On the Weight Distribution of Elliptic Codes

MOHAMMAD AMIN SHOKROLLAHI
Institut für Informatik V, Universität Bonn
Römerstraße 164, D-5300 Bonn

September 6, 1990

Abstract

We discuss the problem of determining the weight distribution of linear codes on elliptic curves. It is shown that the weight enumerators of non-MDS elliptic codes are completely determined by the number of minimum weight codewords. Using the structure of the group of rational points of the curves in question, we give a procedure to determine the number of minimum weight codewords. For two classes of elliptic codes we shall give explicit formulas for this number. Several examples are presented.

1 Introduction

One of the problems of coding theory is to determine for a given linear code C its *weight enumerator* which is defined as follows: If C is an (n, k, d)-code, i.e. it has block length n, dimension k and minimum distance d, then for each w with $0 \leq w \leq n$ let $A_w(C)$ denote the number of codewords in C which have Hamming-weight equal to w (i.e. which have nonvanishing coordinates at exactly w positions). It is customary to replace $A_w(C)$ simply by A_w if there is no danger of confusion. Then the weight enumerator $A_C(z)$ of C is defined by $A_C(z) := \sum_{w=0}^{n} A_w z^w \in \mathbf{Z}[x]$.

The importance of the weight enumerator of a linear code is partly due to the fact that the weight enumerator of a (linear) code C is directly related to the probability of decoding failure of C [1, Chapter 16].

The problem of determining the weight enumerator of an arbitrary linear code which is e.g. given by its generator matrix seems to be very hard. Since the minimum distance of a code can be easily derived from its weight

enumerator, the determination of the latter is at least as hard as the determination of the minimum distance which is conjectured to be NP-complete in the case of linear codes. (To be more accurate, the corresponding decision problem is conjectured to be NP-complete [2, 11]).

Surprisingly there are classes of linear codes, for which the weight enumerators are known. For the class of MDS-codes there exist explicit formulae for the coefficients of the weight enumerators [5, Chapter 11]. (Recall that an (n, k, d)-code C is called MDS, if $d = n - k + 1$.)

In this paper we are going to investigate the problem of computing the weight enumerator of *elliptic codes*. By elliptic codes we mean linear codes constructed on elliptic curves over finite fields using Goppa's construction of algebraico-geometric codes [3, 7]. It turns out that elliptic codes are *almost* MDS: If the elliptic code C is an (n, k, d)-code, then either $d = n - k$ or $d = n - k + 1$. In the latter case C is of course MDS. The computation of the exact minimum distance of an elliptic code over the finite field $\mathbf{F}_q$ leads to the investigation of the group of $\mathbf{F}_q$-rational points of the underlying elliptic curve [7].

The fact that the dual of an elliptic code C is again an elliptic code can be used to compute the dual distance of C, i.e. the minimum distance of the dual $C^\perp$ of C. Since the duals of MDS-codes are again MDS, it turns out that the dual distance of an $(n, k, n - k)$-elliptic-code is k (Lemma 1). Once the dual distance of C is known, one can use the MacWilliams-identities in order to give explicit formulae for the coefficients of the weight enumerator of C. The method involved is similar to the method for computing the weight enumerator of an MDS-code as is done in [5, Chapter 11]. It turns out that the weight enumerator of an elliptic $(n, k, n - k)$-code is completely determined by the number of minimum weight codewords, that is codewords of weight $n - k$. The problem of determining the weight distribution of an elliptic code thus leads to the problem of computing the number of its minimum weight codewords. As a consequence of this one gets the assertion that elliptic codes of even block-length $2k$ and dimension k are *formally selfdual* which means that their weight enumerators coincide with those of their duals.

