Extended and Generalized Weight Enumerators

Relinde Jurrius and Ruud Pellikaan

Department of Mathematics and Computer Science,
Eindhoven University of Technology, The Netherlands
r.m.p.j.jurrius@tue.nl and g.r.pellikaan@tue.nl

Appeared in:
Proceedings of the International Workshop on Coding and Cryptography
WCC 2009, Ullensvang
May 10-15, Selmer Center, Bergen, pp. 76-91, 2009.

Abstract. This paper gives a survey on extended and generalized weight enumerators of a linear code and the Tutte polynomial of the matroid of the code [16]. Furthermore ongoing research is reported on the coset leader and list weight enumerator and its extensions using the derived code and its arrangement of hyperplanes.

1 Introduction

The weight enumerator of a linear code is a classifying polynomial associated with the code. Besides its intrinsic importance as a mathematical object, it is used in the probability theory around codes. For example, the weight enumerator of a binary code is very useful if we want to study the probability that a received message is closer to a different codeword than to the codeword sent [1]. (Or, rephrased: the probability that a maximum likelihood decoder makes a decoding error.)
We will generalize the weight enumerator in two ways, which lead to polynomials which are better invariants for a code. A procedure for the determination of these polynomials is given. We will show that the two generalizations determine each other, and that they connect to the Tutte polynomial of a matroid, thus linking coding theory and matroid theory. An overview of all the connections is given.
Where the weight enumerator is used for calculating the probability of error-correction, the coset leader weight enumerator does the same for error-detection [20]. We report on ongoing research on the coset leader and list weight enumerator using arrangements of hyperplanes, and consider the connections with other classifying polynomials.

2 Generalized Weight Enumerator

We start with generalizing the weight distribution of C in the following way, first formulated by Kløve [18] and re-discovered by Wei [23]. Let C be a linear $[n, k]$

code over $\mathbb{F}_q$. 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* is equal to n minus the number of coordinates which are zero for every word in the subcode. The smallest weight for which a subcode of dimension r exists, is called the *r-th generalized Hamming weight* of C and denoted by d_r. For each r we can define the *r-th generalized weight distribution* of the code, which forms the coefficients of the following polynomial.

Definition 1. *The* generalized weight enumerator *is given by*

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

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

We can see from this definition that $A_0^0 = 1$ and $A_0^r = 0$ for all $0 < r \leq k$. Furthermore, every 1-dimensional subspace of C contains $q-1$ non-zero codewords, so $(q-1)A_w^1 = A_w$ for $0 < w \leq n$.
We will give a way to determine the generalized weight enumerator of a $[n,k]$ code C over $\mathbb{F}_q$. This method is based on Katsman and Tsfasman [17]. Full proofs are in [16].

Definition 2. *Let* $[n] = \{1, \ldots, n\}$. *For* $J \subseteq [n]$ *we define:*

$$C(J) = \{\mathbf{c} \in C : c_j = 0 \text{ for all } j \in J\}$$
$$l(J) = \dim C(J)$$

The next next two lemmas on determining $l(J)$ will become useful later.

Lemma 1. *Let C be a $[n,k]$ code with generator matrix G. Let G' 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'. Then the dimension $l(J)$ is equal to $k - r(J)$.*

Lemma 2. *Let d and $d^\perp$ be the minimum distance of C and $C^\perp$ respectively, and let $J \subseteq [n]$. Then we have*

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

From the first lemma it follows that $l(J)$ is independent of the choice of the generator matrix G of C. We furthermore introduce the following:

Definition 3. *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 $B_J^r = \begin{bmatrix} l(J) \\ r \end{bmatrix}_q$, the Gaussian binomial. The relation between B_t^r and A_w^r becomes clear in the next proposition.

Proposition 1. *The following formula holds:*

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

Proof: Count in two ways the elements of the set
$\mathcal{B}_t^r = \{(D, J) : J \subseteq [n], |J| = t, D \subseteq C(J) \text{ subspace of dimension } r\}$. □

We can now give the generalized weight enumerator in terms of the B_t^r.

Theorem 1. *The generalized weight enumerator is given by:*

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

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

Proposition 2. *The following formula holds:*

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

We can find back the original weight enumerator by using $W_C(X, Y) = W_C^0(X, Y) + (q-1)W_C^1(X, Y)$.

