5. Ball of radius $\sqrt{2}$ in the Euclidean plane

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 $B_r(\mathbf{x})$. □

1.2.2. *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.

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 $[n, k, d]_q$ or $[n, k, d]$ the *parameters* of

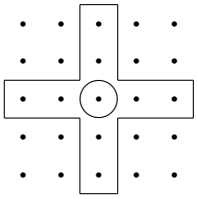


Fig. 1.6. Balls of radius 0 and 1 in the Hamming metric

the code.

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

Let C and D be linear codes in $\mathbb{F}_q^n$. Then C is called *permutation equivalent* to D, if there exists a permutation matrix Π such that $\Pi(C) = D$. If moreover $C = D$, then Π is called an *permutation automorphism* of C. The code C is called *generalized equivalent* or *monomial equivalent* to D, if there exists a monomial matrix M such that $M(C) = D$. If moreover $C = D$, then M is called a *monomial automorphism* of C.

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 $\mathrm{supp}(\mathbf{x}) = \{i : x_i \neq 0\}$. 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 is defined as the minimal value of the weights of the nonzero codewords in case there is a nonzero codeword, and $n + 1$ otherwise.

Proposition 1.3. *The minimum distance of a linear code C of dimension $k > 0$ 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 1.6. 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 1.7. 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 2 by definition, the minimum distance of C is 2 if $q = 2$ and 1 otherwise. The code C is linear if and only if $q = 2$.

Example 1.8. The *Hamming code* C of Example 1.2 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 in 7 variables. The minimum weight is 3 as shown in Example 1.5. So it is a $[7, 4, 3]$ code.

1.2.3. *Generator matrix*

Let C be a linear $[n, k]$ 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.

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

A $k \times n$ matrix G with entries in $\mathbb{F}_q$ is called a *generator matrix* of a $\mathbb{F}_q$-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 a $\mathbb{F}_q$-linear $[n, k]$ code.

Example 1.9. 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 $n \times n$ identity matrix I_n as generator matrix, respectively.

Example 1.10. The repetition code of length n has generator matrix

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

Example 1.11. The Hamming code C of Example 1.2 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

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

Let C be a $[n, k]$ code. The code is called *systematic at the positions* $(j_1, \ldots, j_k)$ if for all $\mathbf{m} \in \mathbb{F}_q^k$ there exists a unique codeword $\mathbf{c}$ such that $c_{j_i} = m_i$ for all $i = 1, \ldots, k$. In that case, the set $(j_1, \ldots, j_k)$ is called an *information set*. A generator matrix G of C is called systematic at the positions $(j_1, \ldots, j_k)$ if the $k \times k$ submatrix G' consisting of the k columns of G at the positions $(j_1, \ldots, j_k)$ is the identity matrix. For such a matrix G the mapping $\mathbf{m} \mapsto \mathbf{m}G$ is called *systematic encoding*.

1.2.4. *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 a $\mathbb{F}_q$-linear $[n, k]$ code. Suppose that H is a $m \times n$ 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 a $(n-k) \times n$ matrix H' with independent rows and with the same null space as H. So H' has rank $n-k$. A $(n-k) \times n$ matrix of rank $n-k$ is called a *parity check matrix* of a $[n, k]$ code C if C is the null space of this matrix.

Remark 1.1. 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 1.4. *Suppose C is a $[n, k]$ 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:

$$\begin{aligned}
& \mathbf{c} \text{ is a codeword,} \\
\iff & -\mathbf{m}P + \mathbf{r} = 0, \\
\iff & -P^T\mathbf{m}^T + \mathbf{r}^T = 0, \\
\iff & \left(-P^T|I_{n-k}\right)(\mathbf{m}, \mathbf{r})^T = 0, \\
\iff & \left(-P^T|I_{n-k}\right)\mathbf{c}^T = 0.
\end{aligned}$$

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

Example 1.12. The trivial codes $\{0\}$ and $\mathbb{F}_q^n$ have I_n and the empty matrix as parity check matrix, respectively.

Example 1.13. As a consequence of Proposition 1.4 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.

Example 1.14. The generator matrix G of the Hamming code C in Example 1.11 is of the form $(I_4|P)$ and in Example 1.8 we see that the parity check matrix is equal to $(P^T|I_3)$.

1.2.5. *Inner product and dual codes*

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.

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

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

Proposition 1.5. *Let C be a $[n, k]$ code with generator matrix G. Then $C^\perp$ is a $[n, n-k]$ code with parity check matrix G.*

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

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

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 $C^\perp$. □

Example 1.15. The trivial codes $\{0\}$ and $\mathbb{F}_q^n$ are dual codes.

Example 1.16. The binary even weight code and the repetition code of the same length are dual codes.

1.2.6. *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 1.6. *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 $\text{supp}(\mathbf{c}) = \{j_1, \ldots, j_w\}$ with $1 \leq j_1 < \cdots < j_w \leq n$. 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. □

Let H be a parity check matrix of a code C. As a consequence of Proposition 1.6 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 1.8. 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 a $\mathbb{F}_q$-linear code C with $d(C) \geq 3$ and redundancy r is at most $(q^r - 1)/(q - 1)$.

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 $\mathcal{S}_r(q)$.The simplex code $\mathcal{S}_r(q)$ and the Hamming code $\mathcal{H}_r(q)$ are dual codes.

Proposition 1.7. *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 1.6. □

Example 1.17. Consider the following ternary Hamming code $\mathcal{H}_3(3)$ 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 1.7 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 1.8. *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 1.7 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 nonzero coordinates which is q^{r-1}. □

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 1.1. (The Singleton Bound) *If C is a $[n, k, d]$ code, then*

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

Proof. Let H be a parity check matrix of C. This is a $(n-k) \times n$ 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 1.6. 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. □

Let C be a $[n,k,d]$ code. If $d = n-k+1$, then C is called a *maximum distance separable code* or a *MDS code*, for short. 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 1.18. The minimum distance of 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 a $[n,n-1,2]$ code and both are MDS.

Proposition 1.9. *For a $[n,k,d]$ 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 of C are linearly independent,
(3) every k columns of a generator matrix G of C are linearly independent.

Proof. Let H be a parity check matrix of a $[n,k,d]$ code C. As the minimum distance of C is d, any $d-1$ columns of H are linearly independent by Proposition 1.6. 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 that is zero at k given coordinates. 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 coordinates 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 a $[n,k,n-k+1]$ 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 a 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. □

Proposition 1.10. *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. All $k \times k$ submatrices are Vandermonde matrices, and their determinant is not zero, since the a_i are mutually distinct. □

1.3. Weight enumerators and error probability

1.3.1. *Weight spectrum*

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

Let C be a code of length n. The *weight spectrum* or *weight distribution* is the following set

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

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

The so-called *weight enumerator* of a code C is a convenient representation of the weight spectrum. It 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

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

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:

$$W_C(Z) = W_C(1, Z) \quad \text{and} \quad W_C(X, Y) = X^n W_C(X^{-1}Y).$$

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

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

and

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

Example 1.19. 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 1.20. 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 1.21. 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 1.22. The simplex code $\mathcal{S}_r(q)$ is a constant weight code by Proposition 1.8 with parameters $[(q^r-1)/(q-1), r, q^{r-1}]$. Hence its homogeneous weight enumerator is

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

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:

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

It also determines the dimension k of C, since

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

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

Theorem 1.2 (MacWilliams). *Let C be a $[n,k]$ code over $\mathbb{F}_q$. Then*

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

Proof. See [8, Ch.5. §2. Theorem 1] for a proof for binary codes. A general proof will be given via matroids in Theorem 1.13. □

The computation of the minimum distance and the weight enumerator of a code is NP-hard [11–13].

Example 1.23. 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 1.19, which is indeed equal to $q^0 W_C(X + (q-1)Y, X - Y)$ and confirms MacWilliams identity.

Example 1.24. 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 1.20, which is equal to $2^{-1} W_C(X+Y, X-Y)$, confirming the MacWilliams identity for $q = 2$. For arbitrary q we have

$$\begin{aligned} W_{C^\perp}(X,Y) &= q^{-1} W_C(X + (q-1)Y, X - Y) \\ &= q^{-1}((X + (q-1)Y)^n + (q-1)(X-Y)^n) \\ &= \sum_{w=0}^{n} \binom{n}{w} \frac{(q-1)^w + (q-1)(-1)^w}{q} X^{n-w} Y^w. \end{aligned}$$

1.3.2. *The decoding problem*

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 $\mathrm{wt}(\mathbf{e})$ 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. The number $\lfloor (d(C)-1)/2 \rfloor$ is called the *error-correcting capacity* of the code C and is denoted by $e(C)$.

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.

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.

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 ca-

pacity of C.

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 $\mathbf{r}$.

Let $\mathbf{r}$ be a received word with respect to a code C. We call the set $\mathbf{r} + C = \{\mathbf{r} + \mathbf{c} : \mathbf{c} \in C\}$ the *coset* of $\mathbf{r}$ in C. If $\mathbf{r}$ is a codeword, the coset is equal to the code itself. If $\mathbf{r}$ is not a codeword, the coset is not a linear subspace. A *coset leader* of $\mathbf{r}+C$ is a choice of an element of minimal weight in the coset $\mathbf{r} + C$.

The choice of a coset leader of the coset $\mathbf{r}+C$ is unique if $d(C, \mathbf{r}) \leq (d-1)/2$. 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 $\rho(C)$.

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. The coset leader decoder is a nearest neighbor decoder. 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 [11].

1.3.3. *The q-ary symmetric channel*

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 $0 \leq p \leq \frac{q-1}{q}$, 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 1.2. 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.

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*.

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

$$P_{cd}(p) + P_{de}(p) + P_{df}(p) = 1 \text{ for all } 0 \leq p \leq \frac{q-1}{q}.$$

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

$$P_{err}(p) = P_{de}(p) + P_{df}(p).$$

Proposition 1.11. *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

$$\begin{aligned} P_{cd}(p) &= \sum_{\mathbf{x}\in C} P(\mathbf{x}) \sum_{d(\mathbf{x},\mathbf{y})\leq t} P(\mathbf{y}|\mathbf{x}) \\ &= \frac{1}{|C|} \sum_{\mathbf{x}\in C} \sum_{d(\mathbf{x},\mathbf{y})\leq t} P(\mathbf{y}|\mathbf{x}) \\ &= \sum_{w=0}^{t} \binom{n}{w} (q-1)^w \left(\frac{p}{q-1}\right)^w (1-p)^{n-w} \end{aligned}$$

by Proposition 1.2 and Remark 1.2. Clearing the factor $(q-1)^w$ in the numerator and the denominator gives the desired result. □

In Proposition 1.14 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 1.25. 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 1.26. Consider the binary n-fold repetition code. Let $t = (n-1)/2$. Use the decoding algorithm correcting all patterns of t errors. Then by Proposition 1.11 we have

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

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 [10] 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 of 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.

1.3.4. *Error probability*

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 1.1. 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 [14].

Proposition 1.12. *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 1.2. Hence

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

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

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

Example 1.27. 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 1.20.

Example 1.28. 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 1.21.

Proposition 1.13. *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*

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

Proof. See [6]. 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

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

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

Proposition 1.14. *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. □

1.4. 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 [15–17].

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 1.15. *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 [15–17]. □

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 *simple* if all the n hyperplanes are mutually distinct. The arrangement is called *central* if all the hyperplanes are linear subspaces. A central arrangement is called *essential* if the intersection of all its hyperplanes is equal to $\{0\}$.

Let $G = (g_{ij})$ be a generator matrix of a nondegenerate code C of dimension k. So G has no zero columns. Let H_j be the linear hyperplane in $\mathbb{F}_q^k$ with equation

$$g_{1j}X_1 + \cdots + g_{kj}X_k = 0.$$

The arrangement $(H_1, \ldots, H_n)$ associated with G will be denoted by $\mathcal{A}_G$.

In case of a central 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 nondegenerate $[n, k, d]$ 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 $\mathbb{P}^{k-1}(\mathbb{F}_q)$.

We can translate Proposition 1.15 for an arrangement.

Proposition 1.16. *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. See [15–17]. □

A code C is called *projective* if $d(C^\perp) \geq 3$. Let G be a generator matrix of 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.

1.5. 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 1.16. 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*:

$$\begin{aligned} A_n &= q^k - |H_1 \cup \cdots \cup H_n| \\ &= q^k + \sum_{w=1}^{n} (-1)^w \sum_{\substack{i_1 < \cdots < i_w \\ i_j \in [n]}} |H_{i_1} \cap \cdots \cap H_{i_w}|. \end{aligned}$$

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 [15]. Later we will encounter two more methods: by matroids and the Tutte polynomial in Section 1.6.3 and by geometric lattices and the characteristic polynomial in Section 1.7.

Definition 1.1. For a subset J of $[n] := \{1, 2, \ldots, n\}$ define

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

Remark 1.3. 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 1.16. Furthermore B_J is equal to the number of nonzero codewords $\mathbf{c}$ that are zero at all j in J, and this is equal to the number of nonzero elements of the intersection $\bigcap_{j \in J} H_j$.

Proposition 1.17. *We have the following connection between the B_t and the weight distribution of a code:*

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

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 determine the generalized weight enumerators.

1.5.1. *Generalized weight enumerators*

The notion of the generalized weight enumerator was first introduced by Helleseth, Kløve and Mykkeltveit [18, 19] and later studied by Wei [20]. See also [21]. This notion has applications in the wire-tap channel II [22] and trellis complexity [23].

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 that 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:

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

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*.

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

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 $q - 1$ nonzero 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 1.1. *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 the projection map $\pi : C \to \mathbb{F}_q^t$ given by deleting the coordinates that are not indexed by J. Then π is a linear map, the image of C under π is C_J and the kernel is $C(J)$ by definition. It follows that $\dim C_J + \dim C(J) = \dim C$. So $l(J) = k - r(J)$. □

Lemma 1.2. *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*

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

Proof. Let $t > n - d$ 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 1.1 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 1.6, so $r(J) = t$ and $l(J) = k - t$. □

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 $k - t = 0$.
We introduce the following notations:

$$\begin{aligned} [m, r]_q &= \prod_{i=0}^{r-1} (q^m - q^i) \\ \langle r \rangle_q &= [r, r]_q \\ \begin{bmatrix} k \\ r \end{bmatrix}_q &= \frac{[k, r]_q}{\langle r \rangle_q}. \end{aligned}$$

Remark 1.4. 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 $\mathbb{F}_q^k$.

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

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

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 1.18. *Let r be a positive integer. 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 a direct corollary of Lemma 1.2, 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 $t > n - d_r$. □

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 1.19. *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. □

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 1.3. *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 1.20. *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 1.3; this method is left as an exercise. Instead, we first prove the following general lemma:

Lemma 1.3. *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 with entries $\binom{i}{j}$ and $(-1)^{i+j}\binom{i}{j}$, respectively. 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 identity matrix of size $n+1$, and therefore the matrices are each other's inverse. □

Proof. **(Proposition 1.20)** The proposition is now a direct consequence of Proposition 1.19 and Lemma 1.3. □

1.5.2. *Extended weight enumerator*

Let G be the generator matrix of a linear $[n, k]$ code C over $\mathbb{F}_q$. 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 $C \otimes \mathbb{F}_{q^m}$ of weight w by $A_{C\otimes\mathbb{F}_{q^m},w}$ and the number of subspaces in $C \otimes \mathbb{F}_{q^m}$ of dimension r and weight w by $A^{(r)}_{C\otimes\mathbb{F}_{q^m},w}$. 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 1.1 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, that is an extension of Definition 1.1.

Definition 1.2. 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 1.21. *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 proposition and its proof are generalizations of Proposition 1.17 and its proof. The proof is a direct consequence of Lemma 1.2. 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 1.4. *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

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

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 1.22. *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 1.3 and Theorem 1.4. □

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 extension code $C \otimes \mathbb{F}_{q^m}$.

Proposition 1.23. *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 $A_0(q^m) = A_{C\otimes\mathbb{F}_{q^m},0} = 1$, so assume $w \neq 0$. It is enough to show that $A_w(q^m) = (q^m-1) A^{(1)}_{C\otimes\mathbb{F}_{q^m},w}$. 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. □

Because of Proposition 1.23 we interpret $W_C(X, Y, T)$ 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 [24].

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

Proposition 1.24. *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

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

□

1.5.3. *Puncturing and shortening of codes*

There are several ways to get new codes from existing ones. In this section, we will focus on *puncturing* and *shortening* of codes and show how they are used in an alternative algorithm for finding the extended weight enumerator. The algorithm is based on the Tutte-Grothendieck decomposition of matrices introduced by Brylawski [25]. Greene [26] used this decomposition for the determination of the weight enumerator.

Let C be a linear $[n,k]$ code and let $J \subseteq [n]$. Then the code C *punctured by* J is obtained by deleting all the coordinates indexed by J from the codewords of C. The length of this punctured code is $n-|J|$ and the dimension is at most k. Let C be a linear $[n,k]$ code and let $J \subseteq [n]$. If we puncture the code $C(J)$ by J, we get the code C *shortened by* J. The length of this shortened code is $n-|J|$ and the dimension is $l(J)$.
The operations of puncturing and shortening a code are each others dual: puncturing a code C by J and then taking the dual, gives the same code as shortening $C^{\perp}$ by J.

We have seen that we can determine the extended weight enumerator of a $[n,k]$ code C with the use of a $k \times n$ generator matrix of C. This concept

can be generalized for arbitrarily matrices, not necessarily of full rank.

Let $\mathbb{F}$ be a field. Let G be a $k \times n$ matrix over $\mathbb{F}$, possibly of rank smaller than k and with zero columns. Then for each $J \subseteq [n]$ we define

$$l(J) = l(J, G) = k - r(G_J).$$

as in Lemma 1.1. Define the extended weight enumerator $W_G(X, Y, T)$ as in Definition 1.2. We can now make the following remarks about $W_G(X, Y, T)$.

Proposition 1.25. *Let G be a $k \times n$ matrix over $\mathbb{F}$ and $W_G(X, Y, T)$ the associated extended weight enumerator. Then the following statements hold:*

(i) $W_G(X, Y, T)$ *is invariant under row-equivalence of matrices.*
(ii) Let G' be a $l \times n$ matrix with the same row-space as G, then we have $W_G(X, Y, T) = T^{k-l} W_{G'}(X, Y, T)$. In particular, if G is a generator matrix of a $[n, k]$ code C, we have $W_G(X, Y, T) = W_C(X, Y, T)$.
(iii) $W_G(X, Y, T)$ *is invariant under permutation of the columns of G.*
(iv) $W_G(X, Y, T)$ *is invariant under multiplying a column of G with an element of $\mathbb{F}^*$.*
(v) If G is the direct sum of G_1 and G_2, i.e. of the form

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

then $W_G(X, Y, T) = W_{G_1}(X, Y, T) \cdot W_{G_2}(X, Y, T)$.