The method used so far works for every linear $(n, k, n - k)$-code with dual distance k. The determination of the number of minimum weight codewords now calls for specific properties of elliptic codes. One of the properties of elliptic codes one can use is the fact that the rational points of the underlying

elliptic curve form an abelian group. With the aid of this so-called *class group* of the curve the problem of determining the number of minimum weight codewords can be interpreted as a combinatorial problem in the theory of finite abelian groups. To be more specific, for a given finite additive abelian group $\mathfrak{G}$, a given subset M of $\mathfrak{G}$, a given element $g \in \mathfrak{G}$ and a given integer j satisfying $1 \leq j < |M|$ we are interested in the number of j-subsets T of M such that the sum of the elements of T equals g. A constrained version of this problem is the case where $M = \mathfrak{G}$. Although the general problem seems to be more difficult than the constrained problem, it can be shown that one can solve the general problem for given $\mathfrak{G}$ for every subset M, every j and every $g \in \mathfrak{G}$ if one can solve the constrained problem for every j and every $g \in \mathfrak{G}$ (Theorem 6).

The above constrained problem is closely related to a series of elliptic codes which we have called *elliptic codes of Type I*. We shall give explicit formulae for the solution of the constrained problem in the case where n and j are coprime (Section 4). This leads to the number of minimum weight codewords of elliptic code of Type I in certain cases. The case $M = \mathfrak{G} \setminus \{0\}$ leads to the so-called *elliptic codes of Type II*. These are considered in Section 5.

The paper is organized as follows: Section 2 contains a brief introduction to elliptic codes and curves. Section 3 discusses the problem of the computation of the weight enumerators of elliptic codes. Here explicit formulae for the coefficients of the weight enumerators are derived. This Section also discusses the relationship between the number of minimum weight codewords and the above mentioned combinatorial problem. Sections 4 and 5 are devoted to the investigation of elliptic codes of Type I and Type II, respectively. In Section 6 we show how the solution of the constrained problem leads to the solution of the general problem (in the above terminology). The last section presents some examples how to compute the weight enumerators with the help of the theorems presented in this paper.

2 A Review of Elliptic Curves and Codes

In this section we are going to give a brief summary of elliptic curves and linear codes arising from them. A more thorough treatment can be found in [7]. For general facts about elliptic curves and function fields the reader

can consult [4, 8].

Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbf{F}_q$. Thus, $\mathcal{E}$ is an irreducible smooth algebraic curve of genus one. Let us denote by $\mathcal{E}(L)$ the set of L-rational points of $\mathcal{E}$, where L is an algebraic extension of $\mathbf{F}_q$. Further, let $\mathbf{P}(\mathcal{E})$ denote the set of prime divisors of $\mathcal{E}$. Since $\mathcal{E}$ is assumed to be smooth, there is a bijection between $\mathbf{P}(\mathcal{E})$ and $\mathcal{E}(\overline{\mathbf{F}_q})$, where $\overline{\mathbf{F}_q}$ denotes the algebraic closure of $\mathbf{F}_q$. There is a map $\deg\colon \mathbf{P}(\mathcal{E}) \to \mathbf{Z}$: for $P \in \mathbf{P}(\mathcal{E})$ $\deg(P)$ is defined to be the (absolute) residue class degree of P. Equivalently, using the mentioned bijection, one can say that $\deg(P)$ is the degree of the smallest field extension L of $\mathbf{F}_q$ such that P lies in $\mathcal{E}(L)$.