As an example, we determine the generalized weight enumerators for MDS-codes. By setting $d = n - k + 1$ in Lemma 2, we see that for MDS-codes $l(J)$ only depends on the size of J. Substituting this in Proposition 2 leads to the following theorem.

Theorem 2. *Let C be a linear $[n, k]$ MDS code. Then for $r \leq i \leq k$ the generalized weight distribution is given by*

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

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

3 Extended Weight Enumerator

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

Definition 4. *Let C be a $[n,k]$ 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 $(C \otimes \mathbb{F}_{q^m})(J)$. The following statements are proven similar to the case of the generalized weight enumerator.

Theorem 3. *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 for $0 < w \leq n$

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

Proposition 3. *The following formula holds:*

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

Another way of writing the extended weight enumerator can be very useful for rewriting proofs.

Proposition 4. *The extended weight enumerator of a $[n,k]$ code C can be written as*

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

As we said before, the motivation for looking at the extended weight enumerator comes from the extension codes. The next proposition is already suggested by intuition.

Proposition 5. *Let C be a $[n,k]$ code over $\mathbb{F}_q$. Then*

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

Therefore we can view $W_C(X,Y,T)$ as the weight enumerator of the extension code over the algebraic closure of $\mathbb{F}_q$. This means we can find a relation with the two variable zeta-function of a code, see Duursma [11]. The notion of the extended weight enumerator was first introduced by Helleseth, Kløve and Mykkeltveit [14,18].

4 Connections

There is a connection between the extended and generalized weight enumerator.

Theorem 4. *Let C be a $[n,k]$ code over $\mathbb{F}_q$. Then the extended weight numerator and the generalized weight enumerators are connected via*

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

We need the following proposition to prove the case $T = q^m$.

Proposition 6. *Let C be a $[n,k]$ 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 (sketch): Choose a primitive m-th root of unity $\alpha \in \mathbb{F}_{q^m}$. For a word in $C \otimes \mathbb{F}_{q^m}$, write all the coordinates on the basis $(1, \alpha, \alpha^2, \ldots, \alpha^{m-1})$. This gives an $m \times n$ matrix over $\mathbb{F}_q$ whose rows are words of C. □

Note that this isomorphism depends on the choice of a primitive element α. The use of this isomorphism for the proof of Theorem 4 was suggested in [21] by Simonis. We also need the next subresult, proven with a counting-argument. We will use the following notations:

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

Proposition 7. *Let C be a $[n,k]$ code over $\mathbb{F}_q$. Then the weight numerator of an extension code and the generalized weight enumerators are connected via*

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

This result first appears in [14, Theorem 3.2], although the term "generalized weight enumerator" was yet to be invented. We now get Theorem 4 by Lagrange interpolation. We can also proof the inverse of this relation. As far as the authors are aware, no such formula existed.

Theorem 5. *Let C be a $[n,k]$ code over $\mathbb{F}_q$. Then the generalized weight enumerator and the extended weight enumerator are connected via*

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

5 Connection with Matroid Theory

Matroid theory generalizes the notion of 'independence'. There are many ways to define a matroid: we will use a definition that makes it easy to see that the generator matrix of a linear code gives rise to a matroid.

Definition 5. *A matroid G is a pair (S,I) where S is a finite set and I is a collection of subsets of S called the independent sets, satisfying the following properties:*

1. *The empty set is independent.*
2. *Every subset of an independent set is independent.*
3. *If A and B be two independent sets with $|A| > |B|$, then there exists an $a \in A$ with $a \notin B$ and $B \cup \{a\}$ an independent set.*

A subset of S which is not independent, is called *dependent*. An independent subset of S for which adding an extra element of S always gives a dependent subset, is a maximal independent set or a *basis*. The set of all bases of a matroid is denoted by $\mathcal{B}$. Just as in linear algebra, it can be showed that every basis has the same number of elements. This is called the *rank* of the matroid. We define the rank of a subset of A to be the size of the largest independent set contained in it, and denote this by $r(A)$. The rank function and the set of bases can each be used to determine a matroid completely. Therefore we can define the *dual* of a matroid in the following way:

Definition 6. *Let $G = (S,\mathcal{B})$ be a matroid defined by its set of bases. Then its dual is the matroid $G^* = (S,\mathcal{B}^*)$ with the same underlying set and set of bases*

$$\mathcal{B}^* = \{S - B : B \in \mathcal{B}\}.$$

Note the similarity with the definition of a dual code. For a matroid we also have $(G^*)^* = G$.