Proof. (i) If we multiply G from the left with an invertible $k \times k$ matrix, the $r(J)$ do not change, and therefore (i) holds.
For (ii), we may assume without loss of generality that $k \geq l$. Because G and G' have the same row-space, the ranks $r(G_J)$ and $r(G'_J)$ are the same. So $l(J, G) = k - l + l(J, G')$. Using Proposition 1.24 we have for G

$$\begin{aligned} W_G(X, Y, T) &= \sum_{t=0}^{n} \sum_{|J|=t} T^{l(J,G)} (X - Y)^t Y^{n-t} \\ &= \sum_{t=0}^{n} \sum_{|J|=t} T^{k-l+l(J,G')} (X - Y)^t Y^{n-t} \\ &= T^{k-l} \sum_{t=0}^{n} \sum_{|J|=t} T^{l(J,G')} (X - Y)^t Y^{n-t} \\ &= T^{k-l} W_{G'}(X, Y, T). \end{aligned}$$

The last part of (ii) and (iii)–(v) follow directly from the definitions. □

With the use of the extended weight enumerator for general matrices, we can derive a recursive algorithm to determine the extended weight enumerator of a code. Let G be a $k \times n$ matrix with entries in $\mathbb{F}$. Suppose that the j-th column is not the zero vector. Then there exists a matrix row-equivalent to G such that the j-th column is of the form $(1, 0, \ldots, 0)^T$. Such a matrix is called *reduced* at the j-th column. In general, this reduction is not unique.

Let G be a matrix that is reduced at the j-th column a. The matrix $G \setminus a$ is the $k \times (n-1)$ matrix G with the column a removed, and G/a is the $(k-1) \times (n-1)$ matrix G with the column a and the first row removed. We can view $G \setminus a$ as G punctured by a, and G/a as G shortened by a.

For the extended weight enumerators of these matrices, we have the following connection (we omit the (X, Y, T) part for clarity):

Proposition 1.26. *Let G be a $k \times n$ matrix that is reduced at the j-th column a. For the extended weight enumerator of a reduced matrix G holds*

$$W_G = (X - Y)W_{G/a} + YW_{G\setminus a}.$$

Proof. We distinguish between two cases here. First, assume that $G \setminus a$ and G/a have the same rank. Then we can choose a G with all zeros in the first row, except for the 1 in the column a. So G is the direct sum of 1 and G/a. By Proposition 1.25 parts (v) and (ii) we have

$$W_G = (X + (T-1)Y)W_{G/a} \qquad \text{and} \qquad W_{G\setminus a} = TW_{G/a}.$$

Combining the two gives

$$\begin{aligned} W_G &= (X + (T-1)Y)W_{G/a} \\ &= (X - Y)W_{G/a} + YTW_{G/a} \\ &= (X - Y)W_{G/a} + YW_{G\setminus a}. \end{aligned}$$

For the second case, assume that $G \setminus a$ and G/a do not have the same rank. So $r(G \setminus a) = r(G/a) + 1$. This implies G and $G \setminus a$ do have the same rank. We have that

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

by Proposition 1.24. This double sum splits into the sum of two parts by distinguishing between the cases $j \in J$ and $j \notin J$.

Let $j \in J$, $t = |J|$, $J' = J \setminus \{j\}$ and $t' = |J'| = t - 1$. Then

$$l(J', G/a) = k - 1 - r((G/a)_{J'}) = k - r(G_J) = l(J, G).$$

So the first part is equal to

$$\sum_{t=0}^{n} \sum_{\substack{|J|=t \\ j \in J}} T^{l(J,G)}(X-Y)^t Y^{n-t} = \sum_{t'=0}^{n-1} \sum_{|J'|=t'} T^{l(J',G/a)}(X-Y)^{t'+1} Y^{n-1-t'}$$

which is equal to $(X - Y)W_{G/a}$.
Let $j \notin J$. Then $(G \setminus a)_J = G_J$. So $l(J, G \setminus a) = l(J, G)$. Hence the second part is equal to

$$\sum_{t=0}^{n} \sum_{\substack{|J|=t \\ j \notin J}} T^{l(J,G)}(X-Y)^t Y^{n-t} = Y \sum_{t'=0}^{n-1} \sum_{\substack{|J|=t' \\ j \notin J}} T^{l(J,G\setminus a)}(X-Y)^{t'} Y^{n-1-t'}$$

which is equal to $YW_{G\setminus a}$. □

Theorem 1.5. *Let G be a $k \times n$ matrix over $\mathbb{F}$ with $n > k$ of the form $G = (I_k|P)$, where P is a $k \times (n-k)$ matrix over $\mathbb{F}$. Let $A \subseteq [k]$ and write P_A for the matrix formed by the rows of P indexed by A. Let $W_A(X, Y, T) = W_{P_A}(X, Y, T)$. Then the following holds:*

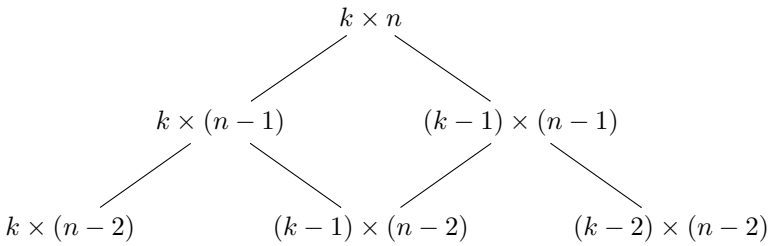
$$W_C(X, Y, T) = \sum_{l=0}^{k} \sum_{|A|=l} Y^l (X-Y)^{k-l} W_A(X, Y, T).$$

Proof. We use the formula of the last proposition recursively. We denote the construction of $G \setminus a$ by G_1 and the construction of G/a by G_2. Repeating this procedure, we get the matrices G_{11}, G_{12}, G_{21} and G_{22}. So we get for the weight enumerator

$$W_G = Y^2 W_{G_{11}} + Y(X-Y)W_{G_{12}} + Y(X-Y)W_{G_{21}} + (X-Y)^2 W_{G_{22}}.$$

Repeating this procedure k times, we get 2^k matrices with $n-k$ columns and $0, \ldots, k$ rows, which form exactly the P_A. In the diagram are the sizes of the matrices of the first two steps: note that only the $k \times n$ matrix on top has to be of full rank. The number of matrices of size $(k-i) \times (n-j)$

is given by the binomial coefficient $\binom{j}{i}$.



On the last line we have $W_0(X, Y, T) = X^{n-k}$. This proves the formula. □

Example 1.29. Let C be the even weight code of length $n = 6$ over $\mathbb{F}_2$. Then a generator matrix of C is the 5×6 matrix $G = (I_5|P)$ with $P = (1, 1, 1, 1, 1)^T$. So the matrices P_A are $l \times 1$ matrices with all ones. We have $W_0(X, Y, T) = X$ and $W_l(X, Y, T) = T^{l-1}(X + (T-1)Y)$ by part (ii) of Proposition 1.25. Therefore the weight enumerator of C is equal to

$$\begin{aligned} W_C(X, Y, T) &= W_G(X, Y, T) \\ &= X(X-Y)^5 + \sum_{l=1}^{5} \binom{5}{l} Y^l (X-Y)^{5-l} T^{l-1} (X + (T-1)Y) \\ &= X^6 + 15(T-1)X^4Y^2 + 20(T^2 - 3T + 2)X^3Y^3 \\ &\quad +15(T^3 - 4T^2 + 6T - 3)X^2Y^4 \\ &\quad +6(T^4 - 5T^3 + 10T^2 - 10T + 4)XY^5 \\ &\quad +(T^5 - 6T^4 + 15T^3 - 20T^2 + 15T - 5)Y^6. \end{aligned}$$

For $T = 2$ we get $W_C(X, Y, 2) = X^6 + 15X^4Y^2 + 15X^2Y^4 + Y^6$, which we indeed recognize as the weight enumerator of the even weight code that we found in Example 1.20.

1.5.4. *Connections*

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

Proposition 1.27. *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 a 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 a $m \times n$ 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 of $\mathbb{F}_q$-vector spaces $C \otimes \mathbb{F}_{q^m} \to C^m$. □

Note that this isomorphism depends on the choice of a primitive element α. The use of this isomorphism for the proof of Theorem 1.6 was suggested by Simonis [21]. We also need the next subresult.

Lemma 1.4. *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 $\mathrm{wt}(\mathbf{c}) = \mathrm{wt}(D)$. □

Proposition 1.28. *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 1.27. 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 a $r \times n$ matrix whose rows are words of C. Left multiplication by a $m \times r$ matrix of rank r gives an element of C^m that 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. □

This result first appears in [18, Theorem 3.2], although the term "generalized weight enumerator" was yet to be invented. In general, we have the following theorem.

Theorem 1.6. *Let C be a linear code over $\mathbb{F}_q$. Then the extended weight numerator and the generalized weight enumerators 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. □

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

Theorem 1.7. *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 1.24. 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, which can be proven by induction:

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

See [19, 27–30]. □

1.5.5. *MDS-codes*

We can use the theory in Sections 1.5.1 and 1.5.2 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 1.29. *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 an MDS code is also MDS, so $d^\perp = k+1$. Now use $d = n - k + 1$ in Lemma 1.2. □

Now that we know all the $l(J)$ for an MDS code, it is easy to find the weight distribution.

Theorem 1.8. *Let C be an MDS code with parameters $[n,k]$. Then the generalized weight distribution is given by*

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

The coefficients of the extended weight enumerator are given by

$$A_w(T) = \binom{n}{w} \sum_{j=0}^{w-d} (-1)^j \binom{w}{j} (T^{w-d+1-j} - 1).$$

Proof. We will give the construction for the generalized weight enumerator here: the case of the extended weight enumerator goes similar and is left as an exercise. We know from Proposition 1.29 that for an 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 and extended weight enumerators are not enough to distinguish between codes with the same parameters. We illustrate the non-equivalence of two MDS codes by an example.

Example 1.30. Let C be a linear $[n, 3]$ MDS code over $\mathbb{F}_q$. It is possible to write the generator matrix G of C in the following form:

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{pmatrix}.$$

Because C is MDS we have $d = n - 2$. We now view the n columns of G as points in the projective plane $\mathbb{P}^2(\mathbb{F}_q)$, say $P_1, \dots, P_n$. The MDS property that every k columns of G are independent is now equivalent with saying that no three points are on a line. To see that these n points do not always determine an equivalent code, consider the following construction. Through the n points there are $\binom{n}{2} = N$ lines, the set $\mathcal{N}$. These lines determine (the generator matrix of) a $[N, 3]$ code $\hat{C}$. The minimum distance of the code $\hat{C}$ is equal to the total number of lines minus the maximum number of lines from $\mathcal{N}$ through an arbitrarily point $P \in \mathbb{P}^2(\mathbb{F}_q)$ by Proposition 1.16. If

$P \notin \{P_1, \ldots, P_n\}$ then the maximum number of lines from $\mathcal{N}$ through P is at most $\frac{1}{2}n$, since no three points of $\mathcal{N}$ lie on a line. If $P = P_i$ for some $i \in [n]$ then P lies on exactly $n-1$ lines of $\mathcal{N}$, namely the lines P_iP_j for $j \neq i$. Therefore the minimum distance of $\hat{C}$ is $d = N - n + 1$.

We now have constructed a $[N, 3, N-n+1]$ code $\hat{C}$ from the original code C. Notice that two codes $\hat{C}_1$ and $\hat{C}_2$ are generalized equivalent if C_1 and C_2 are generalized equivalent. The generalized and extended weight enumerators of an MDS code of length n and dimension k are completely determined by the pair (n, k), but this is not generally true for the weight enumerator of $\hat{C}$.

Take for example $n = 6$ and $q = 9$, so $\hat{C}$ is a $[15, 3, 10]$ code. Look at the codes C_1 and C_2 generated by the following matrices respectively, where $\alpha \in \mathbb{F}_9$ is a primitive element:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & \alpha^5 & \alpha^6 \\ 0 & 0 & 1 & \alpha^3 & \alpha & \alpha^3 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & \alpha^7 & \alpha^4 & \alpha^6 \\ 0 & 0 & 1 & \alpha^5 & \alpha & 1 \end{pmatrix}$$

Being both MDS codes, the weight distribution is $(1, 0, 0, 120, 240, 368)$. If we now apply the above construction, we get $\hat{C}_1$ and $\hat{C}_2$ generated by

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 & \alpha^4 & \alpha^6 & \alpha^3 & \alpha^7 & \alpha & 1 & \alpha^2 & 1 & \alpha^7 & 1 \\ 0 & 1 & 0 & \alpha^7 & 1 & 0 & 0 & \alpha^4 & 1 & 1 & 0 & \alpha^6 & \alpha & 1 & \alpha^3 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & \alpha^7 & \alpha^2 & \alpha^3 & \alpha & 0 & \alpha^7 & \alpha^7 & \alpha^4 & \alpha^7 & \alpha & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & \alpha^3 & 0 & \alpha^6 & \alpha^6 & 0 & \alpha^7 & \alpha & \alpha^6 & \alpha^3 & \alpha \\ 0 & 0 & 1 & \alpha^5 & \alpha^5 & \alpha^6 & \alpha^3 & \alpha^7 & \alpha^4 & \alpha^3 & \alpha^5 & \alpha^2 & \alpha^4 & \alpha & \alpha^5 \end{pmatrix}$$

The weight distribution of $\hat{C}_1$ and $\hat{C}_2$ are, respectively,

$$(1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 48, 0, 16, 312, 288, 64) \text{ and}$$
$$(1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 48, 0, 32, 264, 336, 48).$$

So the latter two codes are not generalized equivalent, and therefore not all $[6, 3, 4]$ MDS codes over $\mathbb{F}_9$ are generalized equivalent.

Another example was given in [31, 32] showing that two $[6, 3, 4]$ MDS codes could have distinct covering radii.

1.6. Matroids and codes

Matroids were introduced by Whitney [33], axiomatizing and generalizing the concepts of "independence" in linear algebra and "cycle-free" in graph theory. In the theory of arrangements one uses the notion of a geometric lattice that will be treated in Section 1.7.2. In graph and coding theory one usually refers more to matroids. See [34–38] for basic facts of the theory of matroids.

1.6.1. *Matroids*

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

(I.1) $\emptyset \in \mathcal{I}$.
(I.2) If $J \subseteq I$ and $I \in \mathcal{I}$, then $J \in \mathcal{I}$.
(I.3) If $I, J \in \mathcal{I}$ and $|I| < |J|$, then there exists a $j \in (J \setminus I)$ such that $I \cup \{j\} \in \mathcal{I}$.

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

If J is a subset of E, 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 $|I_1| = |I_2|$ by condition (I.3). The *rank* or *dimension* of a subset J of E is the number of elements of a maximal independent subset of J. An independent set of rank $r(M)$ is called a *basis* of M. The collection of all bases of M is denoted by $\mathcal{B}$.

Let $M_1 = (E_1, \mathcal{I}_1)$ and $M_2 = (E_2, \mathcal{I}_2)$ be matroids. A map $\varphi : E_1 \to E_2$ 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 $\psi : E_2 \to E_1$ 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.

Example 1.31. Let n and k be nonnegative integers such that $k \leq n$. Let $\mathcal{I}_{n,k} = \{I \subseteq [n] : |I| \leq k\}$. Then $U_{n,k} = ([n], \mathcal{I}_{n,k})$ is a matroid that is called the *uniform matroid* of rank k on n elements. A subset B of $[n]$ is

a basis of $U_{n,k}$ if and only if $|B| = k$. The matroid $U_{n,n}$ has no dependent sets and is called *free*.

Let $(E, \mathcal{I})$ be a matroid. An element x in E is called a *loop* if $\{x\}$ is a dependent set. Let x and y in E 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. Now $U_{n,r}$ is the only simple matroid of rank r if $r \leq 2$.

Let G be a $k \times n$ matrix with entries in a field $\mathbb{F}$. Let E be the set $[n]$ indexing the columns of G and $\mathcal{I}_G$ be the collection of all subsets I of E such that the submatrix G_I consisting of the columns of G at the positions of I are independent. Then $M_G = (E, \mathcal{I}_G)$ 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 $(E, \mathcal{I}_{G_1}) = (E, \mathcal{I}_{G_2})$. So the matroid $M_C = (E, \mathcal{I}_C)$ of a code C is well defined by $(E, \mathcal{I}_G)$ for some generator matrix G of C. If C is degenerate, then there is a position i such that $c_i = 0$ for every codeword $\mathbf{c} \in C$ and all such 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. A $[n, k]$ code C is MDS if and only if the matroid M_C is the uniform matroid $U_{n,k}$.

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

For more on representable matroids we refer to Tutte [39] and Whittle [40, 41]. Let g_n be the number of isomorphism classes of simple matroids on n points. The values of g_n are determined for $n \leq 8$ by [42] and are given in the following table:

n	1	2	3	4	5	6	7	8
g_n	1	1	2	4	9	26	101	950

Extended tables can be found in [43]. Clearly $g_n \leq 2^{2^n}$. Asymptotically the number g_n is given in [44] and is as follows:

$$g_n \leq n - \log_2 n + \mathcal{O}(\log_2 \log_2 n),$$

$$g_n \geq n - \tfrac{3}{2} \log_2 n + \mathcal{O}(\log_2 \log_2 n).$$

A crude upper bound on the number of $k \times n$ matrices with $k \leq n$ and entries in $\mathbb{F}_q$ is given by $(n+1)q^{n^2}$. Hence the vast majority of all matroids on n elements is not representable over a given finite field for $n \to \infty$.

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

The dual matroid is indeed a matroid. Let C be a code over a finite field. Then the matroids $(M_C)^\perp$ and $M_{C^\perp}$ are isomorphic.

Let e be a loop of the matroid M. Then e is not a member of any basis of M. Hence e is in every basis of $M^\perp$. An element of M is called an *isthmus* if it is an element of every basis of M. Hence e is an isthmus of M if and only if e is a loop of $M^\perp$.

Proposition 1.30. *Let $(E, \mathcal{I})$ be a matroid with rank function r. Then the dual matroid has rank function $r^\perp$ given by*

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

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

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

□

1.6.2. *Graphs, codes and matroids*

Graph theory is regarded to start with the paper of Euler [45] with his solution of the problem of the Königbergs bridges. For an introduction to

the theory of graphs we refer to [46, 47].