The group $\mathbf{D}(\mathcal{E})$ of divisors on $\mathcal{E}$ is defined to be the free abelian group generated over $\mathbf{P}(\mathcal{E})$. Each divisor D on $\mathcal{E}$ has thus a representation of the form $D = \sum_{P \in \mathbf{P}(\mathcal{E})} d_P P$ where the d_P are integers vanishing for almost all $P \in \mathbf{P}(\mathcal{E})$. There is a partial order $\leq$ on $\mathbf{D}(\mathcal{E})$: if $C = \sum_{P \in \mathbf{P}(\mathcal{E})} c_P P$ is another divisor, then $D \leq C$ if and only if for all $P \in \mathbf{P}(\mathcal{E})$ we have $d_P \leq c_P$. The divisor D is called *integral* if $d_P \geq 0$ for all P. Each divisor D has a unique decomposition $D = D_0 - D_\infty$ with integral divisors D_0 and D_∞. There is a special kind of divisors called *principal divisors* on $\mathcal{E}$: for each non-vanishing function f on $\mathcal{E}$, the divisor $(f) := \sum_{P \in \mathbf{P}(\mathcal{E})} \mathrm{ord}_P(f) P$ is well defined and called *the principal divisor of* f. The principal divisors form a subgroup $\mathbf{H}(\mathcal{E})$ of $\mathbf{D}(\mathcal{E})$ which is isomorphic to $K^\times / \mathbf{F}_q{}^\times$ where K is the function field of $\mathcal{E}$. They give rise to a partition on $\mathbf{D}(\mathcal{E})$: The divisors C and D are called *equivalent*, if there is a function f such that $C = D + (f)$. The classes under this equivalence relation are called *divisor classes*. The map deg extends to $\mathbf{D}(\mathcal{E})$ by $\deg(D) := \sum_{P \in \mathbf{P}(\mathcal{E})} d_P \deg(P)$. The kernel $\mathbf{D}_0(\mathcal{E})$ of deg contains the group $\mathbf{H}(\mathcal{E})$ and the quotient $\mathbf{D}_0(\mathcal{E})/\mathbf{H}(\mathcal{E})$ is called the *class group* of $\mathcal{E}$. The degree map is thus constant on divisor classes.

The linear space $\mathcal{L}(D)$ attached to D is defined to be the union of $\{0\}$ and the set of all functions f in $K^\times$ which satisfy $(f) \geq -D$. $\mathcal{L}(D)$ is a vector space over $\mathbf{F}_q$ of finite dimension $\dim(D)$. The determination of $\dim(D)$ for all divisors D is the problem of Riemann-Roch, the nontrivial case being the case where $\deg(D) > 0$. (For $\deg(D) < 0$ it is easily seen that $\dim(D) = 0$. If $\deg(D) = 0$, then $\dim(D) = 1$ if D is a principal divisor and $\dim(D) = 0$ otherwise, a fact which can also be seen easily.) By the Theorem of Riemann-Roch we have $\dim(D) = \deg(D)$ if $\deg(D) \geq 1$.

If $P \in \mathcal{E}(\mathbf{F}_q)$ denotes the neutral element of $\mathcal{E}(\mathbf{F}_q)$ and $G \in \mathbf{D}(\mathcal{E})$, we define the *trace* $\mathrm{Tr}(G)$ of a divisor G of nonnegative degree to be the (unique!)

divisor in $\mathcal{E}(\mathbf{F}_q)$ such that $G - \deg(G)P$ and $\mathrm{Tr}(G) - P$ belong to the same class.

The Theorem of Riemann-Roch establishes a bijection between the class group of $\mathcal{E}$ and the set $\mathcal{E}(\mathbf{F}_q)$. This bijection carries over the structure of $\mathbf{D}_0(\mathcal{E})/\mathbf{H}(\mathcal{E})$ to $\mathcal{E}(\mathbf{F}_q)$. Thus, $\mathcal{E}(\mathbf{F}_q)$ is equipped with the structure of an abelian group. The operation in this group is usually written additively. In order to distinguish between the addition in the divisor group and the inherited addition in $\mathcal{E}(\mathbf{F}_q)$, we denote the latter by $\oplus$. Stated in terms of divisors, the addition $\oplus$ (also called o-addition) goes as follows: fix an element $P \in \mathcal{E}(\mathbf{F}_q)$. For $P_1, P_2 \in \mathcal{E}(\mathbf{F}_q)$ there is by the Theorem of Riemann-Roch a (unique) divisor $P_3 \in \mathcal{E}(\mathbf{F}_q)$, such that $P_1 + P_2 - 2P$ and $P_3 - P$ belong to the same class. One defines $P_1 \oplus P_2 := P_3$. Fixing a model for $\mathcal{E}$, one can also give explicit formulae for the addition of points. These can be found in standard textbooks such as [4, 8].