Definition 7. *For a matroid G with rank function r the* Tutte polynomial *is defined by*

$$t_G(X,Y) = \sum_{A \subseteq G} (X-1)^{r(G)-r(A)}(Y-1)^{|A|-r(A)} \ .$$

If we have a $[n,k]$ code C over $\mathbb{F}_q$ with generator matrix G, we can interpret the columns of G as a matroid. This matroid does only depend on the code C and not on the generator matrix G, because for every choice of G the dependencies between the columns of G are the same. (In fact, all generalized equivalent codes give rise to the same matroid.) Using Lemma 1 we can rewrite the Tutte polynomial as follows.

Proposition 8. *Let C be a $[n,k]$ code over $\mathbb{F}_q$ with generator matrix G. Then the Tutte polynomial associated with the code C is*

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

This formula and Proposition 4 suggest a connection between the weight enumerator and the Tutte polynomial.

Theorem 6. *Let C be a $[n,k]$ code over $\mathbb{F}_q$ with generator matrix G. Then the following holds for the Tutte polynomial and the extended weight enumerator:*

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

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

Theorem 7. *Let C be a $[n,k]$ code over $\mathbb{F}_q$ with generator matrix G. Then the following holds for the extended weight enumerator and the Tutte polynomial:*

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

Greene [12] already showed that the Tutte polynomial determines the weight enumerator, but not the other way round. By using the extended weight enumerator, we get a two-way equivalence and the proof reduces to rewriting.

We can also give expressions for the generalized weight enumerator in terms of the Tutte polynomial, and the other way round. The first formula was found by Britz [8] and independently by Jurrius [16].

Theorem 8. *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*

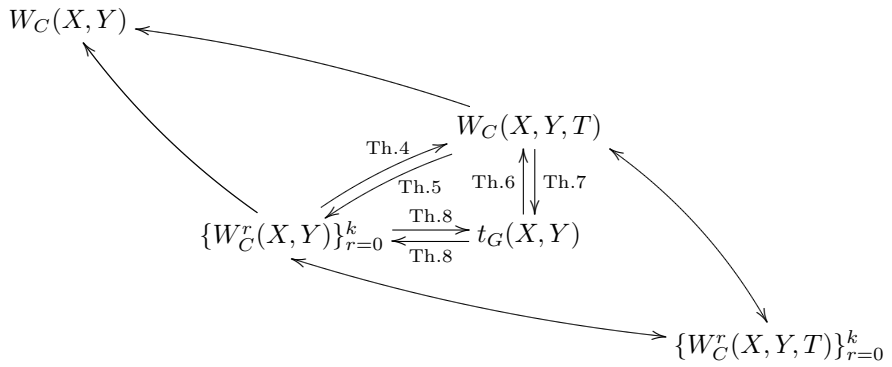
$$\frac{1}{\langle r \rangle_q} \sum_{j=0}^{r} \begin{bmatrix} r \\ j \end{bmatrix}_q (-1)^{r-j} q^{\binom{r}{j}} (X-Y)^k Y^{n-k} \; t_G\left(\frac{X+(q^j-1)Y}{X-Y}, \frac{X}{Y}\right) .$$

And, conversely,

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

6 Overview

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



We see that the Tutte polynomial, the extended weight enumerator and the collection of generalized weight enumerators all contain the same amount of information about a code, because they completely define each other. The original weight enumerator $W_C(X,Y)$ contains less information and therefore does not determine $W_C(X,Y,T)$ or $\{W_C^r(X,Y)\}_{r=0}^k$. See Simonis [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 then the three underlying polynomials.
It was shown by Gray [9] that the matroid of a code is a stronger invariant than its Tutte polynomial.

7 MacWilliams Relations

The Tutte polynomial, the extended and the generalized weight enumerator completely define this polynomials for their dual structures. Because we can switch

freely between polynomials, proofs will become a lot easier. We start with the shortest relation, which can be proven using only basic matroid theory.