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 is called *planar* if the there is an injective map $f : V \to \mathbb{R}^2$ from the set of vertices V to the real plane such that for every edge e with ends u and v there is a simple curve in the plane connecting the ends of the edge such that mutually distinct simple curves do not intersect except at the endpoints. More formally: for every edge e with ends u and v there is an injective continuous map $g_e : [0, 1] \to \mathbb{R}^2$ from the unit interval to the plane such that $\{f(u), f(v)\} = \{g_e(0), g_e(1)\}$, and $g_e(0, 1) \cap g_{e'}(0, 1) = \emptyset$ for all edges e, e' with $e \neq e'$.

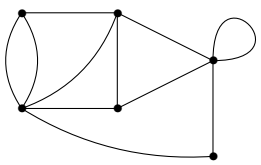


Fig. 1.7. A planar graph

Example 1.32. Consider the next riddle:

> Three new-build houses have to be connected to the three nearest terminals for gas, water and electricity. For security reasons, the connections are not allowed to cross. How can this be done?

The answer is "not", because the corresponding graph (see Figure 1.9) is not planar. This riddle is very suitable to occupy kids who like puzzles, but make sure to have an easy explainable proof of the improbability. We leave it to the reader to find one.

Let $\Gamma_1 = (V_1, E_1)$ and $\Gamma_2 = (V_2, E_2)$ be graphs. A map $\varphi : V_1 \to V_2$ is called a *morphism of graphs* if $\varphi(v)$ and $\varphi(w)$ are connected in Γ_2 for all

$v, w \in V_1$ that are connected in Γ_1. The map is called an *isomorphism of graphs* if it is a morphism of graphs and there exists a map $\psi : V_2 \to V_1$ such that it is a morphism of graphs and it is the inverse of φ. The graphs are called *isomorphic* if there is an isomorphism of graphs between them.

An edge of a graph is called an *isthmus* if the number of components of the graph increases by deleting the edge. It the graph is connected, then deleting an isthmus gives a graph that is no longer connected. Therefore an isthmus is also called a *bridge*. An edge is an isthmus if and only if it is in no cycle. Therefore an edge that is an isthmus is also called an *acyclic* edge.

By deleting loops and parallel edges from a graph Γ one gets a simple graph. There is a choice in the process of deleting parallel edges, but the resulting graphs are all isomorphic. We call this simple graph the *simplification* of the graph and it is denoted by $\bar{\Gamma}$.

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 a map $\gamma : V \to K$ such that $\gamma(u) \neq \gamma(v)$ for all distinct adjacent vertices u and v in V. So vertex u has color $\gamma(u)$ and all other adjacent vertices have a color distinct from $\gamma(u)$. Let $P_\Gamma(k)$ be the number of k-colorings of Γ. Then P_Γ is called the *chromatic polynomial* of Γ.

If the graph Γ has no edges, then $P_\Gamma(k) = k^v$ where $|V| = v$ and $|K| = k$, since it is equal to the number of all maps from V to K. In particular there is no map from V to an empty set in case V is nonempty. So the number of 0-colorings is zero for every graph.

The number of colorings of graphs was studied by Birkhoff [48], Whitney [49, 50] and Tutte [51–55]. Much research on the chromatic polynomial was motivated by the four-color problem of planar graphs.

Let K_n be the *complete graph* on n vertices in which every pair of two distinct vertices is connected by exactly one edge. Then there is no k coloring if $k < n$. Now let $k \geq n$. Take an enumeration of the vertices. Then there are k possible choices of a color of the first vertex and $k - 1$ choices for the second vertex, since the first and second vertex are connected. Now suppose by induction that we have a coloring of the first i vertices, then there are

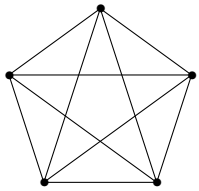


Fig. 1.8. The complete graph K_5

$k-i$ possibilities to color the next vertex, since the $(i+1)$-th vertex is connected to the first i vertices. Hence

$$P_{K_n}(k) = k(k-1)\cdots(k-n+1)$$

So $P_{K_n}(k)$ is a polynomial in k of degree n.

Proposition 1.31. *Let $\Gamma = (V, E)$ be a graph. Then $P_\Gamma(k)$ is a polynomial in k.*

Proof. See [48]. Let $\gamma : V \to K$ be a k-coloring of Γ with exactly i colors. Let σ be a permutation of K. Then the composition of maps $\sigma \circ \gamma$ is also k-coloring of Γ with exactly i colors. Two such colorings are called equivalent. Then $k(k-1)\cdots(k-i+1)$ is the number of colorings in the equivalence class of a given k-coloring of Γ with exactly i colors. Let m_i be the number of equivalence classes of colorings with exactly i colors of the set K. Let $v = |V|$. Then $P_\Gamma(k)$ is equal to

$$m_1k+m_2k(k-1)+\ldots+m_ik(k-1)\cdots(k-i+1)+\ldots+m_vk(k-1)\cdots(k-v+1).$$

Therefore $P_\Gamma(k)$ is a polynomial in k. □

A graph $\Gamma = (V, E)$ is called *bipartite* if V is the disjoint union of two nonempty sets M and N such that the ends of an edge are in M and in N. Hence no two points in M are adjacent and no two points in N are adjacent. Let m and n be integers such that $1 \leq m \leq n$. The *complete bipartite graph* $K_{m,n}$ is the graph on a set of vertices V that is the disjoint union of two sets M and N with $|M| = m$ and $|N| = n$, and such that every vertex in M is connected with every vertex in N by a unique edge.
Another tool to show that $P_\Gamma(k)$ is a polynomial this by deletion-contraction of graphs, a process which is similar to the puncturing and shortening of codes from Section 1.5.3.

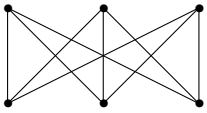


Fig. 1.9. The complete bipartite graph $K_{3,3}$

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 and deleting e. Formally this is defined as follows. Let $\tilde{u} = \tilde{v} = \{u, v\}$, and $\tilde{w} = \{w\}$ if $w \neq u$ and $w \neq v$. Let $\tilde{V} = \{\tilde{w} : w \in V\}$. Then Γ/e is the graph $(\tilde{V}, E \setminus \{e\})$, where an edge $f \neq e$ is incident with $\tilde{w}$ in Γ/e if f is incident with w in Γ.

Notice that the number of k-colorings of Γ does not change by deleting loops and a parallel edge. Hence the chromatic polynomial Γ and its simplification $\bar{\Gamma}$ are the same.

The following proposition is due to Foster. See the concluding note in [50].

Proposition 1.32. *Let $\Gamma = (V, E)$ be a simple 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. Let u and v be the vertices of e. Then $u \neq v$, since the graph is simple. Let γ be a k-coloring of $\Gamma \setminus e$. Then γ is also a coloring of Γ if and only if $\gamma(u) \neq \gamma(v)$. If $\gamma(u) = \gamma(v)$, then consider the induced map $\tilde{\gamma}$ on $\tilde{V}$ defined by $\tilde{\gamma}(\tilde{u}) = \gamma(u)$ and $\tilde{\gamma}(\tilde{w}) = \gamma(w)$ if $w \neq u$ and $w \neq v$. The map $\tilde{\gamma}$ gives a k-coloring of Γ/e. Conversely, every k-coloring of Γ/e gives a k-coloring γ of $\Gamma \setminus e$ such that $\gamma(v) = \gamma(w)$. Therefore

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

This follows also from a more general deletion-contraction formula for matroids that will be treated in Section 1.6.4 and Proposition 1.8.1. □

Let $\Gamma = (V, E)$ be a graph. Suppose that $V' \subseteq V$ and $E' \subseteq E$ and all the endpoints of e' in E' are in V'. Then $\Gamma' = (V', E')$ is a graph and it is

called a *subgraph* of Γ.

Two vertices u to v are *connected* by a *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 a $(t-1)$-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 $\gamma(\Gamma)$.

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 and set of edges E is a disjoint union of subsets E_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 $c(\Gamma)$.

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 a $m \times n$ 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 with $i < k$. An arrangement $\mathcal{A}$ is called *graphic* if $\mathcal{A}$ is isomorphic with $\mathcal{A}_\Gamma$ for some graph Γ.

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

Let Γ be a finite graph without loops. Then the arrangement $\mathcal{A}_\Gamma$ is isomorphic with $\mathcal{A}_{C_\Gamma}$.

Proposition 1.33. *Let Γ be a finite graph. Then C_Γ is a code with parameters $[n, k, d]$, where $n = |E|$, $k = |E| - |V| + c(\Gamma)$ and $d = \gamma(\Gamma)$.*

Proof. See [46, Prop. 4.3] □

Let $M = (E, \mathcal{I})$ be a matroid. A subset C of E is called a *circuit* if it is dependent and all its proper subsets are independent. A circuit of the dual matroid $M^\perp$ is called a *cocircuit* of M.

Proposition 1.34. *Let $\mathcal{C}$ be the collection of circuits of a matroid. Then*

(C.1) $\emptyset \notin \mathcal{C}$.
(C.2) If $C_1, C_2 \in \mathcal{C}$ *and* $C_1 \subseteq C_2$, *then* $C_1 = C_2$.
(C.3) 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 [35, Lemma 1.1.3]. □

Condition (C.3) is called the *circuit elimination axiom*. The converse of Proposition 1.34 holds.

Proposition 1.35. *Let $\mathcal{C}$ be a collection of subsets of a finite set E that satisfies the conditions (C.1), (C.2) and (C.3). Let $\mathcal{I}$ be the collection of all subsets of E that contain no member of $\mathcal{C}$. Then $(E, \mathcal{I})$ is a matroid with $\mathcal{C}$ as its collection of circuits.*

Proof. See [35, Theorem 1.1.4]. □

Proposition 1.36. *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_Γ on E. This matroid is called the* cycle matroid *of Γ.*

Proof. See [35, Proposition 1.1.7]. □

Loops in Γ correspond one-to-one to loops in M_Γ. Two edges that are no loops, are parallel in Γ if and only if they are parallel in M_Γ. So Γ is simple if and only if M_Γ is simple. Let e in E. Then e is an isthmus in the graph Γ if and only is e is an isthmus in the matroid M_Γ.

A matroid M is called *graphic* if M is isomorphic with M_Γ for some graph Γ, and it is called *cographic* if $M^\perp$ is graphic. If Γ is a planar graph, then the matroid M_Γ is graphic by definition but it is also cographic.

Let Γ be a finite graph with incidence matrix $I(\Gamma)$. This is a generator matrix for C_Γ over a field $\mathbb{F}$. Suppose that $\mathbb{F}$ is the binary field. Look at all the

columns indexed by the edges of a cycle of Γ. Since every vertex in a cycle is incident with exactly two edges, the sum of these columns is zero and therefore they are dependent. Removing a column gives an independent set of vectors. Hence the cycles in the matroid M_{C_Γ} coincide with the cycles in Γ. Therefore M_Γ is isomorphic with M_{C_Γ}. One can generalize this argument for any field. Hence graphic matroids are representable over any field.

The matroids of the binary Hamming $[7, 4, 3]$ code is not graphic and not cographic. Clearly the matroids M_{K_5} and $M_{K_{3,3}}$ are graphic by definition, but they are not cographic. Tutte [56] found a classification for graphic matroids.

1.6.3. *The weight enumerator and the Tutte polynomial*

See [1, 26, 57–63] for references of this section.

Definition 1.3. Let $M = (E, \mathcal{I})$ be a matroid. Then the *Whitney rank generating function* $R_M(X, Y)$ is defined by

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

and the *Tutte-Whitney* or *Tutte polynomial* by

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

In other words,

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

Whitney [50] defined the coefficients m_{ij} of the polynomial $R_M(X, Y)$ such that

$$R_M(X, Y) = \sum_{i=0}^{r(M)} \sum_{j=0}^{|M|} m_{ij} X^i Y^j,$$

but he did not define the polynomial $R_M(X, Y)$ as such. It is clear that these coefficients are nonnegative, since they count the number of elements of certain sets. The coefficients of the Tutte polynomial are also nonnegative, but this is not a trivial fact, it follows from the counting of certain *internal* and *external* bases of a matroid. See [64].

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.

Proposition 1.37. *Let C be a $[n,k]$ code over $\mathbb{F}_q$. Then the Tutte polynomial t_C associated with the matroid M_C of the code C is*

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

Proof. This follows from $l(J) = k - r(J)$ by Lemma 1.1 and $r(M) = k$. □

This formula and Proposition 1.24 suggest the next connection between the weight enumerator and the Tutte polynomial. Greene [26] was the first to notice this connection.

Theorem 1.9. *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:*

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

Proof. By using Proposition 1.37 about the Tutte polynomial, rewriting, and Proposition 1.24 we get

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

□

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

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

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

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

□

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 [26] 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 [61] and independently by Jurrius [1].

Theorem 1.11. *For the generalized weight enumerator of a $[n,k]$ code C and the associated Tutte polynomial we have that $W_C^{(r)}(X,Y)$ is equal to*

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

And, conversely,

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

Proof. For the first formula, use Theorems 1.7 and 1.9. Use Theorems 1.6 and 1.10 for the second formula. □

1.6.4. *Deletion and contraction of matroids*

Let $M = (E, \mathcal{I})$ be a matroid of rank k. Let e be an element of E. Then the *deletion* $M \setminus e$ is the matroid on the set $E \setminus \{e\}$ with independent sets of the form $I \setminus \{e\}$ where I is independent in M. The *contraction* M/e is the matroid on the set $E \setminus \{e\}$ with independent sets of the form $I \setminus \{e\}$ where I is independent in M and $e \in I$.

Let C be a code with reduced generator matrix G at position e. So $a = (1, 0, \ldots, 0)^T$ is the column of G at position e. Then $M \setminus e = M_{G\setminus a}$ and $M/e = M_{G/a}$. A puncturing-shortening formula for the extended weight enumerator is given in Proposition 1.26. By virtue of the fact that the extended weight enumerator and the Tutte polynomial of a code determine each other by the Theorems 1.9 and 1.10, one expects that an analogous generalization for the Tutte polynomial of matroids holds.

Proposition 1.38. *Let $M = (E, \mathcal{I})$ be a matroid. Let $e \in E$ that is not a loop and not an isthmus. Then the following* deletion-contraction formula *holds:*

$$t_M(X,Y) = t_{M\setminus e}(X,Y) + t_{M/e}(X,Y).$$

Proof. See [25, 53, 65, 66]. □

Let M be a graphic matroid. So $M = M_\Gamma$ for some finite graph Γ. Let e be an edge of Γ, then $M \setminus e = M_{\Gamma\setminus e}$ and $M/e = M_{\Gamma/e}$.

1.6.5. *MacWilliams type property for duality*

For both codes and matroids we defined the dual structure. These objects obviously completely define there dual. But how about the various polynomials associated to a code and a matroid? We know from Example 1.30 that 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 1.2. We will give a proof of this relation by considering the more general question for the extended weight enumerator. We will prove the MacWilliams identities using the Tutte polynomial. We do this because of the following simple and very useful relation between the Tutte polynomial of a matroid and its dual.

Theorem 1.12. *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. Let M be a matroid on the set E. Then $M^\perp$ is a matroid on the same set. In Proposition 1.30 we proved $r^\perp(J) = |J| - r(E) + r(E \setminus J)$. In particular, we have $r^\perp(E) + r(E) = |E|$. Substituting this relation into the definition of the Tutte polynomial for the dual code, gives

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

In the last step, we use that the summation over all $J \subseteq E$ is the same as a summation over all $E \setminus J \subseteq E$. 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 [26] was the first to use this idea, see also Brylawsky and Oxley [67].

Theorem 1.13 (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 the matroid associated to the code. Using the previous theorem and the relation between the weight enumerator and the Tutte polynomial, we find

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

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

We can use the relations in Theorems 1.6 and 1.7 to prove the MacWilliams identities for the generalized weight enumerator.

Theorem 1.14. *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.

$$\begin{aligned}
W^{(r)}_{C^\perp}(X,Y) &= \frac{1}{\langle r\rangle_q}\sum_{j=0}^{r}\begin{bmatrix} r\\ j\end{bmatrix}_q(-1)^{r-j}q^{\binom{r-j}{2}}\ W_{C^\perp}(X,Y,q^i)\\
&= \sum_{j=0}^{r}(-1)^{r-j}\frac{q^{\binom{r-j}{2}-j(r-j)}}{\langle j\rangle_q\langle r-j\rangle_q}q^{-jk}W_c(X+(q^j-1)Y, X-Y, q^j)\\
&= \sum_{j=0}^{r}(-1)^{r-j}\frac{q^{\binom{r-j}{2}-j(r-j)-jk}}{\langle j\rangle_q\langle r-j\rangle_q}\\
&\quad\times\sum_{l=0}^{j}\frac{\langle j\rangle_q}{q^{l(j-l)}\langle j-l\rangle_q}W^{(l)}_C(X+(q^j-1)Y, X-Y)\\
&= \sum_{j=0}^{r}\sum_{l=0}^{j}(-1)^{r-j}\frac{q^{\binom{r-j}{2}-j(r-j)-l(j-l)-jk}}{\langle r-j\rangle_q\langle j-l\rangle_q}\\
&\quad\times W^{(l)}_C(X+(q^j-1)Y, X-Y).
\end{aligned}$$

□

This theorem was proved by Kløve [68], 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.

1.7. Posets and lattices

In this section we consider the theory of posets and lattices and the Möbius function. Geometric lattices are defined and its connection with matroids is given. See [30, 69–72].

1.7.1. *Posets, the Möbius function and lattices*

Let L be a set and $\leq$ a relation on L such that:

(PO.1) $x \leq x$, for all x in L *(reflexive).*
(PO.2) If $x \leq y$ and $y \leq x$, then $x = y$, for all $x, y \in L$ *(anti-symmetric).*
(PO.3) If $x \leq y$ and $y \leq z$, then $x \leq z$, for all x, y and z in L *(transitive).*

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 $[x, y] = L_x \cap L^y$.

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 $r \geq 0$ be an integer. 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 integer $r \geq 0$.

Proposition 1.39. *Let L be a locally finite poset. Let $x \leq y$ in L. Then*

(N.1) $c_0(x, y) = 0$ *if x and y are not comparable.*
(N.2) $c_0(x, x) = 1$, $c_r(x, x) = 0$ *for all $r > 0$ and* $c_0(x, y) = 0$ *if $x < y$.*
(N.3) $c_{r+1}(x, y) = \sum_{x \leq z < y} c_r(x, z) = \sum_{x < z \leq y} c_r(z, y)$.