Linear codes over the curve $\mathcal{E}$ are constructed in the following way: take n different points $P_1, \ldots, P_n \in \mathcal{E}(\mathbf{F}_q)$ and a divisor $G \in \mathbf{D}(\mathcal{E})$ subject to the condition $\mathrm{ord}_{P_1}(G) = \ldots = \mathrm{ord}_{P_n}(G) = 0$. Denote by D the divisor $P_1 + \cdots + P_n$. There is a morphism $\gamma : \mathcal{L}(G) \to {\mathbf{F}_q}^n$ given by $\gamma(f) := (f(P_1), \ldots, f(P_n))$. The image $C(G, D)$ of this morphism is called an *elliptic code (over $\mathcal{E}$)*. For the rest of this paper we shall make the auxiliary assumptions $0 < \deg(G) < n$. This guarantees $C(G, D)$ to be nontrivial and γ to be injective. Hence, the dimension k of $C(G, D)$ equals $\dim(G) = \deg(G)$. The determination of the minimum distance d of $C(G, D)$ turns out to be more difficult. Using the Theorem of Riemann-Roch, one gets the inequality $d \geq n - k$. In view of the Singleton-inequality $d \leq n - k + 1$ [5, Chapter 2], d can only take the values $n - k$ or $n - k + 1$. In the latter case, $C(G, D)$ is MDS. Utilizing the group structure of $\mathcal{E}(\mathbf{F}_q)$ one can derive a criterion for $C(G, D)$ to be MDS, hence a criterion for determining d:

Theorem 1 *$C(G, D)$ is MDS if and only if for every $\deg(G)$-subset J of $\{1, \ldots, n\}$ the divisor $G - \sum_J P_j$ is not principal.*

Proof [7]. ∎

In this paper we shall primarily be concerned with two types of elliptic codes: with the above assumptions, the first class of codes (called codes of type I) consists of codes for which $\{P_1, \ldots, P_n\} = \mathcal{E}(\mathbf{F}_q)$. The second class of codes (called codes of type II) consists of codes for which $D = P_1 + \cdots + P_{n-1}$

and $G = \alpha P_n$ for some integer α satisfying $0 < \alpha < n$. The minimum distance of these codes has been determined in [7].

3 Weight Distribution of Elliptic Codes

In this section we shall derive explicit formulae for the weight distribution of elliptic codes. Actually, the method involved can be used to determine the weight distribution of any linear code with given dual distance.

Let C be an elliptic code over $\mathbf{F}_q$ as constructed in the last section. Let n denote the block-length, k the dimension and d the minimum distance of C. Then, as was pointed out in the last section, we have $d \in \{n-k, n-k+1\}$. If $d = n-k+1$ then C is MDS. This case is not of interest to us, since the weight distribution of MDS-codes is well-known [5, Chapter 11]. So assume that $d = n-k$. Let $C^\perp$ denote the dual of C. By [9, Theorem 2.5] $C^\perp$ is an $(n, n-k, d')$-code with $d' \in \{k, k+1\}$. (Actually, the mentioned theorem implies that $C^\perp$ is an elliptic code, so the assertion on d' follows.) Since C is assumed to be non-MDS, so is $C^\perp$ [5, Chapter 11]; thus $d' = k$.

Lemma 1 *The dual distance of the $(n, k, n-k)$-elliptic-code C equals k.* ∎

Let A_w and B_w denote the number of codewords of weight w of C resp. $C^\perp$. The MacWilliams identities [5, Chapter 11] relate these numbers:

$$(1) \qquad \sum_{i=0}^{n-j} \binom{n-i}{j} A_i = q^{k-j} \sum_{i=0}^{j} \binom{n-i}{j-i} B_i, \quad j = 0, 1, \ldots, n.$$

But since $A_0 = B_0 = 1$ and $A_1 = \ldots = A_{n-k-1} = B_1 = \ldots = B_{k-1} = 0$, we get

$$(2) \qquad \sum_{i=n-k}^{n-j} \binom{n-i}{j} A_i = \binom{n}{j} (q^{k-j} - 1), \quad j = 0, 1, \ldots, k-1.$$