Theorem 9. *Let $t_G(X,Y)$ be the Tutte polynomial of a matroid G, and let G^* be the dual matroid. Then $t_G(X,Y) = t_{G^*}(Y,X)$.*

With this theorem, Greene [12] proved the MacWilliams relations for the weight enumerator of a code. This can easily be extended to the extended weight enumerator:

Theorem 10 (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:*

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

For the generalized weight enumerator, we have the following.

Theorem 11. *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:*

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

This theorem was first found by Kløve [19], but the proof is reduced to rewriting by the two way equivalence between the generalized and extended weight enumerator. Another form of this formula was found by Simonis [21] and re-proven by Barg [2] and Britz [7] using matroid theory. The equivalence with the previous theorem can also be proven directly.

Theorem 12. *The generalized Hamming weights A^r_w of a $[n,k]$ code C over $\mathbb{F}_q$ and the generalized Hamming weights $\tilde{A}^r_w$ of the dual code $C^\perp$ are related via*

$$\sum_{i=0}^{n} \binom{n-i}{n-m} \tilde{A}^r_i = \sum_{l=0}^{r} q^{l(m-k+l-r)} \begin{bmatrix} m-k \\ r-l \end{bmatrix}_q \sum_{v=0}^{n} \binom{n-v}{m} A^l_v \; .$$

Using our terminology with B^r_t, the theorem takes the following form. The shortness of this theorem gives another motivation for the use of the B^r_t in defining the extended weight enumerator.

Theorem 13. *Let C be a $[n,k]$ code over $\mathbb{F}_q$, with the associated B^r_t. Let $C^\perp$ be its dual, with associated $\tilde{B}^r_t$. Then the B^r_t and $\tilde{B}^r_t$ completely define each other in the following way:*

$$\tilde{B}^r_t = \sum_{l=0}^{r} q^{l(n-t-k+l-r)} \begin{bmatrix} n-t-k \\ r-l \end{bmatrix}_q B^l_{n-t} \; .$$

8 Another Algorithm

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 arbitrary matrices, not necessarily of full rank. With the help of the following definition, we will give another way to determine the extended weight enumerator. The method is based on matrix decomposition like in Greene [12], and by the previous connections it turned out to be similar to Britz et al. [6].

Definition 8. *Let G be an $k \times n$ matrix over $\mathbb{F}_q$, not necessarily of full rank and without zero columns. Then for each $J \subseteq [n]$ we define $l(J) = k - r(J)$ as in Lemma 1, and the extended weight enumerator $W_G(X, Y, T)$ as in Definition 4.*

We make the following remarks about $W_G(X, Y, T)$.

Proposition 9. *Let G be a $k \times n$ matrix over $\mathbb{F}_q$ 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, permutation of the columns of G and multiplying a column of G with an element of $\mathbb{F}_q^*$.*

(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) *If G is the direct sum of G_1 and G_2, that is of the form*

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

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

We use the extended weight enumerator for general matrices to derive a recursive algorithm for determining the extended weight enumerator of a code. If G is a $k \times n$ matrix, we denote by G^* a matrix which is row-equivalent to G and has a column a of the form $(1, 0, \ldots, 0)^T$. In general, this reduction G^* is not unique. The matrix $G^* - a$ 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. Their extended weight enumerators are connected via

Proposition 10. *For the extended weight enumerator of a reduced matrix G^* holds*

$$W_{G^*}(X, Y, T) = (X - Y) W_{G^*/a}(X, Y, T) + Y W_{G^* - a}(X, Y, T) .$$

By recursively using the previous proposition, we get the next way to determine the extended weight enumerator of a code.

Theorem 14. *Let G be a $k \times n$ matrix over $\mathbb{F}_q$ with $n > k$ of the form $G^* = (I_k | P)$, where P is a $k \times (n-k)$ matrix over $\mathbb{F}_q$. Let $A \subseteq [k]$ and write P_A for the matrix formed by the rows of P indexed by A. Then the following holds:*

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

Using this theorem we can formulate the next algorithm for determining the extended weight enumerator.