Proof. Statements (N.1) and (N.2) are 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 1.4. 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 1.40. *Let L be a locally finite poset. Then for all $x, y \in L$:*

(M.1) $\mu(x,y) = 0$ *if x and y are not comparable.*
(M.2) $\mu(x,x) = 1$.
(M.3) If $x < y$, then $\sum_{x \leq z \leq y} \mu(x,z) = \sum_{x \leq z \leq y} \mu(z,y) = 0$.
(M.4) If $x < y$, then $\mu(x,y) = -\sum_{x \leq z < y} \mu(x,z) = -\sum_{x < z \leq y} \mu(z,y)$.

Proof.
(M.1) and (M.2) follow from (N.1) and (N.2), respectively, of Proposition 1.39. (M.3) is clearly equivalent with (M.4). If $x < y$, then $c_0(x,y) = 0$. So

$$\begin{aligned}
\mu(x,y) &= \sum_{r=1}^{\infty} (-1)^r c_r(x,y) \\
&= \sum_{r=0}^{\infty} (-1)^{r+1} c_{r+1}(x,y) \\
&= -\sum_{r=0}^{\infty} (-1)^r \sum_{x \leq z < y} c_r(x,z) \\
&= -\sum_{x \leq z < y} \sum_{r=0}^{\infty} (-1)^r c_r(x,z) \\
&= -\sum_{x \leq z < y} \mu(x,z).
\end{aligned}$$

The first and last equality use the definition of μ. The second equality starts counting at $r = 0$ instead of $r = 1$, the third uses (N.3) of Proposition 1.39 and in the fourth the order of summation is interchanged. □

Remark 1.5. (M.2) and (M.4) of Proposition 1.40 can be used as an alternative way to compute $\mu(x,y)$ by induction.

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.

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

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

Define similarly the sum function $\check{f}$ of f by $\check{f}(x) = \sum_{x \leq y} f(y)$ if L is a locally finite poset with a maximum element.

A poset L is locally finite if and only if $[x, y]$ is finite for all $x \leq y$ in L. So $[0, x]$ is finite if L is a locally finite poset with minimum element 0. Hence the sum function $\hat{f}(x)$ is well-defined, since it is a finite sum of $f(y)$ in A with y in $[0, x]$. In the same way $\check{f}(x)$ is well-defined, since $[x, 1]$ is finite.

Theorem 1.15 (Möbius inversion formula). *Let L be a locally finite poset with a minimum element. Then*

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

Similarly $f(x) = \sum_{x \leq y} \mu(x, y) \check{f}(y)$ if L is a locally finite poset with a maximum element.

Proof. Let x be an element of L. Then

$$\begin{aligned}
\sum_{y \leq x} \mu(y, x) \hat{f}(y) &= \sum_{y \leq x} \sum_{z \leq y} \mu(y, x) f(z) \\
&= \sum_{z \leq x} f(z) \sum_{z \leq y \leq x} \mu(y, x) \\
&= f(x)\mu(x, x) + \sum_{z < x} f(z) \sum_{z \leq y \leq x} \mu(y, x) \\
&= f(x)
\end{aligned}$$

The first equality uses the definition of $\hat{f}(y)$. In the second equality the order of summation is interchanged. In the third equality the first summation is split in the parts $z = x$ and $z < x$, respectively. Finally $\mu(x, x) = 1$ and the second summation is zero for all $z < x$, by Proposition 1.40.
The proof of the second equality is similar. □

Example 1.33. Let $f(x) = 1$ if $x = 0$ and $f(x) = 0$ otherwise. Then the sum function $\hat{f}(x) = \sum_{y \leq x} f(y)$ 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 1.40.

Remark 1.6. 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. Suppose that $(L, \leq)$ is 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 $\mu_R(x, y) = \mu(y, x)$. If $(L, \leq)$ has minimum 0_L or maximum 1_L, then $(L, \leq_R)$ has minimum 1_L or maximum 0_L, respectively.

Definition 1.5. 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 1.34. Let $L = \mathbb{Z}$ be the set of integers with the usual linear order. The Hasse diagram of this poset looks as follows:

$$\ldots \longrightarrow n+1 \longrightarrow n \longrightarrow n-1 \longrightarrow \ldots \longrightarrow 1 \longrightarrow 0 \longrightarrow -1 \longrightarrow \ldots$$

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 $y > x + 1$.

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 1.7. 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. The meet of all the elements in this collection is well defined and is given by

$$x \vee y = \bigwedge \{z : x \leq z, y \leq z\}.$$

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

Example 1.35. Let L be the collection of all finite subsets of a given set $\mathcal{X}$. Let $\leq$ be defined by the inclusion, that means $I \leq J$ if and only if $I \subseteq J$. Then $0_L = \emptyset$, and L has a maximum if and only if $\mathcal{X}$ is finite in which case $1_L = \mathcal{X}$. For $\mathcal{X} = \{a, b, c, d\}$ the Hasse diagram of the poset is given in Figure 1.10.

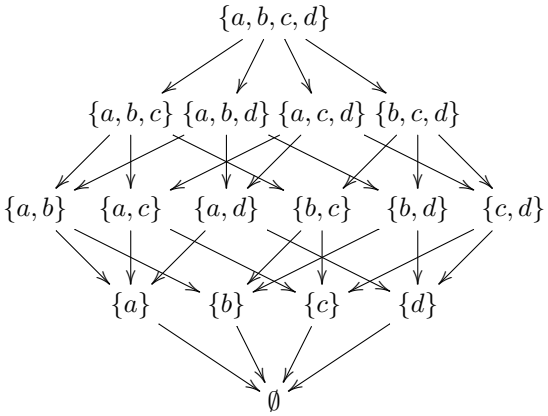


Fig. 1.10. The Hasse diagram of the poset of all subsets of $\{a, b, c, d\}$

Let $I, J \in L$ and $I \leq J$. Then $|I| \leq |J| < \infty$. Let $m = |J| - |I|$. Then

$$c_r(I,J) = \sum_{m_1 < m_2 < \cdots < m_{r-1} < m} \binom{m_2}{m_1}\binom{m_3}{m_2} \cdots \binom{m}{m_{r-1}}.$$

Hence L is locally finite. L is finite if and only if $\mathcal{X}$ is finite. Furthermore $I \vee J = I \cup J$ and $I \wedge J = I \cap J$. So L is a lattice. Using Remark 1.5 we see that $\mu(I,J) = (-1)^{|J|-|I|}$ if $I \leq J$. This is much easier than computing $\mu(I,J)$ by means of Definition 1.4.

Example 1.36. Let $\mathcal{X} = [n]$. Let k be an integer between 0 and n. Let $L_k = \{\mathcal{X}\}$ and L_i be the collection of all subsets of $\mathcal{X}$ of size i for all $i < k$. Let the partial order be given by the inclusion. Then L is a poset and $\mu(I,J) = (-1)^{|J|-|I|}$ if $I \leq J$ and $|J| < k$ as in Example 1.35, and $\mu(I,\mathcal{X}) = -\sum_{I \leq J < \mathcal{X}} (-1)^{|J|-|I|}$ for all $I < \mathcal{X}$ by Proposition 1.40.

Example 1.37. Now suppose again that $\mathcal{X} = [n]$. 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 J of $\mathcal{X}$

$$A_J = \bigcap_{j \in J} A_j \quad \text{and} \quad f(J) = |A_J \setminus \left(\bigcup_{I < J} A_I \right)|.$$

Then A_J is the disjoint union of the subsets $A_I \setminus (\bigcup_{K<I} A_K)$ for all $I \leq J$. Hence the sum function is equal to

$$\hat{f}(J) = \sum_{I \leq J} f(I) = \sum_{I \leq J} |A_I \setminus \left(\bigcup_{K < I} A_K \right)| = |A_J|.$$

Möbius inversion gives that

$$|A_J \setminus \left(\bigcup_{I < J} A_I \right)| = \sum_{I \leq J} (-1)^{|J|-|I|} |A_I|,$$

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

Example 1.38. A variant of the principle of inclusion/exclusion is given as follows. Let $H_1, \ldots, H_n$ be a collection of subsets of a finite set H. Let L be the poset of all intersections of the H_j with the reverse inclusion as partial order. Then H is the minimum of L and $H_1 \cap \cdots \cap H_n$ is the maximum of L. Let $x \in L$. Define

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

Then

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

Hence

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

Example 1.39. 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 $m \vee n = \text{lcm}(m,n)$ and $m \wedge n = \gcd(m,n)$. Hence L is a lattice. By Remark 1.5 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*. Define

$$V_d = \{i \in [n] : \gcd(i,n) = \tfrac{n}{d}\}$$

for $d|n$. Then

$$\{\ i \cdot \tfrac{n}{d} : i \in [d]\ , \gcd(i,d) = 1\ \} = V_d$$

so $|V_d| = \varphi(d)$. Now $[n]$ is the disjoint union of the subsets V_d with $d|n$. Hence the sum function of $\varphi(n)$ is given by

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

Therefore by Möbius inversion

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

Example 1.40. Consider the poset L of Example 1.39 with the divisibility as partial order. Let $\text{Irr}_q(n)$ be the number of irreducible monic polynomials over $\mathbb{F}_q$ of degree n. Define $f(d) = d \cdot \text{Irr}_q(d)$. Then the sum function

$\hat{f}(n) = \sum_{d|n} f(d)$ is equal to q^n. See [73, Corollary 3.21]. The Möbius inversion formula of Theorem 1.15 implies that

$$\mathrm{Irr}_q(n) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d.$$

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.

If $\varphi : L_1 \to L_2$ is an isomorphism between locally finite posets with a minimum, then $\mu_2(\varphi(x), \varphi(y)) = \mu_1(x, y)$ for all x, y in L_1. 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 1.41. 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 1.39. Let L_2 be the collection of all subsets of $[r]$ with the inclusion as partial order $\leq_2$ as in Example 1.35. 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.

1.7.2. *Geometric lattices*

Let $(L, \leq)$ be a lattice without infinite chains. Then L has a minimum and a maximum. 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 for 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 $x, y, z \in L$ 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 1.42. Let L be the collection of all finite subsets of a given set $\mathcal{X}$ as in Example 1.35. 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 1.43. Let L be the set of positive integers with the divisibility relation as in Example 1.39. 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.

Let L be a geometric lattice. Let $x, y \in L$ and $x \leq y$. The chain $x = y_0 < y_1 < \cdots < y_s = y$ from x to y is called an *extension* of the chain $x = x_0 < x_1 < \cdots < x_r = y$ if $\{x_0, x_1, \ldots, x_r\}$ is a subset of $\{y_0, y_1, \ldots, y_s\}$. A chain from x to y is called *maximal* if there is no extension to a longer chain from x to y.

Proposition 1.41. *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) *If $x \leq y$, then every chain from x to y can be extended to a maximal chain with the same end points, and all such maximal chains have the same length $r(y) - r(x)$.* (Jordan-Hölder property).

Proof. See [30, Prop. 3.3.2] and [72, Prop. 3.7]. □

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* holds for all $x, y \in L$:

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

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 1.5 is the directed graph with an arrow from y to x if $x, y \in L$,

$x < y$ and $r(y) = r(x) + 1$.

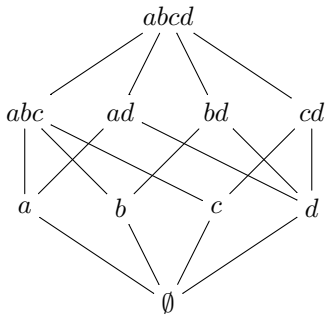


Fig. 1.11. The Hasse diagram of the geometric lattice in Example 1.48

Let L be a geometric lattice. Then L_x is a geometric lattice with x as minimum element and 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 1.44. 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. Let x and y be linear subspaces of V. Then $x \cap y$ the intersection of x and y is the largest linear subspace that is contained in x and y. So $x \wedge y = x \cap y$. The sum $x + y$ of x and y is by definition the set of elements $a + b$ with a in x and b in y. Then $x + y$ is the smallest linear subspace containing both x and y. Hence $x \vee y = x + y$. 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.

Example 1.45. 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

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

by Remark 1.7. 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 1.46. 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 element and $\{0\}$ is the maximum element. Furthermore

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

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.

Example 1.47. Let G be a generator matrix of the simplex code $\mathcal{S}_r(q)$ of Example 1.22. Let $\mathcal{A} = \mathcal{A}_G$ be the arrangement of the matrix G. Then the projective hyperplanes of the arrangement of $\mathcal{A}$ are all the $(q^r - 1)/(q - 1)$ hyperplanes of $\mathbb{P}^{r-1}(\mathbb{F}_q)$. The geometric lattice $L(\mathcal{A})$ consists of all possible intersections of these hyperplanes, so they are all projective subspaces of $\mathbb{P}^{r-1}(\mathbb{F}_q)$ with the reverse inclusion as partial order. This geometric order is self dual, that means that it is isomorphic under map that sends points to hyperplanes with the inclusion and the reverse inclusion as partial orders. In this way we see that the geometric lattice of the simplex code $\mathcal{S}_r(q)$ is isomorphic with the geometric lattice of all linear subspaces of a given vector space V of Example 1.44 with $\mathbb{F} = \mathbb{F}_q$ and $V = \mathbb{F}_q^r$.

Example 1.48. Consider the following matrix over a field $\mathbb{F}$.

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

Denote the columns of the matrix by a, b, c, d respectively. We can interpret the columns either as vectors in $\mathbb{F}^3$, or as the coefficients of hyperplanes in $\mathbb{F}^3$. The corresponding lattices will be isomorphic by Examples 1.45 and 1.46. The corresponding Hasse diagram is given in Figure 1.11.

1.7.3. *Geometric lattices and matroids*

The notion of a geometric lattice is "*cryptomorphic*" , that is almost equivalent to the concept of a matroid. See [35, 63, 69, 72, 74].

Proposition 1.42. *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. □

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.

Let $M = (E, \mathcal{I})$ be a matroid. A *k-flat* of M is a maximal subset of E of rank k. Let $L(M)$ be the collection of all flats of M, it is called the *lattice of flats* of M. Let J be a subset of E. Then the *closure* $\bar{J}$ is by definition the intersection of all flats that contain J. Flats of size $k-1$ are sometimes called *hyperplanes*, but in this chapter we will avoid this terminology.

The whole set E 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 E 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 1.7. Let J be a subset of E, then $\bar{J}$ is a flat, since it is a nonempty, finite intersection of flats. So $\bar{\emptyset}$ is the minimum of $L(M)$.

An element x in E is a loop if and only if $\bar{x} = \bar{\emptyset}$. If $x, y \in E$ are not loops, then x and y are parallel if and only if $\bar{x} = \bar{y}$. Let $\bar{E} = \{\bar{x} : x \in M, \bar{x} \neq \bar{\emptyset}\}$. Let $\bar{\mathcal{I}} = \{\bar{I} : I \in \mathcal{I}, \bar{\emptyset} \notin \bar{I}\}$. Then $\bar{M} = (\bar{E}, \bar{\mathcal{I}})$ is a simple matroid.

Let G be a generator matrix of a code C. The *simplified 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 *simplified code* $\bar{C}$ of C is the code with generator matrix $\bar{G}$.

Remark 1.8. Let G be a generator matrix of a code C. The definition of the simplified code $\bar{C}$ by means of $\bar{G}$ does not depend on the choice of the generator matrix G of C. The matroids $\bar{M}_G$ and $M_{\bar{G}}$ are isomorphic.
Let J be a subset of $[n]$. Then the closure $\bar{J}$ is equal to the complement in $[n]$ of the support of $C(J)$ and $C(J) = C(\bar{J})$.

Proposition 1.43. *Let M 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. See [75] and [72, Theorem 3.8]. □

Example 1.49. The geometric lattice of the matroid $U_{n,k}$ is isomorphic with the lattice consisting of $[n]$ and all its subsets of size at most $k-1$ of Example 1.36.

1.8. The characteristic polynomial

The characteristic polynomial of geometric lattices is defined. This is generalized in two variable polynomials in two ways: the coboundary polynomial and the Möbius polynomial. For simple matroids and codes the coboundary polynomial is equivalent with the Tutte polynomial and the extended weight enumerator. The Möbius polynomial contains information on the number of minimal subcodes and codewords. The coboundary and Möbius polynomial do not determine each other. This will be shown by examples of three dimensional codes.

1.8.1. *The characteristic and coboundary polynomial*

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

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

The *two variable characteristic polynomial* or *coboundary polynomial* is defined by

$$\chi_L(S,T) = \sum_{x \in L} \sum_{x \leq y \in L} \mu(x,y) S^{a(x)} T^{r(L)-r(y)},$$

where $a(x)$ is the number of atoms a in L such that $a \leq x$.

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

Remark 1.9. Let n be the number of atoms of L. Then the following relation holds for the coboundary polynomial in terms of characteristic polynomials:

$$\chi_L(S,T) = \sum_{i=0}^{n} S^i \chi_i(T) \quad \text{with} \quad \chi_i(T) = \sum_{\substack{x \in L \\ a(x)=i}} \chi_{L_x}(T).$$

$\chi_i(T)$ is called the *i-defect* polynomial . See [63, 76].

Example 1.50. Let L be the lattice of all subsets of a given finite set of r elements as in Examples 1.35 and 1.42. Then $r(x) = a(x)$ and $\mu(x,y) = (-1)^{a(y)-a(x)}$ if $x \leq y$. Hence

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

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

Example 1.51. Let L be the lattice of the simplex code, that is of all linear subspaces of a given vector space of dimension r over the finite field $\mathbb{F}_q$ as in Example 1.47. Then the number of atoms of L is $n = \frac{q^r-1}{q-1}$ and $r(x)$ is the dimension of x over $\mathbb{F}_q$. The number of subspaces of dimension i is counted in Remark 1.4. 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

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

See [19] and the proof of Theorem 1.7. More generally if $0 \leq i \leq \frac{q^r-1}{q-1}$ and $0 \leq j \leq r$, then

$$\chi_i(T) = \begin{cases} 0 & \text{if } \frac{q^{j-1}-1}{q-1} < i < \frac{q^j-1}{q-1}, \\ 1 & \text{if } i = \frac{q^r-1}{q-1}, \\ \begin{bmatrix} r \\ j \end{bmatrix}_q \prod_{l=0}^{r-j-1}(T-q^l) & \text{if } i = \frac{q^j-1}{q-1}, j < r. \end{cases}$$

See [77, Prop. 3.3].