If one knows the number A_{n-k} of minimum weight codewords of C, then it is possible to compute recursively the A_w for all w and hence the weight enumerator $A(z) := \sum_{w=0}^{n} A_w z^w$ of C. This is done by substituting for j the values $k-1, k-2, \ldots, 0$. Using induction we get the following formula for the numbers A_w:

Theorem 2 *Let C be an $(n, k, n-k)$ code with dual distance k. Further, let A_w denote the number of codewords in C of weight w. Then we have*

$$A_{n-k+r} = \binom{n}{k-r} \sum_{j=0}^{r-1} (-1)^j \binom{n-k+r}{j} (q^{r-j} - 1) + (-1)^r \binom{k}{k-r} A_{n-k}$$

for all $r \in \{0, \ldots, k\}$.

Proof Induction on r. ∎

Remark. Comparing this result with the weight distribution of MDS-codes [5, Chapter 11, Theorem 6], one realizes that the assertion of the theorem remains valid, if one only requests C to be an (n, k, d)-code with $d \geq k$.

Corollary 1 *Let C be an $(n, k, n-k)$-elliptic-code. Then we have:*

1. *The weight distribution of C is completely determined by the number A_{n-k} of minimum weight codewords of C.*

2. *If B_w denotes the number of codewords of weight w in $C^\perp$, we have $B_k = A_{n-k}$, i.e. $C^\perp$ and C have the same number of minimum weight codewords.*

3. *If $n = 2k$, then C and $C^\perp$ have the same weight distribution, i.e. are formally selfdual.*

Proof 1. This is an immediate consequence of the formula in Theorem 2.

2. Substitute in the MacWilliams-identities (1) for j the value k. Then, noting that $A_1 = \ldots = A_{n-k-1} = B_1 = \ldots = B_{k-1} = 0$ and $A_0 = B_0 = 1$, we get the assertion.

3. Follows from 1 and 2. ∎

For the computation of the weight distribution of C it is thus sufficient to compute the number of minimum weight codewords. For this, we shall utilize the following

Lemma 2 *Let $\mathcal{E}$ be an elliptic curve over $\mathbf{F}_q$, $P_1, \ldots, P_n \in \mathcal{E}(\mathbf{F}_q)$ and G a divisor of $\mathcal{E}$ subject to the conditions $\mathrm{ord}_{P_1}(G) = \ldots = \mathrm{ord}_{P_n}(G) = 0$ and $0 < \deg(G) < n$. Further, assume that $C(G, D)$ is not MDS. Let $Q \in \mathcal{E}(\mathbf{F}_q)$ correspond to the neutral element of $\mathcal{E}(\mathbf{F}_q)$.*

Then the number of minimum weight codewords of $C(G,D)$ *equals* $(q-1)$ *times the number of* $\deg(G)$*-subsets* J *of* $\{1,\ldots,n\}$ *such that*

$$\bigoplus_J P_j = \mathrm{Tr}(G). \tag{3}$$

Proof Let $C(G,D)$ be an (n,k,d)-code. Then $n = \deg(D)$ and $k = \deg(G)$. If J is a subset of $\{1,\ldots,n\}$ satisfying (3), then we have by definition of $\oplus$: $\sum_J P_j - |J|\mathrm{Tr}(G) \in \mathbf{H}(\mathcal{E})$, i.e.