Procedure Extended Weight Enumerator
Input: $k \times n$ matrix G over $\mathbb{F}_q$
Output: extended weight enumerator $W_G(X, Y, T)$

1. If $n = 1$ then $W_G(X, Y, T) = T^{k-1}(X + (T-1)Y)$.
2. By Gaussian elimination and deleting any zero rows and columns, write $G^* = (I_{k'}|P)$ of full rank with $k' \leq k$ and length $n' \leq n$.
3. If $k' = n'$ then $W_{G^*}(X, Y, T) = (X + (T-1)Y)^{n'}$.
4. If $k' = 1$, then $W_{G^*}(X, Y, T) = X^{n'} + (T-1)X^{n'-w}Y^w$.
5. If $k' < \frac{1}{2}n'$, switch to $(I_{n'-k'}|P^T)$ and use the MacWilliams relations after step 7.
6. For $A = \emptyset$ set $W_{P_A}(X, Y, T) = X^{n'}$. Then go to step 1 for all P_A with $A \subseteq [k']$ and $|A| > 0$.
7. Use Theorem 14 to determine $W_{G^*}(X, Y, T)$.
8. Now $W_G(X, Y, T) = T^{k-k'}X^{n-n'}W_{G^*}(X, Y, T)$.

9 Complexity

In general, computing the weight enumerator of a code is NP-hard [3,4,22]. We have discussed multiple ways to determine the extended weight enumerator of a linear code. In this section, we look at the complexity of these calculations.

Definition 9. *The complexity* $\mathrm{Comp}(n, R)$ *of the calculation of the weight enumerator of a linear* $[n, k]$ *code over* $\mathbb{F}_q$ *with information rate* $R = \frac{k}{n}$ *is given as a function of* n *and* R. *The* exponent *of this function is defined as*

$$E_q(R) = \lim_{n \to \infty} \frac{\log_2 \mathrm{Comp}(n, R)}{n}.$$

Remark that the complexity also depends on the used algorithm. Which algorithm is used, should be clear from the context.

The most straightforward way to calculate the weight enumerator is the brute force-method: we simply go through all words and determining their weight. There are q^k words of length n, so the complexity of this brute force-method is nq^k and therefore $E_q(R) = \log_2 q \cdot \min\{R, 1-R\}$ because of the MacWilliams relations.
To find the extended weight enumerator by using the brute force method, we have to calculate the weight distribution of $k+1$ extension codes and use Lagrange interpolation to find the coefficients of the polynomial. The fastest way to do this is to take the first k extension codes and the zero code. This gives for the complexity $n(1+q^k+(q^2)^k+(q^3)^k+\ldots+(q^k)^k)$ and therefore $E_q(R) \to \infty$.

If we use the $B_t(T)$ to determine the weight enumerator, we have to look at all $J \subseteq [n]$ and determine the dimension of $C(J)$. Therefore the complexity is $\mathcal{O}(2^n \cdot n^3)$ and $E_q(R) = 1$. So if we just want to know the weight enumerator, this method is faster than the brute force-method if $\log_q 2 < \min\{R, 1-R\} < \frac{1}{2}$, so $q > 4$. Unfortunately, in practice we often use binary codes, so $q = 2$.
On the other hand, using the $B_t(T)$ gives us the generalized extended weight enumerator, not only the ordinary weight enumerator. That means we also determine the generalized weight enumerators $\{W_C^r(X, Y, T)\}_{r=0}^k$ and the weight enumerator $W_{C \otimes \mathbb{F}_{q^m}}(X, Y)$ of all extension codes of C. If that is the goal, the method using B_t^r is much faster than the brute force method.

We can also determine the (extended) weight enumerator using the algorithm in section 8. By Björklund et al. [5] we find the complexity of the deletion–contraction algorithm for computing the Tutte polynomial of a graph. Translating this to codes, we find the complexity is $\mathcal{O}(3^{n-k})$ and thus $E_q(R) = (1-R)\log_2 3$. This is faster than our algorithm.

10 Coset Leader and List Weight Enumerator

The following is a summary of [15]

Definition 10. *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. 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, that is $\mathrm{wt}(\mathbf{y}) = \mathrm{wt}(\mathbf{y} + C)$. 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*

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

See [20,13]. The covering radius *$\rho(C)$ of C is the maximal i such that $\alpha_{C,i} \neq 0$.*

For instance $\alpha_i = \binom{n}{i}(q-1)^i$ 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 *probability of error*, 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. See [20, Chap.1 §5]. The list weight enumerator is of interest in case the decoder has as output the list of all nearest codewords.