Proposition 1.44 (Rota's Crosscut Theorem). *Let L be a finite geometric lattice. Let $M(L) = (E, \mathcal{I})$ be the matroid associated with L. Then*

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

Proof. See [71] and [78, Theorem 3.1]. □

As a consequence of Proposition 1.44 we have the following way to describe the characteristic polynomial of L in terms of the Tutte polynomial of $M(L)$:

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

Theorem 1.16. *The two variable characteristic or coboundary polynomial $\chi_L(S,T)$ of a finite geometric lattice L is related to the Whitney rank generating function of $M(L)$ by the formula*

$$\chi_L(S,T) = (S-1)^{r(L)} R_{M(L)}\left(\frac{T}{S-1}, S-1\right).$$

Proof. See [79, p. 605] and [76, Theorem II]. □

Remark 1.10. Because of Theorem 1.16 we have the following relations between $t_{M(L)}(X,Y)$ and $\chi_L(S,T)$:

$$\chi_L(S,T) = (S-1)^{r(L)} t_{M(L)}\left(\frac{S+T-1}{S-1}, S\right)$$

and, vice versa,

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

Therefore the polynomials $\chi_L(S,T)$ and $t_{M(L)}(X,Y)$ completely determine each other.

Starting with an arbitrary matroid M one has the associated geometric lattice $L(M)$, but $M(L(M))$ is isomorphic with M if and only if M is simple by Proposition 1.43. Therefore $t_M(X,Y)$ and $\chi_{L(M)}(S,T)$ completely determine each other if M is simple, but $t_M(X,Y)$ is a stronger invariant than $\chi_{L(M)}(S,T)$ if M is not simple. We will see a counterexample in Example 1.59

Remark 1.11. The relation between $t_{M(L)}(X,Y)$ and $\chi_L(S,T)$ shows great similarity with the formula in Theorem 1.10. Combining the relations we find that for projective codes

$$\chi_i(T) = A_{n-i}(T)$$

for all $0 \leq i \leq n$.

Let $\Gamma = (V,E)$ be a finite simple graph. Let χ_Γ be the characteristic polynomial of the geometric lattice $L(M_\Gamma)$. Then for all positive integers k, $P_\Gamma(k) = \chi_\Gamma(k)$. So the chromatic polynomial of a graph is the prime example of a characteristic polynomial and the two variable characteristic polynomial of a graph is also called the *dichromatic* polynomial of the graph. See [51, 53, 65].

Let γ be a coloring of Γ. Then an edge is called *bad* if if joins two vertices with the same color. The i-defect polynomial $\chi_i(T)$ counts up to a factor of T the number of ways of coloring Γ with i bad edges. See [63, §6.3.F].

1.8.2. *The Möbius polynomial and Whitney numbers*

Let L be a finite geometric lattice. The *two variable Möbius polynomial* $\mu_L(S,T)$ in S and T is defined by

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

Now $\mu(L) = \chi_L(0)$ and $\mu_L(0,T) = \chi_L(0,T) = \chi_L(T)$.

Remark 1.12. Let r be the rank of L. Then the following relation holds for the Möbius polynomial in terms of characteristic polynomials

$$\mu_L(S,T) = \sum_{i=0}^{r} S^i \mu_i(T) \quad \text{with} \quad \mu_i(T) = \sum_{x\in L_i} \chi_{L_x}(T).$$

Example 1.52. In Examples 1.42 and 1.50 we considered the lattice L of all subsets of a given finite set of r elements. Since $r(x) = a(x)$ for all $x \in L$, the Möbius polynomial of L is equal to the coboundary polynomial of L, so $\mu_L(S,T) = (S+T-1)^r$.

Example 1.53. 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 1.44. In Example 1.51 we calculated the characteristic polynomial of this lattice. In the same way, we find that

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

Remark 1.13. Let L be a geometric lattice. Then

$$\begin{aligned} \sum_{i=0}^{r(L)} \mu_i(T) &= \mu_L(1,T) \\ &= \sum_{y\in L} \sum_{0\leq x\leq y} \mu(x,y) T^{r(L)-r(y)} \\ &= T^{r(L)} \end{aligned}$$

since $\sum_{0\leq x\leq y} \mu(x,y) = 0$ for all $0 < y$ in L by Proposition 1.40. Similarly $\sum_{i=0}^{n} \chi_i(T) = \chi_L(1,T) = T^{r(L)}$. Also $\sum_{w=0}^{n} A_w(T) = T^k$ for the extended weights of a code of dimension k by Propositions 1.21 and 1.22 for $t = 0$.

Example 1.54. Let L be the lattice consisting of $[n]$ and all its subsets of size at most $k-1$ as in Example 1.36, which is also the lattice of the uniform matroid $U_{n,k}$ and the lattice of an MDS code with parameters

$[n, k, n-k+1]$. Then $\mu_i(T)$ and $\chi_i(T)$ are both equal to $\binom{n}{i}(T-1)^{n-i}$ for all $i < k$ as in Example 1.50, and $\chi_i(T) = 0$ for all $k \leq i < n$, since $a(1_L) = n$, $r(1_L) = k$ and $a(x) = r(x)$ for all x in L_i and $i < k$. Remark 1.13 implies

$$\mu_k(T) = T^k - \sum_{i<k} \binom{n}{i}(T-1)^{n-i} \text{ and } \chi_n(T) = T^k - \sum_{i<k} \binom{n}{i}(T-1)^{n-i}.$$

Let L be a finite geometric lattice. The *Whitney numbers* w_i and W_i of *the first and second kind*, respectively, are defined by

$$w_i = \sum_{x \in L_i} \mu(x) \text{ and } W_i = |L_i|.$$

The *doubly indexed Whitney numbers* w_{ij} and W_{ij} of *the first and second kind*, respectively, are defined by

$$w_{ij} = \sum_{x \in L_i} \sum_{y \in L_j} \mu(x, y) \text{ and } W_{ij} = |\{(x, y) : x \in L_i, y \in L_j, x \leq y\}|.$$

In particular $w_j = w_{0j}$ and $W_j = W_{0j}$. See [80] and [63, §6.6.D] and [35, Chapter 14] and [30, §3.11].

Remark 1.14. We have that

$$\chi_L(T) = \sum_{i=0}^{r(L)} w_i T^{r(L)-i} \text{ and } \mu_L(S, T) = \sum_{i=0}^{r(L)} \sum_{j=0}^{r(L)} w_{ij} S^i T^{r(L)-j}.$$

Hence the (doubly indexed) Whitney numbers of the first kind are determined by $\mu_L(S, T)$. The leading coefficient of

$$\mu_i(T) = \sum_{x \in L_i} \sum_{x \leq y} \mu(x, y) T^{r(L_x) - r_{L_x}(y)}$$

is equal to $\sum_{x \in L_i} \mu(x, x) = |L_i| = W_i$. Hence the Whitney numbers of the second kind W_i are also determined by $\mu_L(S, T)$. We will see in Example 1.60 that the Whitney numbers are not determined by $\chi_L(S, T)$. Finally, let $r = r(L)$. Then

$$\mu_{r-1}(T) = (T-1) \cdot W_{r-1}.$$

1.8.3. *Minimal codewords and subcodes*

A *minimal codeword* of a code C is a codeword whose support does not properly contain the support of another codeword.

The zero word is a minimal codeword. Notice that a nonzero scalar multiple of a minimal codeword is again a minimal codeword. Nonzero minimal codewords play a role in minimum distance decoding. Minimal codewords play a role in minimum distance decoding algorithms [11, 58, 81] and secret sharing schemes and access structures [82, 83]. We can generalize this notion to subcodes instead of words.

A *minimal subcode of dimension* r of a code C is an r-dimensional subcode whose support is not properly contained in the support of another r-dimensional subcode.

A minimal codeword generates a minimal subcode of dimension one, and all the elements of a minimal subcode of dimension one are minimal codewords. A codeword of minimal weight is a nonzero minimal codeword, but the converse is not always the case.

In Example 1.60 we will see two codes that have the same Tutte polynomial, but a different number of minimal codewords. Hence the number of minimal codewords and subcodes is not determined by the Tutte polynomial. However, the number of minimal codewords and the number of minimal subcodes of a given dimension are given by the Möbius polynomial.

Theorem 1.17. *Let C be a code of dimension k. Let $0 \leq r \leq k$. Then the number of minimal subcodes of dimension r is equal to W_{k-r}, the $(r-k)$-th Whitney number of the second kind, and it is determined by the Möbius polynomial.*

Proof. Let D be a subcode of C of dimension r. Let J be the complement in $[n]$ of the support of D. If $\mathbf{d} \in D$ and $d_j \neq 0$, then $j \in \mathrm{supp}(D)$ and $j \notin J$. Hence $D \subseteq C(J)$. Now suppose moreover that D is a minimal subcode of C. Without loss of generality we may assume that D is systematic at the first r positions. So D has a generator matrix of the form $(I_r|A)$. Let $\mathbf{d}_i$ be the i-th row of this matrix. Let $\mathbf{c} \in C(J)$. If $\mathbf{c} - \sum_{i=1}^{r} c_i \mathbf{d}_i$ is not the zero word, then the subcode D' of C generated by $\mathbf{c}, \mathbf{d}_2, \ldots, \mathbf{d}_r$ has dimension r and its support is contained in $\mathrm{supp}(D) \setminus \{1\}$ and $1 \in \mathrm{supp}(D)$. This contradicts the minimality of D. Hence $\mathbf{c} - \sum_{i=1}^{r} c_i \mathbf{d}_i = 0$ and $\mathbf{c} \in D$. Therefore $D = C(J)$.
To find a minimal subcode of dimension r, we fix $l(J) = r$ and minimize the support of $C(J)$ with respect to inclusion. Because J is contained in the complement in $[n]$ of the support of $C(J)$, this is equivalent to maximize J

with respect to inclusion. In matroid terms this means we are maximizing J for $r(J) = k - l(J) = k - r$. This means $J = \bar{J}$ is a flat of rank $k - r$ by Remark 1.8. The flats of a matroid are the elements in the geometric lattice $L = L(M)$. The number of $(k-r)$-dimensional elements in $L(M)$ is equal to $|L_{k-r}|$, which is equal to the Whitney number of the second kind W_{k-r} and thus equal to the leading coefficient of $\mu_{k-r}(T)$ by Remark 1.14. Hence the Möbius polynomial determines all the numbers of minimal subcodes of dimension r for $0 \leq r \leq k$. □

Note that the flats of dimension $k - r$ in a matroid are exactly the hyperplanes in the $(r-1)$-th truncated matroid $T^{r-1}(M)$. This gives another proof of the result of Britz [61, Theorem 3] that the minimal supports of dimension r are the cocircuits of the $(r-1)$-th truncated matroid. For $r = 1$, this gives the well-known equivalence between nonzero minimal codewords and cocircuits. See [35, Theorem 9.2.4] and [39, 1.21].

The number of minimal subcodes of dimension r does not change after extending the code under a finite field extension, since this number is determined by the Möbius polynomial of the lattice of the code, and this lattice does not change under a finite field extension.

1.8.4. *The characteristic polynomial of an arrangement*

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 $\mathbb{F}_q[X_1, \ldots, X_k]$. Then $\mathcal{X}(\mathbb{F}_{q^m})$ is the set of all points $\mathcal{X}$ with coordinates in $\mathbb{F}_{q^m}$, also called the the set of $\mathbb{F}_{q^m}$-*rational points* of $\mathcal{X}$. Note the similarity with extension codes.

A central 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 [84]. Something similar can be said about arrangements over finite fields.

Proposition 1.45. *Let q be a prime power, and let $\mathcal{A} = (H_1, \ldots, H_n)$ be a simple and central 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 [57, Theorem 2.2] and [85, Proposition 3.2] and [74, Sect. 16] and [70, Theorem 2.69].

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 with the reverse inclusion as partial order. The principle of inclusion/exclusion as formulated in Example 1.38 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 geometric lattice $L(\mathcal{A})$ of the arrangement $\mathcal{A} = (H_1, \ldots, H_n)$ with rank function $r = r_L$, so $\dim(x) = r(L) - r(x)$. □

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 1.46 (Deletion-restriction formula).
Let $\mathcal{A} = (H_1, \ldots, H_n)$ be a simple and central arrangement in $\mathbb{F}^k$ over the field $\mathbb{F}$. Let $H = H_i$. Then

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

Proof. Note the similarity of this theorem with the corresponding Proposition 1.32 for graphs. A proof can be given using the deletion-contraction formula for matroids in Proposition 1.38, Remark 1.11 that gives the relation between the two variable characteristic and the Tutte polynomial and the fact that $\chi_{\mathcal{A}}(T) = \chi_{\mathcal{A}}(0, T)$. Another proof for an arbitrary field can be found in Orlik [70, Theorem 2.56]. Here a proof of the special case of a central arrangement over the finite field $\mathbb{F}_q$ will be given by a counting argument. Without loss of generality we may assume that $H = H_1$. Denote $H_j(\mathbb{F}_{q^m})$ by H_j and $\mathbb{F}_{q^m}^k$ by V. Then we have the following disjoint union:

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

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 1.45. 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. □

Let $\mathcal{A} = (H_1, \ldots, H_n)$ be a central simple arrangement over the field $\mathbb{F}$ in $\mathbb{F}^k$. Let $J \subseteq [n]$. Define $H_J = \cap_{j \in J} H_j$. Consider the decreasing sequence

$$\mathcal{N}_k \subset \mathcal{N}_{k-1} \subset \cdots \subset \mathcal{N}_1 \subset \mathcal{N}_0$$

of algebraic subsets of the affine space $\mathbb{A}^k$, defined by

$$\mathcal{N}_i = \bigcup_{\substack{J \subseteq [n] \\ r(H_J) = i}} H_J.$$

Define $\mathcal{M}_i = (\mathcal{N}_i \setminus \mathcal{N}_{i+1})$.

Note that $\mathcal{N}_0 = \mathbb{A}^k$, $\mathcal{N}_1 = \cup_{j=1}^n H_j$, $\mathcal{N}_k = \{0\}$ and $\mathcal{N}_{k+1} = \emptyset$. Furthermore, $\mathcal{N}_i$ is a union of linear subspaces of $\mathbb{A}^k$ all of dimension $k-i$. Remember from Remark 1.3 that H_J is isomorphic with $C(J)$ in case $\mathcal{A}$ is the arrangement of the generator matrix G of the code C.

Proposition 1.47. *Let $\mathcal{A} = (H_1, \ldots, H_n)$ be a central simple arrangement over the field $\mathbb{F}$ in $\mathbb{F}^k$. Let $z(\mathbf{x}) = \{j \in [n] : \mathbf{x} \in H_j\}$ and $r(\mathbf{x}) = r(H_{z(\mathbf{x})})$ the* rank *of $\mathbf{x}$ for $\mathbf{x} \in \mathbb{A}^k$. Then*

$$\mathcal{N}_i = \{\mathbf{x} \in \mathbb{A}^k : r(\mathbf{x}) \geq i\} \quad \textit{and} \quad \mathcal{M}_i = \{\mathbf{x} \in \mathbb{A}^k : r(\mathbf{x}) = i\}.$$

Proof. Let $\mathbf{x} \in \mathbb{A}^k$ and $\mathbf{c} = \mathbf{x}G$. Let $\mathbf{x} \in \mathcal{N}_i$. Then there exists a $J \subseteq [n]$ such that $r(H_J) = i$ and $\mathbf{x} \in H_J$. So $c_j = 0$ for all $j \in J$. So $J \subseteq z(\mathbf{x})$. Hence $H_{z(\mathbf{x})} \subseteq H_J$. Therefore $r(\mathbf{x}) = r(H_{z(\mathbf{x})}) \geq r(H_J) = i$. The converse implication is proved similarly.
The statement about $\mathcal{M}_i$ is a direct consequence of the one about $\mathcal{N}_i$. □

Proposition 1.48. *Let $\mathcal{A}$ be a central simple arrangement over $\mathbb{F}_q$. Let $L = L(\mathcal{A})$ be the geometric lattice of $\mathcal{A}$. Then*

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

Proof. See also [57, Theorem 6.3]. Remember from Remark 1.12 that $\mu_i(T) = \sum_{r(x)=i} \chi_{L_x}(T)$. 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 1.45. Now $\mathcal{M}_i$ is the disjoint union of complements of the arrangements of $\mathcal{A}_x$ for all $x \in L$ such that $r(x) = i$ by Proposition

1.47. Hence

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

□

1.8.5. *The characteristic polynomial of a code*

Proposition 1.49. *Let C be a nondegenerate $\mathbb{F}_q$-linear code. Then*

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

Proof. The short proof is given by $\chi_C(T) = \chi_C(0,T) = \chi_0(T) = A_n(T)$. The geometric interpretation is as follows.
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 1.16 and because the arrangements corresponding to C and $C \otimes \mathbb{F}_{q^m}$ are the same. So $A_n(q^m) = \chi_C(q^m)$ for all positive integers m by Proposition 1.45. Hence $A_n(T) = \chi_C(T)$. □

Let G be a generator matrix of a $[n,k]$ code C over $\mathbb{F}_q$. Define

$$\mathcal{Y}_i = \{\mathbf{x} \in \mathbb{A}^k : \mathrm{wt}(\mathbf{x}G) \le n-i\} \quad \text{and} \quad \mathcal{X}_i = \{\mathbf{x} \in \mathbb{A}^k : \mathrm{wt}(\mathbf{x}G) = n-i\}.$$

The $\mathcal{Y}_i$ form a decreasing sequence

$$\mathcal{Y}_n \subseteq \mathcal{Y}_{n-1} \subseteq \cdots \subseteq \mathcal{Y}_1 \subseteq \mathcal{Y}_0$$

of algebraic subsets of $\mathbb{A}^k$, and $\mathcal{X}_i = (\mathcal{Y}_i \setminus \mathcal{Y}_{i+1})$. Suppose that G has no zero column. Let $\mathcal{A}_G$ be the arrangement of G. Then

$$\mathcal{X}_i = \{\mathbf{x} \in \mathbb{A}^k : \mathbf{x} \text{ is in exactly } i \text{ hyperplanes of } \mathcal{A}_G\}.$$

Proposition 1.50. *Let C be a projective code of length n. Then*

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

Proof. Every $\mathbf{x} \in \mathbb{F}_{q^m}^k$ corresponds one-to-one to codeword in $C \otimes \mathbb{F}_{q^m}$ via the map $\mathbf{x} \mapsto \mathbf{x}G$. So $|\mathcal{X}_i(\mathbb{F}_{q^m})| = A_{n-i}(q^m)$. Also, $A_{n-i}(q^m) = \chi_i(q^m)$ for all i, by Remark 1.11. See also [86, Theorem 3.3]. □

The similarity between Proposition 1.50 and Proposition 1.16 gives a geometric argument for the relation between $\chi_C(S,T)$ and $W_C(X,Y,T)$ in Remark 1.11.