$$\sum_J P_j - G \in \mathbf{H}(\mathcal{E}). \tag{4}$$

Conversely, each subset J satisfying (4) also satisfies (3). Since each set J satisfying (4) gives rise to a divisor $D' := \sum_J P_j \leq D$ for which $D' - G$ is principal, and conversely, each divisor $D' \leq D$ such that $D' - G \in \mathbf{H}(\mathcal{E})$ determines a subset J satisfying (4) in an obvious manner, the number of subsets J satisfying (3) equals the number of elements in the set

$$\{D' \mid D' \leq D, D' - G \in \mathbf{H}(\mathcal{E})\}. \tag{5}$$

Now consider the number A_{n-k} of minimum weight codewords of $C(G,D)$. By definition of $C(G,D)$ this is equal to the number of functions $f \in \mathcal{L}(G)$ such that there exists a k-subset J of $\{1,\ldots,n\}$ for which we have $f(P_j) = 0$ for all $j \in J$. It follows that $(f)_0 \geq \sum_J P_j$. But since $f \in \mathcal{L}(G)$ we have $(f) \geq -G$, thus $(f) \geq \sum_J P_j - G$. Since $\deg(G) = k = |J|$, we have $(f) = \sum_J P_j - G$. So it is not difficult to see that

$$A_{n-k} = |\{f \mid f \in \mathcal{L}(G), \exists\, J \subseteq \{1,\ldots,n\}\colon (f) = \sum_J P_j - G\}|. \tag{6}$$

Denote the above set by $M(D,G)$. We define a mapping ψ

$$\begin{array}{rccl} \psi: & M(D,G) & \to & \{D' \mid D' \leq D, D' - G \in \mathbf{H}(\mathcal{E})\} \\ & f & \mapsto & (f) + G. \end{array}$$

Clearly, ψ is surjective and well-defined. Moreover $\psi(f_1) = \psi(f_2)$ implies $(f_1) = (f_2)$, i.e. $f_2/f_1 \in \mathbf{F}_q{}^{\times}$. Thus the assertion follows. ∎

Using Lemma 2 one can reformulate the problem of determining the number of minimum weight codewords of a non-MDS elliptic code $C(G,D)$ as follows:

Given a finite additive abelian group $\mathfrak{G}$, a subset M of $\mathfrak{G}$, a positive integer k and an element g of $\mathfrak{G}$, determine the number of k-subsets $\{g_1, \ldots, g_k\}$ of M subject to the condition $\sum_{i=1}^{k} g_i = g$.

In our applications $\mathfrak{G}$ is the group $(\mathcal{E}(\mathbf{F}_q), \oplus)$, M is the subset $\{P_1, \ldots, P_n\}$, k equals $\deg(G)$ and g is $\mathrm{Tr}(G)$. It should be pointed out that we are merely interested in a constrained version of the above problem since the group $(\mathcal{E}(\mathbf{F}_q), \oplus)$ is of the type $C_n \times C_m$, where C_n and C_m denote the cyclic groups of orders m and n [6].

In the next section we are going to solve the above problem in special cases for the elliptic codes of types I and II.

4 Weight Distribution of Elliptic Codes of Type I

Let $\mathcal{E}$ be an elliptic curve over $\mathbf{F}_q$. Further let $\mathcal{E}(\mathbf{F}_q) = \{P_1, \ldots, P_n\}$ and choose a divisor $G \in \mathbf{D}(\mathcal{E})$ subject to the conditions $0 < \deg(G) < n$ and $\mathrm{ord}_{P_1}(G) = \ldots = \mathrm{ord}_{P_n}(G) = 0$. Put $D := P_1 + \ldots + P_n$. We call the code $C(G, D)$ an elliptic code of type I (cf. Section 2). In the following we are going to determine the weight distribution of $C(G, D)$.

Let $\mathfrak{G}$ be an additive finite abelian group of order n and j be an integer satisfying $0 < j < n$. For $g \in \mathfrak{G}$ define

$$S_j(g; \mathfrak{G}) := \Big\{ T \subseteq \mathfrak{G} \;\Big|\; |T| = j, \sum_{t \in T} t = g \Big\}$$

and $s_j(g; \mathfrak{G}) := |S_j(g; \mathfrak{G})|$. The following is a reformulation of Lemma 2:

Lemma 3 *Assume that the code $C(G, D)$ is not MDS. Then the number of minimum weight codewords of $C(G, D)$ equals $(q-1) \cdot s_{\deg(G)}(\mathrm{Tr}(G); \mathcal{E}(\mathbf{F}_q))$, where $\mathrm{Tr}(G) \in \mathcal{E}(\mathbf{F}_q)$ is the trace of G.* ∎