Although the generalized weight enumerator, the Tutte polynomial and the matroid of a code contain a lot of information of a code, they do not determine

the coset leader weight enumerator or even the covering radius of a code. See [9]. For instance all $[n, k, n-k+1]$ codes over $\mathbb{F}_q$ are MDS and have the same generalized weight enumerator, uniform matroid and Tutte polynomial but the covering radius varies for fixed n, k and q with $n > q$.

There is a one-to-one correspondence between cosets and syndromes. It is a well known fact that a coset leader corresponds to a minimal way to write its syndrome as a linear combination of the columns of a parity check matrix. This idea is formalized as follows.

Definition 11. *Let H be a parity check matrix of a $[n, k]$ code C over $\mathbb{F}_q$. Let $\mathbf{y}$ be a received word. Then $\mathbf{s} = H\mathbf{y}^T$ is the* syndrome *of this word with respect to H. Define the* weight of $\mathbf{s}$ with respect of H *also called the* syndrome weight *of $\mathbf{s}$, by $\mathrm{wt}_H(\mathbf{s}) = \mathrm{wt}(\mathbf{y} + C)$.*

Then α_i is the number of syndromes in $\mathbb{F}_q^{n-k}$ with respect to H that are of weight i. See [13, Definition 2.1]. Note that the correspondence between cosets en syndromes depends on the choice of H, but that the structure of the cosets themselves is independent of H.

Definition 12. *Let $\mathbf{h}_j$ be the j-th column of H. Let $J \subseteq [n]$. Let V_J be the vector subspace of $\mathbb{F}_q^{n-k}$ that is generated by the vectors $\mathbf{h}_j^T$, $j \in J$. Let*

$$V_t = \bigcup_{|J|=t} V_J \ .$$

Proposition 11. *Let $\mathbf{s}$ in $\mathbb{F}_q^{n-k}$ be a syndrome with respect to H. Then*

$$\mathrm{wt}_H(\mathbf{s}) = t \quad \textit{if and only if} \quad \mathbf{s} \in V_t \setminus V_{t-1} \ .$$

Let $J \subseteq [n]$ consist of t elements. If V_J has dimension t', then there is a $J' \subseteq J$ consisting of t' elements such that $\mathbf{h}_i$, $i \in J'$ are independent. So $V_J = V_{J'}$. Now V_J is a subspace of the column space of H, which has dimension $n-k$. Hence there is an $I \subseteq [n]$ consisting of $n-k$ elements such that and $J' \subseteq I$ and $\mathbf{h}_i$, $i \in I$ are independent. So

$$V_J = \bigcap_{i \in (I \setminus J')} V_{I \setminus \{i\}} \ .$$

is an intersection of the $n-k-t'$ hyperplanes $V_{I\setminus\{i\}}$.

Consider the *arrangement* $\mathcal{V}_C$ of all hyperplanes V_I where $I \subseteq [n]$ consists of $n-k-1$ elements such that the $\mathbf{h}_i$, $i \in I$ are independent. This arrangement gives the *derived code* code $D(C)$ with generator matrix $D(G)$ whose columns correspond to the hyperplanes of the arrangement and where the entries of a column correspond to the coefficients of the defining equation of the hyperplane.

An inclusion/exclusion counting argument, or more abstractly the characteristic polynomial of the *geometric lattice* of the arrangement $\mathcal{V}_C$, see [10], gives that there are polynomials $\alpha_i(T)$ and $\lambda_i(T)$ of degree i 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 extension code $C \otimes \mathbb{F}_{q^m}$. Then

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

are the *extended coset leader weight enumerator* and the *extended list weight enumerator*, respectively. For instance $\lambda_i(T) = \alpha_i(T) = \binom{n}{i}(T-1)^i$ for all $i \leq (d-1)/2$, where d is the minimum distance of C. Let $i(C)$ be the number of information sets of C. Then $\lambda_{n-k}(T) = i(C)\alpha_{n-k}(T)$.

Example 1. Let $C = \mathbb{F}_q^n$. Then $\lambda_C(X,Y,T) = \alpha_C(X,Y,T) = X^n$.

Example 2. Let $C = \{0\}$. Then $\lambda_i(T) = \alpha_i(T) = \binom{n}{i}(T-1)^i X^{n-i} Y^i$ and $\lambda_C(X,Y,T) = \alpha_C(X,Y,T) = (X + (T-1)Y)^n$.