Remark 1.15. Another way to define $\mathcal{X}_i$ is the collection of all points $P \in \mathbb{A}^k$ such that P is on exactly i distinct hyperplanes of the arrangement $\mathcal{A}_G$. Denote the arrangement of hyperplanes in $\mathbb{P}^{k-1}$ also by $\mathcal{A}_G$, and let $\bar{P}$ be the point in $\mathbb{P}^{k-1}$ corresponding to $P \in \mathbb{A}^k$. Define

$$\bar{\mathcal{X}}_i = \{\bar{P} \in \mathbb{P}^{k-1} : \bar{P} \text{ is on exactly } i \text{ hyperplanes of } \mathcal{A}_G\}.$$

For all $i < n$ the polynomial $\chi_i(T)$ is divisible by $T-1$. Define $\bar{\chi}_i(T) = \chi_i(T)/(T-1)$. Then $\bar{\chi}_i(q^m) = |\bar{\mathcal{X}}_i(\mathbb{F}_{q^m})|$ for all $i < n$ by Proposition 1.50.

Theorem 1.18. *Let G be a generator matrix of a nondegenerate code C. Let $\mathcal{A}_G$ be the associated central arrangement. Let $d^\perp = d(C^\perp)$. Then $\mathcal{N}_i \subseteq \mathcal{Y}_i$ for all i, equality holds for all $i < d^\perp$. Also, $\mathcal{M}_i = \mathcal{X}_i$ for all $i < d^\perp - 1$. If furthermore C is projective, then*

$$\mu_i(T) = \chi_i(T) = A_{n-i}(T) \text{ for all } i < d^\perp - 1.$$

Proof. Let $\mathbf{x} \in \mathcal{N}_i$. Then $\mathbf{x} \in H_J$ for some $J \subseteq [n]$ such that $r(H_J) = i$. So $|J| \geq i$ and $\mathrm{wt}(\mathbf{x}G) \leq n-i$ by Proposition 1.16. Hence $\mathbf{x} \in \mathcal{Y}_i$. Therefore $\mathcal{N}_i \subseteq \mathcal{Y}_i$.
Let $i < d^\perp$ and $\mathbf{x} \in \mathcal{Y}_i$. Then $\mathrm{wt}(\mathbf{x}G) \leq n - i$. Let $J = \mathrm{supp}(\mathbf{x}G)$. Then $|J| \geq i$. Take a subset I of J such that $|I| = i$. Then $\mathbf{x} \in H_I$ and $r(I) = |I| = i$ by Lemma 1.2, since $i < d^\perp$. Hence $\mathbf{x} \in \mathcal{N}_i$. Therefore $\mathcal{Y}_i \subseteq \mathcal{N}_i$. So $\mathcal{Y}_i = \mathcal{N}_i$ for all $i < d^\perp$, and $\mathcal{M}_i = \mathcal{X}_i$ for all $i < d^\perp - 1$.
The code is nondegenerate. So $d(C^\perp) \geq 2$. Suppose furthermore that C is projective. Then $\mu_i(T) = \chi_i(T) = A_{n-i}(T)$ for all $i < d^\perp - 1$, by Remark 1.11 and Propositions 1.50 and 1.48. □

The extended and generalized weight enumerators are determined by the pair (n, k) for a $[n, k]$ MDS code by Theorem 1.8. If C is a $[n, k]$ code, then $d(C^\perp)$ is at most $k+1$. Furthermore $d(C^\perp) = k+1$ if and only if C is MDS if and only if $C^\perp$ is MDS. A $[n, k, d]$ code is called *almost MDS* if $d = n-k$. So $d(C^\perp) = k$ if and only if $C^\perp$ is almost MDS. If C is almost MDS, then $C^\perp$ is not necessarily almost MDS. The code C is called *near MDS* if both C and $C^\perp$ are almost MDS. See [87].

Proposition 1.51. *Let C be a $[n, k, d]$ code such that $C^\perp$ is MDS or almost MDS and $k \geq 3$. Then both $\chi_C(S, T)$ and $W_C(X, Y, T)$ determine $\mu_C(S, T)$. In particular*

$$\mu_i(T) = \chi_i(T) = A_{n-i}(T) \text{ for all } i < k - 1,$$

$$\mu_{k-1}(T) = \sum_{i=k-1}^{n-1} \chi_i(T) = \sum_{i=k-1}^{n-1} A_{n-i}(T),$$

and $\mu_k(T) = 1$.

Proof. Let C be a code such that $d(C^\perp) \geq k \geq 3$. Then C is projective and $\mu_i(T) = \chi_i(T) = A_{n-i}(T)$ for all $i < k-1$ by Theorem 1.18. Furthermore, $\mu_k(T) = \chi_n(T) = A_0(T) = 1$.
Finally let $L = L(C)$. Then $\sum_{i=0}^{k} \mu_i(T) = T^k$, $\sum_{i=0}^{n} \chi_i(T) = T^k$ and $\sum_{i=0}^{n} A_i(T) = T^k$ by Remark 1.13. Hence the formula for $\mu_{k-1}(T)$ holds. Therefore $\mu_C(S,T)$ is determined both by $W_C(X,Y,T)$ and $\chi_C(S,T)$. □

Projective codes of dimension 3 are examples of codes C such that $C^\perp$ is almost MDS. In the following we will give explicit formulas for $\mu_C(S,T)$ for such codes.

Let C be a projective code of length n and dimension 3 over $\mathbb{F}_q$ with generator matrix G. The arrangement $\mathcal{A}_G = (H_1, \ldots, H_n)$ of planes in $\mathbb{F}_q^3$ is simple and essential, and the corresponding arrangement of lines in $\mathbb{P}^2(\mathbb{F}_q)$ is also denoted by $\mathcal{A}_G$. We defined in Remark 1.15 that

$$\bar{\mathcal{X}}_i(\mathbb{F}_{q^m}) = \{\bar{P} \in \mathbb{P}^2(\mathbb{F}_{q^m}) : \bar{P} \text{ is on exactly } i \text{ lines of } \mathcal{A}_G\}$$

and $\bar{\chi}_i(q^m) = |\bar{\mathcal{X}}_i(\mathbb{F}_{q^m})|$ for all $i < n$.

Notice that for projective codes of dimension 3 we have $\bar{\mathcal{X}}_i(\mathbb{F}_{q^m}) = \bar{\mathcal{X}}_i(\mathbb{F}_q)$ for all positive integers m and $2 \leq i < n$. Abbreviate in this case $\bar{\chi}_i(q^m) = \bar{\chi}_i$ for $2 \leq i < n$.

Proposition 1.52. *Let C be a projective code of length n and dimension 3 over $\mathbb{F}_q$. Then*

$$\begin{cases} \mu_0(T) = (T-1)\left(T^2 - (n-1)T + \sum_{i=2}^{n-1}(i-1)\bar{\chi}_i - n + 1\right), \\ \mu_1(T) = (T-1)\left(nT + n - \sum_{i=2}^{n-1} i\bar{\chi}_i\right), \\ \mu_2(T) = (T-1)\left(\sum_{i=2}^{n-1} \bar{\chi}_i\right), \\ \mu_3(T) = 0. \end{cases}$$

Proof. A more general statement and proof is possible for $[n,k]$ codes C such that $d(C^\perp) \geq k$, using Proposition 1.51, the fact that $B_t(T) = T^{k-t} - 1$ for all $t < d(C^\perp)$ by Lemma 1.2 and the expression of $B_t(T)$ in terms of $A_w(T)$ by Proposition 1.17. We will give a second geometric proof for the

special case of projective codes of dimension 3.
By Lagrange interpolation it is enough to show this proposition with $T = q^m$ for all m. Notice that $\mu_i(q^m)$ is the number of elements of $\mathcal{M}_i(\mathbb{F}_{q^m})$ by Proposition 1.48. 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 $\mathcal{M}_i(\mathbb{F}_{q^m})$ by $\mathcal{M}_i$. Define $\bar{\mathcal{M}}_i = \{\bar{P} : P \in \mathcal{M}_i\}$. Then $|\mathcal{M}_i| = (q^m - 1)|\bar{\mathcal{M}}_i|$ for all $i < 3$.
If $\bar{P} \in \bar{\mathcal{M}}_2$, then $\bar{P} \in H_j \cap H_k$ for some $j \neq k$. Hence $\bar{P} \in \bar{\mathcal{X}}_i(\mathbb{F}_q)$ for some $i \geq 2$, since the code is projective. So $\bar{\mathcal{M}}_2$ is the disjoint union of the $\bar{\mathcal{X}}_i(\mathbb{F}_q)$ for $2 \leq i < n$. Therefore $|\bar{\mathcal{M}}_2| = \sum_{i=2}^{n-1} \bar{\chi}_i$.
$\bar{P} \in \bar{\mathcal{M}}_1$ if and only if $\bar{P}$ 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 $\bar{P} \in \bar{\mathcal{X}}_i(\mathbb{F}_q)$ is on i lines H_j. Hence $|\bar{\mathcal{M}}_1| = n(q^m + 1) - \sum_{i=2}^{n-1} i\bar{\chi}_i$.
Finally, $\mathbb{P}^2(\mathbb{F}_{q^m})$ is the disjoint union of $\bar{\mathcal{M}}_0$, $\bar{\mathcal{M}}_0$ and $\bar{\mathcal{M}}_2$. The numbers $|\bar{\mathcal{M}}_2|$ and $|\bar{\mathcal{M}}_1|$ are already computed, and $|\mathbb{P}^2(\mathbb{F}_{q^m})| = q^{2m} + q^m + 1$. From this we derive the number of elements of $\bar{\mathcal{M}}_0$. □

Note that $\mu_i(T)$ is divisible by $T - 1$ for all $0 \leq i < k$. Define $\bar{\mu}_i = \mu_i(T)/(T - 1)$. Define similarly $\bar{A}_w = A_w(T)/(T - 1)$ for all $0 < w \leq n$.

1.8.6. *Examples and counterexamples*

Example 1.55. Consider the matrices G given by

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

Let C be the code over $\mathbb{F}_q$ with generator matrix G. For $q = 2$, this is the simplex code $\mathcal{S}_2(2)$. The columns of G represent also the coefficients of the lines of $\mathcal{A}_G$. The projective picture of $\mathcal{A}_G$ is given in Figure 1.12.

If q is odd, then there are 3 points on two lines and 6 points on three lines, so $\bar{\chi}_2 = 3$ and $\bar{\chi}_3 = 6$. The number of points that are on one line is equal to the number of points on each of the seven lines, minus the points we already counted, with multiplicity: $7(T + 1) - 3 \cdot 2 - 6 \cdot 3 = 7T - 17$. There are no points on more than three lines, so $\bar{\chi}_i = 0$ for $i > 3$. We calculate $\bar{\chi}_0$ via $\bar{\chi}_0 + \bar{\chi}_1 + \bar{\chi}_2 + \bar{\chi}_3 = T^2 + T + 1$.
If q is even, we can do the same kind of calculation. The values of $\bar{\mu}_i$ can be calculated using Proposition 1.52, but they follow more directly from Proposition 1.51. The results are in the next table:

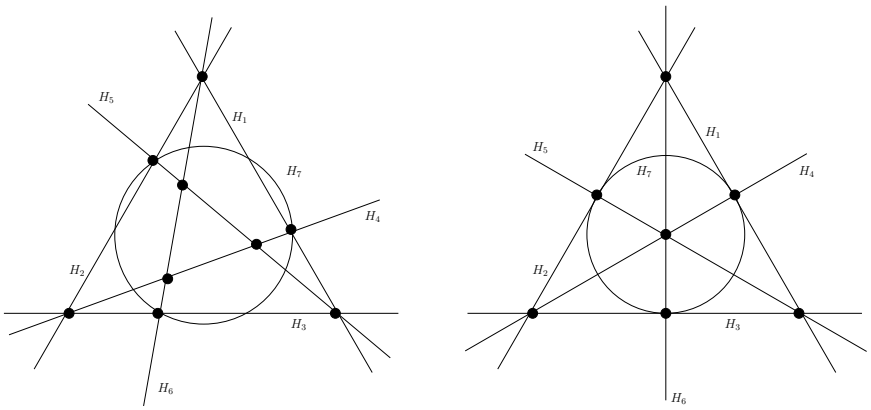


Fig. 1.12. The arrangement of G for q odd and q even

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

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 $\bar{A}_7 = (T-2)(T-4)$ or $\bar{A}_7 = (T-3)^2$, respectively.

Example 1.56. Let G be a $3 \times n$ generator matrix of an MDS code. As mentioned in Example 1.30, 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 $\bar{\chi}_2 = \binom{n}{2}$ and $\bar{\chi}_i = 0$ for all $i > 2$. By Proposition 1.52 we have $\bar{\mu}_2 = \binom{n}{2}$ and $\bar{\mu}_1 = nT + 2n - n^2$ and $\bar{\mu}_0 = T^2 - (n-1)T + \binom{n-1}{2}$. By Proposition 1.48 we find $A_i = 0$ for $0 < i < n-2$, $\bar{A}_{n-2} = \bar{\chi}_2$ and $\bar{A}_{n-1} = \bar{\chi}_1 = \bar{\mu}_1$ and $\bar{A}_n = \bar{\chi}_0 = \bar{\mu}_0$. The values found for the extended weight enumerator are in agreement with Theorem 1.8.

Example 1.57. 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 $a+b$ lines of the projective arrangement of $\mathcal{A}_G$ consists of a distinct lines incident with P, and b distinct lines incident with Q and there is no line incident with P and Q. Then $\bar{\chi}_2 = \bar{A}_{n-2} = ab$, $\bar{\chi}_a = \bar{A}_b = 1$ and $\bar{\chi}_b = \bar{A}_a = 1$. Hence $\bar{\mu}_2(T) = ab + 2$. Furthermore

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

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

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

Example 1.58. Let a, b and c be positive integers such that $2 < a < b < c$. Let $n = a + b + c$. Let G be a $3 \times n$ generator matrix of a nondegenerate code $C(a, b, c)$. Suppose that there are three points P, Q and R in the projective plane over $\mathbb{F}_q$ such that the lines of the projective arrangement of $\mathcal{A}_G$ consist of a distinct lines incident with P and not with Q and R, b distinct lines incident with Q and not with P and R, and c distinct lines incident with R and not with P and Q. The a lines through P intersect the b lines through Q in ab points. Similarly statements hold for the lines through P and R intersecting in ac points, and the lines through Q and R intersecting in bc points. Suppose that all these $ab + bc + ac$ intersection points are mutually distinct, so every intersection point lies on exactly two lines of the arrangement. If q is large enough, then such a configurations exists.

The number of points on two lines of the arrangement is $\bar{\chi}_2 = ab + bc + ca$. Since P is the unique point on exactly a lines of the arrangement, we have $\bar{\chi}_a = 1$. Similarly $\bar{\chi}_b = \bar{\chi}_c = 1$. Finally $\bar{\chi}_i = 0$ for all $2 \leq i < n$ and $i \notin \{2, a, b, c\}$. Propositions 1.51 and 1.52 imply that $\bar{A}_{n-a} = \bar{A}_{n-b} = \bar{A}_{n-c} = 1$ and $\bar{A}_{n-2} = ab + bc + ca$ and $\bar{\mu}_2 = ab + bc + ca + 3$. Furthermore

$$\bar{\mu}_1 = \bar{\chi}_1 = \bar{A}_{n-1} = nT - 2(ab + bc + ca),$$

$$\bar{\mu}_0 = \bar{\chi}_0 = \bar{A}_n = T^2 - (n-1)T + ab + bc + ca - 2$$

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

Therefore $W_{C(a,b,c)}(X, Y, T) = W_{C(a',b',c')}(X, Y, T)$ if and only if $(a, b, c) = (a', b', c')$, and $\mu_{C(a,b,c)}(S, T) = \mu_{C(a',b',c')}(S, T)$ if and only if $a+b+c = a'+b'+c'$ and $ab+bc+ca = a'b'+b'c'+c'a'$. In particular let $C_1 = C(3, 9, 14)$ and $C_2 = C(5, 6, 15)$. Then C_1 and C_2 are two projective codes with the same Möbius polynomial $\mu_C(S, T)$ but distinct extended weight enumerators and coboundary polynomial $\chi_C(S, T)$.

Now $d(C(a, b, c)) = n - c$. Hence $d(C_1) = 12$ and $d(C_2) = 11$. Therefore

$\mu_C(S,T)$ does not determine the minimum distance although it gives the number of minimal codewords.

Example 1.59. Consider the codes C_3 and C_4 over $\mathbb{F}_q$ with $q > 2$ with generator matrices G_3 and G_4 given by

$$G_3 = \begin{pmatrix} 1 & 1\,0\,0\,1\,0\,0 \\ 0 & 1\,1\,1\,0\,1\,0 \\ -1 & 0\,1\,1\,0\,0\,1 \end{pmatrix} \quad \text{and} \quad G_4 = \begin{pmatrix} 1\,1\,0\,0\,1\,0\,0 \\ 0\,1\,1\,1\,0\,1\,0 \\ 0\,1\,1\,a\,0\,0\,1 \end{pmatrix}$$

where $a \in \mathbb{F}_q \setminus \{0,1\}$. It was shown in Brylawsky [63, Exercise 6.96] that the duals of these codes have the same Tutte polynomial. So the codes C_3 and C_4 have the same Tutte polynomial

$$t_C(X,Y) = 2X + 2Y + 3X^2 + 5XY + 4Y^2 + X^3 + X^2Y + 2XY^2 + 3Y^3 + Y^4.$$

Hence C_3 and C_4 have the extended weight enumerator given by

$$X^7 + (2T-2)X^4Y^3 + (3T-3)X^3Y^4 + (T^2-T)X^2Y^5 + \\ + (5T^2 - 15T + 10)XY^6 + (T^3 - 6T^2 + 11T - 6)Y^7.$$

The codes C_3 and C_4 are not projective and their simplifications $\bar{C}_3$ and $\bar{C}_4$, respectively, have generator matrices given by

$$\bar{G}_3 = \begin{pmatrix} 1 & 1\,0\,1\,0\,0 \\ 0 & 1\,1\,0\,1\,0 \\ -1 & 0\,1\,0\,0\,1 \end{pmatrix} \quad \text{and} \quad \bar{G}_4 = \begin{pmatrix} 1\,1\,0\,0\,0\,0 \\ 0\,1\,1\,1\,1\,0 \\ 0\,1\,1\,a\,0\,1 \end{pmatrix}$$

where $a \in \mathbb{F}_q \setminus \{0,1\}$.
From the arrangement $\mathcal{A}(\bar{C}_3)$ and $\mathcal{A}(\bar{C}_4)$ in Figure 1.13 we deduce the $\bar{\chi}_i$ that are given in the following table.

code $\setminus\, i$	0	1	2	3	4	5
C_3	$T^2 - 5T + 6$	$6T - 12$	3	4	0	0
C_4	$T^2 - 5T + 6$	$6T - 13$	6	1	1	0

Therefore $t_{C_3}(X,Y) = t_{C_4}(X,Y)$ but $\chi_{C_3}(S,T) \neq \chi_{C_4}(S,T)$ and $t_{\bar{C}_3}(X,Y) \neq t_{\bar{C}_4}(X,Y)$.

Example 1.60. Let $C_5 = C_3^\perp$ and $C_6 = C_4^\perp$. Their generator matrices are

$$G_5 = \begin{pmatrix} 1\,0\,0\,0\,1\,0 & -1 \\ 0\,1\,0\,0\,1\,1 & 0 \\ 0\,0\,1\,0\,0\,1 & 1 \\ 0\,0\,0\,1\,0\,1 & 1 \end{pmatrix} \quad \text{and} \quad G_6 = \begin{pmatrix} 1\,0\,0\,0\,1\,0\,0 \\ 0\,1\,0\,0\,1\,1\,1 \\ 0\,0\,1\,0\,0\,1\,1 \\ 0\,0\,0\,1\,0\,1\,a \end{pmatrix}$$

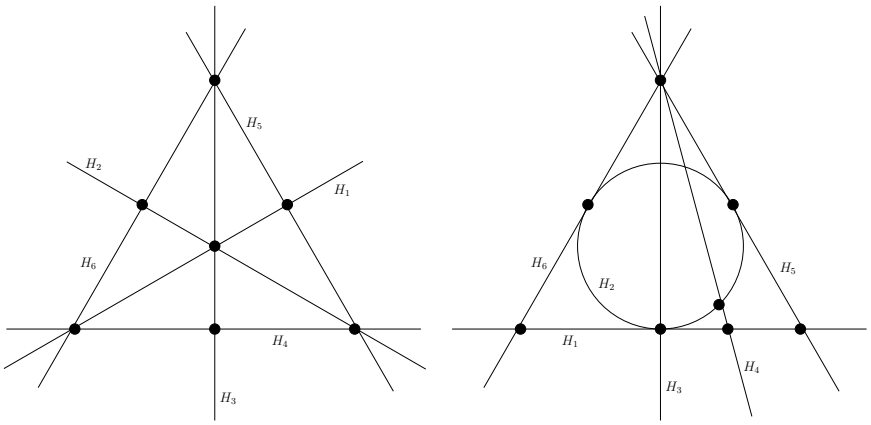


Fig. 1.13. The arrangements of $\bar{G}_3$ and $\bar{G}_4$

where $a \in \mathbb{F}_q \setminus \{0,1\}$. Then C_5 and C_6 have the same Tutte polynomial $t_{C^\perp}(X,Y) = t_C(Y,X)$ as given by Example 1.59:

$$2X + 2Y + 4X^2 + 5XY + 3Y^2 + 3X^3 + 2X^2Y + XY^2 + Y^3 + 3X^4.$$

Hence C_5 and C_6 have the same extended weight enumerator given by

$$\begin{aligned} &X^7 + (T-1)X^5Y^2 \\ &\quad + (6T-6)X^4Y^3 + (2T^2 - T - 1)X^3Y^4 + (15T^2 - 43T + 28)X^2Y^5 \\ &+ (7T^3 - 36T^2 + 60T - 31)XY^6 + (T^4 - 7T^3 + 19T^2 - 23T + 10)Y^7. \end{aligned}$$

The geometric lattice $L(C_5)$ has atoms a, b, c, d, e, f, g corresponding to the first, second, etc. column of G_5. The second level of $L(C_5)$ consists of the following 17 elements:

$$abe,\ ac,\ ad,\ af,\ ag,\ bc,\ bd,\ bf,\ bg,\ cd,\ ce,\ cf,\ cg,\ de,\ df,\ dg,\ efg.$$

The third level consists of the following 12 elements:

$$abce,\ abde,\ abefg,\ acdg,\ acf,\ adf,\ bcdf,\ bcg,\ bdg,\ cde,\ cefg,\ defg.$$

Similarly, the geometric lattice $L(C_6)$ has atoms a, b, c, d, e, f, g corresponding to the first, second, etc. column of G_6. The second level of $L(C_6)$ consists of the following 17 elements:

$$abe,\ ac,\ ad,\ af,\ ag,\ bc,\ bd,\ bf,\ bg,\ cd,\ ce,\ cf,\ cg,\ de,\ dfg,\ ef,\ eg.$$

The third level consists of the following 13 elements:

$$abce,\ abde,\ abef,\ abeg,\ acd,\ acf,\ acg,\ adfg,\ bcdfg,\ cde,\ cef,\ ceg,\ defg.$$

Theorem 1.18 implies that $\mu_0(T)$ and $\mu_1(T)$ are the same for both codes and equal to

$$\mu_0(T) = \chi_0(T) = A_7(T) = (T-1)(T-2)(T^2-4T+5)$$

$$\mu_1(T) = \chi_1(T) = A_6(T) = (T-1)(7T^2-29T+31).$$

The polynomials $\mu_3(T)$ and $\mu_2(T)$ are given in the following table using Remarks 1.14 and 1.13.

	C_5	C_6
$\mu_2(T)$	$17T^2-49T+32$	$17T^2-50T+33$
$\mu_3(T)$	$12T-12$	$13T-13$

This example shows that for projective codes the Möbius polynomial $\mu_C(S,T)$ is not determined by the coboundary polynomial $\chi_C(S,T)$.

1.9. Overview of polynomial relations

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 Figure 1.14.

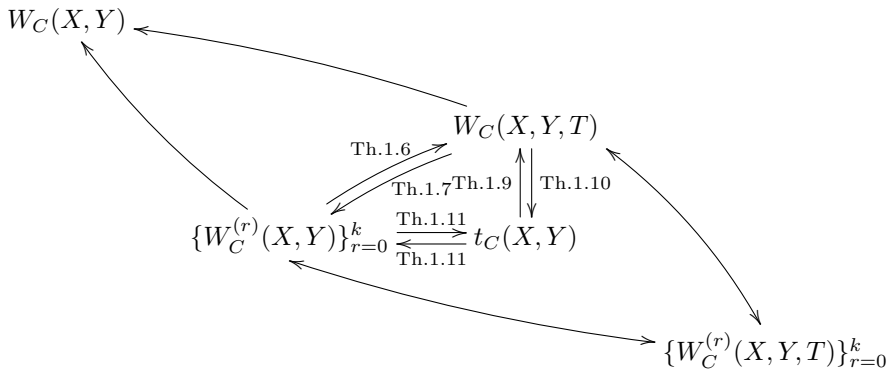


Fig. 1.14. Relations between the weight enumerator and Tutte polynomial

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 [21].

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 than the three underlying polynomials.

Now $t_C(X,Y)$, $R_{M_C}(X,Y)$ and $\chi_C(S,T)$ determine each other on the class of projective codes by Theorem 1.16. This is summarized in Figure 1.15. The dotted arrows only apply if the matroid is simple or, equivalently, if the code is projective.

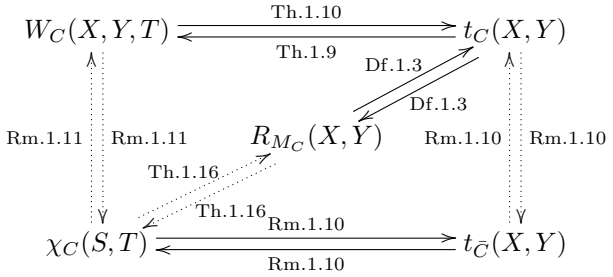


Fig. 1.15. Relations between the weight enumerator, characteristic, and Tutte polynomial

The polynomials $\chi_C(S,T)$ and $\mu_C(S,T)$ do not determine each other by Examples 1.58 and 1.60.

1.10. Further reading and open problems

1.10.1. *Multivariate and other polynomials*

The *multivariate* Tutte or *polychromatic* polynomial of a graph and a matroid is considered in [88–91] and is related to the partition function of the Potts-model in statistical mechanics [92, 93]. The multivariate weight enumerator of a code is considered in [94]. The characteristic and multivariate Tutte polynomial of arrangements are studied in [57, 86, 95].

The *tree polynomial* of a graph is generalized to the *basis polynomial* of a matroid [96]. The characteristic polynomial of a graph is the characteristic polynomial $\det(\lambda I - A)$ of the adjacency matrix A of the graph [46] and is distinct from the chromatic polynomial of the graph and from the characteristic polynomial of the geometric lattice of the graph. The *spectrum* of a graph is the set of eigenvalues of the characteristic polynomial of the graph.

Gray gave an example of two non-isomorphic graphs that have the same Tutte polynomial. This result was generalized in [54] on codichromatic graphs and in [88] on copolychromatic graphs.

Every polynomial in one variable with coefficients in a field $\mathbb{F}$ factorizes in linear factors over the algebraic closure $\bar{\mathbb{F}}$ of $\mathbb{F}$. In Examples 1.50 and 1.51 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 [70].

The theory of matroid complexes gives rise to the *spectrum polynomial* [97]. A recurrence relation is proved in [98, 99] for the spectrum polynomial that is a variation of the deletion-contraction formula for the Tutte polynomial. The Tutte polynomial does not determine the spectrum polynomial. The converse problem is an open question. The multivariate spectrum polynomial is considered in [100].

The theory of knots and links and their Kauffman, Jones and Homfly polynomials have connections with graph theory and the Tutte polynomial. See [101–103].

1.10.2. *The coset leader weight enumerator*

Let C be a linear code of length n over $\mathbb{F}_q$. Let $\mathbf{y} \in \mathbb{F}_q^n$. The weight of the coset $\mathbf{y} + C$ is defined by

$$\mathrm{wt}(\mathbf{y} + C) = \min\{\mathrm{wt}(\mathbf{y} + \mathbf{c}) : \mathbf{c} \in C\}.$$

A *coset leader* is a choice of an element $\mathbf{y} \in \mathbb{F}_q^n$ of minimal weight in its coset, that is $\mathrm{wt}(\mathbf{y}) = \mathrm{wt}(\mathbf{y} + C)$. Let α_i be the number of cosets of C that are of weight i. Let λ_i be the number of $\mathbf{y}$ in $\mathbb{F}_q^n$ that are of minimal weight i in its coset. Then $\alpha_C(X, Y)$, the *coset leader weight enumerator* of C and

$\lambda_C(X,Y)$, the *list weight enumerator* of C are polynomials defined by

$$\alpha_C(X,Y) = \sum_{i=0}^{n} \alpha_i X^{n-i}Y^i \quad \text{and} \quad \lambda_C(X,Y) = \sum_{i=0}^{n} \lambda_i X^{n-i}Y^i.$$

See [8, 104]. The *covering radius* $\rho(C)$ of C is the maximal i such that $\alpha_i(C) \neq 0$. We have $\alpha_i = \lambda_i = \binom{n}{i}(q-1)^i$ for all $i \leq (d-1)/2$, where d is the minimum distance of C. The coset leader weight enumerator gives a formula for the *error probability*, that is the probability that the output of the decoder is the wrong codeword. In this decoding scheme the decoder uses the chosen coset leader as the error vector as explained in Section 1.3.4 and [8, Chap.1 §5]. The list weight enumerator is of interest in case the decoder has as output the list of all nearest codewords [105, 106]. The coset leader weight enumerator is also used in *steganography* to compute the average of changed symbols [107, 108].

The covering radius is determined by the coset leader weight enumerator of a code. The covering radius of a binary code is in general not determined by the Tutte polynomial of the code by [32]. Hence the Tutte polynomial and the extended weight enumerator of a code do not determine the coset leader weight enumerator.

Consider the functions $\alpha_i(T)$ and $\lambda_i(T)$ such that $\alpha_i(q^m)$ and $\lambda_i(q^m)$ are equal to the number of cosets of weight i and the number of elements in $\mathbb{F}_{q^m}^n$ of minimal weight i in its coset, respectively, with respect to the extended coded $C \otimes \mathbb{F}_{q^m}$. Define the *extended coset leader weight enumerator* and the *extended list weight enumerator* [3], respectively, by:

$$\alpha_C(X,Y,T) = \sum_{i=0}^{n} \alpha_i(T) X^{n-i}Y^i \quad \text{and} \quad \lambda_C(X,Y,T) = \sum_{i=0}^{n} \lambda_i(T) X^{n-i}Y^i.$$

In [104, Theorem 2.1] it is shown that the function $\alpha_i(T)$ is determined by finitely many data for all extensions of $\mathbb{F}_q$. In fact, the $\alpha_i(T)$ are polynomials in the variable T. There are well defined nonnegative integers F_{ij} such that

$$\alpha_C(X,Y,T) = 1 + \sum_{i=1}^{n-k} \sum_{j=1}^{n-k} F_{ij}(T-1)(T-q)\cdots(T-q^{j-1})X^{n-i}Y^i.$$

This is similar to the expression of the extended weight enumerator in terms of the generalized weight enumerator as given in Proposition 1.28.

1.10.3. *Graph codes*

Graph codes were studied in [109] and used in [110] to show that decoding linear codes is hard, even if preprocessing is allowed. *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 [111, 112].

1.10.4. *The reconstruction problem*

The reconstruction problem is whether a structure can be reconstructed from certain substructures. The original *vertex reconstruction problem* of Ulam and Kelly is whether a graph with at least three vertices can be reconstructed form the collection of its vertex deleted subgraphs. The *edge reconstruction problem* of a graph $\Gamma = (V, E)$ with at least four edges is whether this graph can be reconstructed form the collection of its edge deleted subgraphs $\Gamma \setminus e$. Both reconstruction problems are still open. See [113] for a survey. One can formulate a corresponding reconstruction problem for matroids. Let $M = (E, \mathcal{I})$ and $N = (E, \mathcal{J})$ be matroids on the same set E. Are M and N isomorphic if $M \setminus e$ and $N \setminus e$ are isomorphic for all e in E? In [114] a counterexample is given for this reconstruction problem for matroids. One can reconstruct the Tutte polynomial of M if one knows the Tutte polynomial of $M \setminus e$ for all e in E, see [115]. See a similar result for the polychromatic polynomial of graphs in [116].

1.10.5. *Questions concerning the Möbius polynomial*

Is it true that the Möbius polynomial of M is determined by the collection of Möbius polynomials of all $M \setminus e$ with e in E?

The doubly indexed Whitney numbers of the first kind and the Whitney numbers of the second kind are determined by the Möbius polynomial as was shown in Remark 1.14 and Theorem 1.17. Are the doubly indexed Whitney numbers of the second kind determined by the Möbius polynomial?

The geometric lattice of a matroid M is equal to the geometric lattice of its simplification $\bar{M}$ by Proposition 1.43. So information is lost by this process. The dual of a simple matroid is not necessarily simple. Similarly the dual of a projective code is not necessarily projective. Now suppose that both C

and its dual are projective. Is there a MacWilliams type of formula for the Möbius polynomial? In other words: Is $\mu_C(S,T)$ determined by $\mu_{C^\perp}(S,T)$? A similar question could be asked for matroids M such that M and $M^\perp$ are simple.

We have seen in Example 1.58 that the Tutte polynomial and the coboundary polynomial are not determined by the Möbius polynomial of a projective code. Is $\chi_C(S,T)$ determined by the polynomials $\mu_C(S,T)$ and/or $\mu_{C^\perp}(S,T)$ if C and $C^\perp$ are projective?

1.10.6. *Monomial conjectures*

A sequence of real numbers $(v_0, v_1, \ldots, v_r)$ is called *unimodal* if

$$v_i \geq \min\{v_{i-1}, v_{i+1}\} \text{ for all } 0 < i < r.$$

The sequence is called *logarithmically concave* or *log-concave* if

$$v_i^2 \geq v_{i-1}v_{i+1} \text{ for all } 0 < i < r.$$

The Whitney numbers of the first kind are alternating in sign. That is

$$w_i^+ := (-1)^i w_i > 0 \text{ for all } i.$$

It was conjectured by Rota [117] that the Whitney numbers w_i^+ are unimodal. See [118, Problem 12]. Welsh [36] conjectured that the Whitney numbers w_i^+ are log-concave by generalizing a conjecture of Read [119] on graphs. It was shown that the following weaker version of the unimodal property is true for a matroid M of rank r:

$$w_i^+ < w_j^+ \text{ for all } 0 \leq i \leq r/2 \text{ and } i < j \leq r - i.$$

See [120, Corollary 8.4.2].

1.10.7. *Complexity issues*

The computation of the minimum distance and the weight enumerator of a code are NP hard problems [11, 13]. The computation of the coefficients of the Tutte polynomial of planar graphs is #P hard, but also the evaluation at a specific point (x, y) is #P-hard except for 9 points and two special curves [103, 121–123].

1.10.8. *The zeta function*

The counting of rational points over field extensions $\mathbb{F}_{q^m}$ is computed by the zeta function. 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 $\mathbb{F}_q[X_1, \ldots, X_k]$. Then $\mathcal{X}(\mathbb{F}_{q^m})$ is the set of all points $\mathcal{X}$ with coordinates in $\mathbb{F}_{q^m}$, also called the the set of $\mathbb{F}_{q^m}$*-rational points* of $\mathcal{X}$. 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 1.19. *Let $\mathcal{A}$ be a central simple 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 [85, Theorem 3.6]. □

The numbers $|c_j|$ can be interpreted as the Betti numbers of the cohomology of the complement of the arrangement over the algebraic closure of the finite field, which is analogous to the situation over the complex numbers [70, 124].

The (two variable) zeta function of a code as defined by Duursma [24, 125, 126] is motivated by algebraic geometry codes on curves and the zeta function of the curve. It is related to the extended and generalized weight enumerator of the code and not to the zeta function of the arrangement of the code.

References

[1] R. Jurrius. Classifying polynomials of linear codes. Master's thesis, Leiden University, (2008).