Example 3. Let C be the binary Hamming code of length 7. Then $\lambda_i(T) = \alpha_i(T)$ for $i \leq 1$, and $\alpha_0(T) = 1$, $\alpha_1(T) = 7(T-1)$, $\lambda_2(T) = 3\alpha_2(T) = 21(T-1)(T-2)$ and $\lambda_3(T) = 28\alpha_3(T) = 28(T-1)(T-2)(T-4)$.
So $\rho(C) = 1$, $\rho(C \otimes \mathbb{F}_4) = 2$ and $\rho(C \otimes \mathbb{F}_{2^m}) = 3$ for $m \geq 3$.

Example 4. Let C be the dual of the $[n,1,n]$ repetition code. Then
$\lambda_C(X,Y,T) = X^n + n(T-1)X^{n-1}Y$ and $\alpha_C(X,Y,T) = X^n + (T-1)X^{n-1}Y$.

Example 5. Let C be the $[n,1,n]$ repetition code. Then this code has not such an easy description of $\lambda_C(X,Y,T)$ and $\alpha_C(X,Y,T)$ as the previous example. Apart from the known expressions for $\lambda_i(T)$ and $\alpha_i(T)$ for $i \leq (n-1)/2$ that hold for every code we have that $\alpha_{n-1}(T) = (T-1)(T-2)\cdots(T-n+1)$ and $\lambda_{n-1}(T) = n\alpha_{n-1}(T)$. We will determine the formulas for all i in further research.

We will compute the extended coset leader weight enumerator and the extended list weight enumerator of several classes of codes such Simplex, Hamming and Golay codes, MDS codes and Reed-Solomon codes, and cyclic codes.

Research Problem 5.1 in [20, Chapter 5] asked whether the coset leader weight enumerator of C determines the coset leader weight enumerator of $C^\perp$, as is the case for the ordinary weight enumerator by the MacWilliams relations. This problem has a negative answer by [1]. We can also give a counter example for the extended coset leader weight enumerator.

Theorem 15. *The extended coset leader weight enumerators of C and $C^\perp$ do not determine each other.*

Theorem 16. *The extended coset leader weight enumerator $\alpha_C(X, Y, T)$ does not determine the generalized weight enumerator $W_C(X, Y, T)$ of a code, nor the extended list weight enumerator $\lambda_C(X, Y, T)$.*

Both theorems follow from a counter example. Let C_1 and C_2 be the two $[6, 3]$ codes over $\mathbb{F}_2$ with generator matrices

$$\begin{pmatrix} 1\,0\,0\,1\,0\,1 \\ 0\,1\,0\,0\,1\,1 \\ 0\,0\,1\,0\,0\,1 \end{pmatrix}, \qquad \begin{pmatrix} 1\,0\,0\,1\,0\,1 \\ 0\,1\,0\,0\,1\,1 \\ 0\,0\,1\,1\,0\,0 \end{pmatrix}.$$

The next table shows the coefficients of the extended coset leader weight enumerator of the codes and their duals, the extended list weight enumerator, and the extended weight enumerator. The values for $i = 0$ are left out, since they are all equal to 1 because of the zero word.

	i	C_1	C_2
$\alpha_{C,i}$	1	$5(T-1)$	$5(T-1)$
	2	$2(T-1)(3T-5)$	$2(T-1)(3T-5)$
	3	$(T-1)(T-2)(T-3)$	$(T-1)(T-2)(T-3)$
$\alpha_{C^\perp,i}$	1	$4(T-1)$	$5(T-1)$
	2	$3(T-1)(2T-3)$	$2(T-1)(3T-5)$
	3	$(T-1)(T-2)(T-3)$	$(T-1)(T-2)(T-3)$
$\lambda_{C,i}$	1	$6(T-1)$	$6(T-1)$
	2	$2(T-1)(7T-12)$	$2(T-1)(7T-11)$
	3	$12(T-1)(T-2)(T-3)$	$13(T-1)(T-2)(T-3)$
$\lambda_{C^\perp,i}$	1	$6(T-1)$	$6(T-1)$
	2	$13(T-1)^2$	$2(T-1)(7T-11)$
	3	$12(T-1)(T-2)(T-3)$	$13(T-1)(T-2)(T-3)$
$A_{C,i}$	1	0	0
	2	$(T-1)$	$(T-1)$
	3	$4(T-1)$	$3(T-1)$
	4	$(T-1)(2T-3)$	$T(T-1)$
	5	$2(T-1)(T-2)$	$(T-1)(4T-7)$
	6	$(T-1)(T^2-3T+3)$	$(T-1)(T-2)^2$

We see that the extended coset leader weight enumerator of the two codes are equal, but none of the other polynomials, so they are not defined by the extended coset leader weight enumerator.