[2] R. Jurrius and R. Pellikaan. Extended and generalized weight enumerators. In eds. T. Helleseth and Ø. Ytrehus, *Proc. Int. Workshop on Coding and Cryptography WCC-2009*, pp. 76–91. Selmer Center, Bergen, (2009).

[3] R. Jurrius and R. Pellikaan. The extended coset leader weight enumerator. In eds. F. Willems and T. Tjalkens, *Proc. 30th Symposium 2009 on Information Theory in the Benelux*, pp. 217–224. WIC, Eindhoven, (2009).
[4] R. Jurrius and R. Pellikaan. Codes, arrangemens and weight enumerators. Soria Summer School on Computational Mathematics (S3CM): Applied Computational Algebraic Geomertric Modelling, (2009).
[5] E. Berlekamp, *Algebraic coding theory.* (Aegon Park Press, Laguna Hills, 1984).
[6] R. Blahut, *Theory and practice of error control codes.* (Addison-Wesley, Reading, 1983).
[7] J. v. Lint, *Introduction to coding theory. Third edition, Graduate Texts in Math. vol. 86.* (Springer, Berlin, 1999).
[8] F. MacWilliams and N. Sloane, *The theory of error-correcting codes.* (North-Holland Mathematical Library, Amsterdam, 1977).
[9] R. Hamming, Error detecting and error correcting codes, *Bell System Technical Journal.* **29**, 147–160, (1950).
[10] A. Shannon, A mathematical theory of communication, *Bell System Technical Journal.* **27**, 379–423 and 623–656, (1948).
[11] A. Barg. Complexity issues in coding theory. In eds. V. Pless and W. Huffman, *Handbook of coding theory, vol. 1*, pp. 649–754. North-Holland, Amsterdam, (1998).
[12] E. Berlekamp, R. McEliece, and H. van Tilborg, On the inherent intractability of certain coding problems, *IEEE Transactions on Information Theory.* **24**, 384–386, (1978).
[13] A. Vardy, The intractability of computing the minimum distance of a code, *IEEE Transactions on Information Theory.* **43**, 1757–1766, (1997).
[14] T. Kløve, *Codes for error detection.* (Series on Coding Theory and Cryptology, vol. 2. World Scientific Publishing Co. Pte. Ltd., Hackensack, 2007).
[15] G. Katsman and M. Tsfasman, Spectra of algebraic-geometric codes, *Problemy Peredachi Informatsii.* **23**, 19–34, (1987).
[16] M. Tsfasman and S. Vlădut, *Algebraic-geometric codes.* (Kluwer Academic Publishers, Dordrecht, 1991).
[17] M. Tsfasman and S. Vlădut, Geometric approach to higher weights, *IEEE Transactions on Information Theory.* **41**, 1564–1588, (1995).
[18] 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).
[19] 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).
[20] V. Wei, Generalized Hamming weights for linear codes, *IEEE Transactions on Information Theory.* **37**, 1412–1418, (1991).
[21] J. Simonis, The effective length of subcodes, *AAECC.* **5**, 371–377, (1993).
[22] L. Ozarev and A. Wyner, Wire-tap channel II, *AT&T Bell Labs Technical Journal.* **63**, 2135–2157, (1984).
[23] G. Forney, Dimension/length profiles and trellis complexity of linear block codes, *IEEE Transactions on Information Theory.* **40**, 1741–1752, (1994).

[24] I. Duursma. Combinatorics of the two-variable zeta function. In eds. G. Mullen, A. Poli, and H. Stichtenoth, *International Conference on Finite Fields and Applications*, vol. 2948, *Lecture Notes in Computer Science*, pp. 109–136. Springer, (2003). ISBN 3-540-21324-4.
[25] T. Brylawski, A decomposition for combinatorial geometries, *Tans. Am. Math. Soc.* **171**, 235–282, (1972).
[26] C. Greene, Weight enumeration and the geometry of linear codes, *Studies in Applied Mathematics.* **55**, 119–128, (1976).
[27] M. Aigner, *Combinatorial theory.* (Springer, New York, 1979).
[28] T. Britz, Extensions of the critical theorem, *Discrete Mathematics.* **305**, 55–73, (2005).
[29] J. van Lint and R. M. Wilson, *A course in combinatorics.* (Cambridge University Press, Cambridge, 1992).
[30] R. Stanley, *Enumerative combinatorics, vol. 1.* (Cambridge University Press, Cambridge, 1997).
[31] A. Skorobogatov. Linear codes, strata of grassmannians, and the problems of segre. In eds. H. Stichtenoth and M. Tfsafsman, *Coding Theory and Algebraic Geometry, Lecture Notes Math. vol 1518*, pp. 210–223. Springer-Verlag, Berlin, (1992).
[32] T. Britz and C. Rutherford, Covering radii are not matroid invariants, *Discrete Mathematics.* **296**, 117–120, (2005).
[33] H. Whitney, On the abstract properties of linear dependence, *American Journal of Mathematics.* **57**, 509–533, (1935).
[34] J. Kung, *A source book in matroid theory.* (Birkhäuser, Boston, 1986).
[35] J. Oxley, *Matroid theory.* (Oxford University Press, Oxford, 1992).
[36] D. Welsh, *Matroid theory.* (Academic Press, London, 1976).
[37] N. White, *Theory of matroids.* (Encyclopedia of Mathmatics and its Applications, vol. 26, Cambridge University Press, Cambridge, 1986).
[38] N. White, *Matroid applications.* (Encyclopedia of Mathmatics and its Applications, vol. 40, Cambridge University Press, Cambridge, 1992).
[39] W. Tutte, Lectures on matroids, *Journal of Research of the National Bureau of Standards, Sect. B.* **69**, 1–47, (1965).
[40] G. Whittle, A charactrization of the matroids representable over GF(3) and the rationals, *Journal of Combinatorial Theory, Ser. B.* **65**(2), 222–261, (1995).
[41] G. Whittle, On matroids representable over GF(3) and other fields, *Transactions of the American Mathematical Society.* **349**(2), 579–603, (1997).
[42] J. Blackburn, N. Crapo, and D. Higgs, A catalogue of combinatorial geometries, *Mathematics of Computation.* **27**, 155–166, (1973).
[43] W. Dukes, on the number of matroids on a fintie set, *Séminaire Lotharingien de Combinatoire.* **51**, Art. B51g, 12 pp., (2004).
[44] D. Knuth, The asymptotic number of geometries, *Journal of Combinatorial Theory, Ser. A.* **16**, 398–400, (1974).
[45] L. Euler, Solutio problematis ad geometriam situs pertinentis, *Commentarii Academiae Scientiarum Imperialis Petropolitanae.* **8**, 128–140, (1736).
[46] N. Biggs, *Algebraic graph theory.* (Cambridge University Press, Cambridge,

1993).
[47] R. Wilson and J. Watkins, *Graphs; An introductory approach.* (J. Wiley & Sons, New York, 1990).
[48] G. Birkhoff, On the number of ways of coloring a map, *Proc. Edinburgh Math. Soc.* **2**, 83–91, (1930).
[49] H. Whitney, A logical expansion in mathematics, *Bulletin of the American Mathematical Society.* **38**, 572–579, (1932).
[50] H. Whitney, The coloring of graphs, *Annals of Mathematics.* **33**, 688–718, (1932).
[51] W. Tutte, A contribution to the theory of chromatic polynomials, *Canadian Journal of Mathematics.* **6**, 80–91, (1954).
[52] W. Tutte, On the algebraic theory of graph coloring, *Journal of Combinatorial Theory.* **1**, 15–50, (1966).
[53] W. Tutte, On dichromatic polynomials, *Journal of Combinatorial Theory.* **2**, 301–320, (1967).
[54] W. Tutte, Cochromatic graphs, *Journal of Combinatorial Theory.* **16**, 168–174, (1974).
[55] W. Tutte, Graphs-polynomials, *Advances in Applied Mathematics.* **32**, 5–9, (2004).
[56] W. Tutte, Matroids and graphs, *Transactions of the American Mathematical Society.* **90**, 527–552, (1959).
[57] C. Athanasiadis, Characteristic polynomials of subspace arrangements and finite fields, *Advances in Mathematics.* **122**, 193–233, (1996).
[58] A. Barg, The matroid of supports of a linear code, *AAECC.* **8**, 165–172, (1997).
[59] T. Britz. *Relations, matroids and codes.* PhD thesis, Univ. Aarhus, (2002).
[60] T. Britz, MacWilliams identities and matroid polynomials, *The Electronic Journal of Combinatorics.* **9**, R19, (2002).
[61] T. Britz, Higher support matroids, *Discrete Mathematics.* **307**, 2300–2308, (2007).
[62] T. Britz and K. Shiromoto, A MacWillimas type identity for matroids, *Discrete Mathematics.* **308**, 4551–4559, (2008).
[63] T. Brylawski and J. Oxley. The Tutte polynomial and its applications. In ed. N. White, *Matroid Applications*, pp. 173–226. Cambridge University Press, Cambridge, (1992).
[64] G. Etienne and M. Las Vergnas, Computing the Tutte polynomial of a hyperplane arrangement, *Advances in Applied Mathematics.* **32**(1), 198–211, (2004).
[65] W. Tutte, A ring in graph theory, *Proc. Cambridge Philos. Soc.* **43**, 26–40, (1947).
[66] W. Tutte. *An algebraic theory of graphs.* PhD thesis, Univ. Cambridge, (1948).
[67] T. Brylawski and J. Oxley, Several identities for the characteristic polynomial of a combinatorial geometry, *Discrete Mathematics.* **31**(2), 161–170, (1980).
[68] T. Kløve, Support weight distribution of linear codes, *Discrete Matematics.*

106/107, 311–316, (1992).
[69] P. Cartier, Les arrangements d'hyperplans: un chapitre de géométrie combinatoire, *Seminaire N. Bourbaki.* **561**, 1–22, (1981).
[70] P. Orlik and H. Terao, *Arrangements of hyperplanes.* vol. 300, (Springer-Verlag, Berlin, 1992).
[71] G.-C. Rota, On the foundations of combinatorial theory I: Theory of möbius functions, *Zeit. für Wahrsch.* **2**, 340–368, (1964).
[72] R. Stanley. An introduction to hyperplane arrangements. In *Geometric combinatorics, IAS/Park City Math. Ser., 13*, pp. 389–496. Amer. Math. Soc., Providence, RI, (2007).
[73] R. Lidl and H. Niederreiter, *Introduction to finite fields and their applications.* (Cambridge University Press, Cambridge, 1994).
[74] H. Crapo and G.-C. Rota, *On the foundations of combinatorial theory: Combinatorial geometries.* (MIT Press, Cambridge MA, 1970).
[75] G. Birkhoff, Abstract linear dependence and lattices, *Amer. Journ. Math.* **56**, 800–804, (1935).
[76] H. Crapo, The Tutte polynomial, *Aequationes Math.* **3**, 211–229, (1969).
[77] E. Mphako, Tutte polynomials of perfect matroid designs, *Combinatorics, Probability and Computing.* **9**, 363–367, (2000).
[78] A. Blass and B. Sagan, Möbius functions of lattices, *Advances in Mathematics.* **129**, 94–123, (1997).
[79] H. Crapo, Möbius inversion in lattices, *Archiv der Mathematik.* **19**, 595–607, (1968).
[80] C. Greene and T. Zaslavsky, On the interpretation of Whitney numbers through arrangements of hyperplanes, zonotopes, non-radon partitions and orientations of graphs, *Transactions of the American Mathematical Society.* **280**, 97–126, (1983).
[81] A. Ashikhmin and A. Barg, Minimal vectors in linear codes, *IEEE Transactions on Information Theory.* **44**(5), 2010–2017, (1998).
[82] J. Massey. Minimal codewords and secret sharing. In *In Proc. Sixth Joint Swedish-Russian Workshop on Information theory, Molle, Sweden*, pp. 276–279, (1993).
[83] D. Stinson, *Cryptography, theory and practice.* (CRC Press, Boca Raton, 1995).
[84] 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).
[85] A. Björner and T. Ekedahl, Subarrangments over finite fields: Chomological and enumerative aspects, *Advances in Mathematics.* **129**, 159–187, (1997).
[86] F. Ardila, Computing the tutte polynomial of a hyperplane arrangement, *Pacific J. Math.* **230**(5), 1–26, (2007).
[87] M. de Boer, Almost MDS codes, *Designs, Codes and Cryptography.* **9**, 143–155, (1996).
[88] T. Brylawski, Intersection theory for graphs, *J. Comb. Theory, Ser. B.* **30** (2), 233–246, (1981).
[89] J. Kung. Twelve views of matroid theory. In eds. K. K. S. Hong, J.H. Kwak

and F. Roush, *Combinnatorial and Computational Mathematics*, pp. 56–96. World Scientific, River Edge, (2001).

[90] A. Sokal. The multivariate Tutte polynomial (alias Potts model) for graphs and matroids. In *Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser. vol. 327*, pp. 173–226. Cambridge University Press, Cambridge, (2005).

[91] G. Farr. Tutte-whitney polynomials: some history and generalizations. In eds. G. Grimmett and C. MacDiarmid, *Combinatorics, Complexity and Chance: A Tribute to D. Welsh*, pp. 28–52. Oxford Univ. Press, Oxford, (2007).

[92] C. Fortuin and P. Kasteleyn, On the random cluster-model. I. Introduction and realtion to other models, *Physica*. **57**, 536–564, (1972).

[93] P. Kasteleyn and C. Fortuin, Phase transitions in lattice systems with random local properties, *J. Phys. Soc. Japan*. **26**, 11–14, (1969).

[94] T. Britz and K. Shiromoto, Designs from subcode support of linear codes, *Designs, Codes and Cryptography*. **46**, 175–189, (2008).

[95] D. Welsh and G. Whittle, Arrangements, channel assignments and associated polynomials, *Advances in Applied Mathematics*. **23**, 375–406, (1999).

[96] J. Kung, Preface: Old and new perspectives on the Tutte polynomial, *Annals of Combinatorics*. **12**, 133–137, (2008).

[97] V. Kook, W. Reiner and D. Stanton, Combinatorial laplacians on matroid complexes, *Journal of the American Mathematical Society*. **13**, 129–148, (2000).

[98] W. Kook, Recurrence relations for the spectrum polynomial of a matroid, *Discrete Applied Mathematics*. **143**, 312–317, (2004).

[99] A. Duval, A common recursion for Laplacians of matroids and shifted simplicial complexes, *Documneta Mathematica*. **10**, 583–618, (2005).

[100] G. Denham, The combinatorial Laplacian of the Tutte complex, *J. Algebra*. **242**(1), 160–175, (2001).

[101] M. Aigner and J. Seidel, Knoten, Spin modelle und Grahen, *Jber. Dt. Math-Verein*. **97**, 75–96, (1995).

[102] L. Kaufmann, *On knots*. (Ann. Math. Stud. 115, Princeton Univ. Press, Princeton, 1987).

[103] D. Welsh, *Complexity: knots, colourings and counting*. (London Mathematical Society Lecture Note Series vol. 186, Cambridge University Press, Cambridge, 1993).

[104] T. Helleseth, The weight distribution of the coset leaders of some classes of codes with related parity-check matrices, *Discrete Mathematics*. **28**, 161–171, (1979).

[105] J. Justesen and T. Høholdt, Bounds on list decoding of MDS codes, *IEEE Transactions on Information Theory*. **47**, 1604–1609, (2001).

[106] M. Sudan, Decoding of reed-solomon codes beyond the error-correction bound, *J. Complexity*. **13**, 180–193, (1997).

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

[108] M. Munuera. Steganography from a coding theory point of view. In ed.

E. Martínez-Moro, *Algebraic Geometry Modeling in Information Theory Cryptography*. World Scientific, River Edge, (2011).

[109] S. Hakami and H. Frank, Cut-set matrices and linear codes, *IEEE Transactions on Information Theory*. **11**, 457–458, (1965).

[110] J. Bruck and M. Naor, The hardness of decoding linear codes with preprocessing, *IEEE Transactions on Information Theory*. **36**(2), 381–385, (1990).

[111] D. MacKay, *Information theory, inference and learning algorithms*. (Cambridge University Press, Cambridge, 2003).

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

[113] F. Harary. A survey of the reconstructing conjecture. In *Lecture Notes in Mathematics, vol. 406*, pp. 18–28, (1974).

[114] T. Brylawski, On the nonreconstructibility of combinatorial geometries, *J. Comb. Theory, Ser. B*. **19**(1), 72–76, (1975).

[115] T. Brylawski. Reconstructing combinatorial geometries. In *Lecture Notes in Mathematics, vol. 406*, pp. 226–235, (1974).

[116] T. Brylawski, Hyperplane reconstruction of the tutte polynomial of a geometric lattice, *Discrete Mathematics*. **35**(1-3), 25–38, (1981).

[117] G.-C. Rota. Combinatorial theory, old and new. In *Proc. Int. Congress Math. 1970 (Nice)*, vol. 3, pp. 229–233, Paris, (1971). Gauthier-Villars.

[118] D. Welsh. Combinatorial problems in matroid theory. In ed. D. Welsh, *Combinatorial mathematics and its applications*, pp. 291–306. Academic Press, London, (1972).

[119] R. Read, An introduction to chromatic polynomials, *Journal of Combinatorial Theory, Series A*. **4**, 52–71, (1968).

[120] M. Aigner. Whitney numbers. In ed. N. White, *Combinatorial geometries, Encyclopedia Math. Appl. vol. 29*, pp. 139–160. Cambridge Univ. Press, Cambridge, (1987).

[121] J. Jaeger, D. Vertigan, and D. Welsh, On the computational complexity of the Jones and Tutte polynomials, *math. Proc. Camb. Phil. Soc.* **108**, 35–53, (1990).

[122] D. Welsh, The computational complexity of knot and matroid polynmials, *Discrete Mathematics*. **124**, 251–269, (1994).

[123] P. K. A. Björklund, T. Husfeldt and M. Koivisto. Computing the Tutte polynomial in vertex-exponential time. In *FOCS*, pp. 677–686. IEEE Computer Society, (2008).

[124] P. Orlik and L. Solomon, Combinatorics and topology of complements of hyperplanes, *Invent. Math.* **56**, 167–189, (1980).

[125] I. Duursma, Weight distributions of geometric Goppa codes, *Transactions of the American Mathematical Society*. **351**, 3609–3639, (1999).

[126] I. Duursma, From weight enumerators to zeta functions, *Discrete Applied Mathematics*. **111**(1-2), 55–73, (2001).