We will research whether the extended list weight enumerator of C determines the corresponding enumerator of $C^\perp$, and if the extended list weight enumerator determines the extended weight enumerator.

References

1. Baicheva, T., Bouyukliev, I., Dodunekov, D., Willems, W.: Teaching linear codes. In: Int. Con. MASSEE (2003)

2. Barg, A.: The matroid of supports of a linear code. AAECC 8, 165–172 (1997)
3. Barg, A.: Complexity issues in coding theory. In: Pless, V.S., Hu?man, W.C. (eds.) Handbook of coding theory, vol. 1, pp 649–754. North-Holland, Amsterdam (1998)
4. Berlekamp, E.R., McEliece, R.J., Van Tilborg, H.C.A.: On the inherent intractability of certain coding problems. IEEE Trans. on Inf. Th. 24, 384–386 (1978)
5. Björklund, A., Husfeldt, T., Kaski, P., Koivisto, M.: Computing the Tutte Polynomial in Vertex-Exponential Time. In: 49th Annual IEEE Symposium on Foundations of Computer Science, pp 677–686. IEEE Computer Society (2008)
6. Britz, D., Britz, T., Shiromoto, K., Sørensen, H.K.: The higher weight enumerators of the douby-even, self-dual $[48, 24, 12]$ code. IEEE Trans. on Inf. Th. 53, 2567–2571 (2007)
7. Britz, T.: MacWilliams identities and matroid polynomials. El. J. of Com. 9, R19 (2002)
8. Britz, T.: Higher support matroids. Discr. Math. 307, 2300–2308 (2007)
9. Britz, T., Rutherford, C.G.: Covering radii are not matroid invariants. Discr. Math. 296, 117–120 (2005)
10. Cartier, P.: Les arrangements d'hyperplans: un chapitre de géométrie combinatoire. Sem. N. Bourbaki 561, 1–22 (1981)
11. Duursma, I.M.: Combinatorics of the two-variable zeta function. In: Finite Fields and Applications, pp. 1–19. Springer-Verlag, Berlin (2004)
12. Greene, C.: Weight enumeration and the geometry of linear codes. St. in App. Math. 55, 119–128 (1976)
13. Helleseth, T.: The weight distribution of the coset leaders of some classes of codes with related parity-check matrices. Discr. Math. 28, 161–171 (1979)
14. Helleseth, T., Kløve, T., Mykkeltveit, J.: The weight distribution of irreducible cyclic codes with block lengths $n_1 = ((q^l - 1)/n)$. Discr. Math. 18, 179–211 (1977)
15. Jurrius, R., Pellikaan, R.: The extended coset leader weight enumerator. To appear in: 30th Symposium on Information Theory in the Benelux (2009)
16. Jurrius, R.P.M.J.: Classifying polynomials of linear codes. Masters thesis, Leiden University, 2008.
17. Katsman, G.L., Tsfasman, M.A.: Spectra of algebraic-geometric codes. Prob. Per. Inf. 23, 19–34 (1987)
18. Kløve, T.: The weight distribution of linear codes over $\mathrm{GF}(q^l)$ having generator matrix over $\mathrm{GF}(q)$. Discr. Math. 23, 159–168 (1978)
19. Kløve, T.: Support weight distribution of linear codes. Discr. Math. 106/107, 311–316 (1992)
20. MacWilliams, F.J., Sloane, N.J.A.: The Theory of Error-Correcting Codes. North-Holland Mathematical Library, Amsterdam (1977)
21. Simonis, J.: The effective length of subcodes. AAECC 5, 371–377 (1993)
22. Vardy, A.: The intractability of computing the minimum distance of a code. IEEE Trans. on Inf. Th. 43, 1757–1766 (1997)
23. Wei, V.: Generalized Hamming Weights for Linear Codes. IEEE Trans. on Inf. Th. 37, 1412-1418 (1991)