The main result about the numbers $s_j(g; \mathfrak{G})$ is given in the following

Lemma 4 *Let $\mathfrak{G}$ be an additive finite abelian group. Further, let j be an integer prime to $|\mathfrak{G}|$. Then for all $g \in \mathfrak{G}$ we have*

$$s_j(g; \mathfrak{G}) = \frac{1}{|\mathfrak{G}|} \binom{|\mathfrak{G}|}{j}.$$

Proof Since the sets $S_j(g_1;\mathfrak{G})$ and $S_j(g_2;\mathfrak{G})$ are disjoint if g_1 and g_2 are different elements of $\mathfrak{G}$ we get the relation

$$\sum_{g\in\mathfrak{G}} s_j(g;\mathfrak{G}) = \binom{|\mathfrak{G}|}{j}. \tag{7}$$

Now let $g,\mathfrak{g}$ be elements of $\mathfrak{G}$. Then there is a bijection between $S_j(g;\mathfrak{G})$ and $S_j(g+j\mathfrak{g};\mathfrak{G})$ which maps a set $T \in S_j(g;\mathfrak{G})$ onto the set $\{t+\mathfrak{g} \mid t \in T\} \in S_j(g+j\mathfrak{g};\mathfrak{G})$. So we have

$$\forall g,\mathfrak{g} \in \mathfrak{G}: \quad s_j(g;\mathfrak{G}) = s_j(g+j\mathfrak{g};\mathfrak{G}).$$

Since j and $|\mathfrak{G}|$ are assumed to be coprime, multiplication by j is an automorphism of $\mathfrak{G}$, i.e. $j\mathfrak{G} = \mathfrak{G}$. Hence together with (7) we get the assertion. ∎

Combining the results of this section with Theorem 2 we get the following explicit weight distribution for elliptic codes of type I:

Theorem 3 *Let $C(G,D)$ be defined as above. Suppose that $C(G,D)$ is not MDS and $j := \deg(G)$ and $n := |\mathcal{E}(\mathbf{F}_q)|$ are coprime. Denote by A_w the number of codewords of weight w in $C(G,D)$. Then we have*

$$A_{n-j+r} = \binom{n}{j-r}\sum_{i=0}^{r-1}(-1)^i\binom{n-j+r}{i}(q^{r-i}-1)+\frac{(-1)^r(q-1)}{n}\binom{j}{j-r}\binom{n}{j}$$

for all $r \in \{0,\ldots,j\}$. ∎

The question, when the code $C(G,D)$ is MDS has been answered in [7]. The reformulated result reads as follows.

Theorem 4 *The code $C(G,D)$ is MDS if and only if $\mathcal{E}(\mathbf{F}_q) \simeq C_2 \times C_2$ and $\deg(G) = 2$ and the trace of G is the neutral element in $\mathcal{E}(\mathbf{F}_q)$.*

Proof [7]. ∎

5 Weight Distribution of Elliptic Codes of Type II

Let $\mathcal{E}$ be an elliptic curve over $\mathbf{F}_q$. Denote by $P, P_1, \ldots, P_n$ *all* the different $\mathbf{F}_q$-rational points of $\mathcal{E}$. Without loss of generality, let P correspond to the neutral element of $(\mathcal{E}(\mathbf{F}_q), \oplus)$. Further let j, $0 < j < n$ be an integer and denote by G the divisor jP. Put $D := P_1 + \cdots + P_n$. Then $C(G, D)$ is an elliptic code of Type II.

In accordance to the present situation where $|\mathcal{E}(\mathbf{F}_q)| = n+1$, let $\mathfrak{G}$ denote an additive finite abelian group of order $n+1$ and j be an integer satisfying $0 < j < n$. Further let $\mathfrak{G}^*$ denote the set of nonzero elements of $\mathfrak{G}$. For $g \in \mathfrak{G}$ define

$$\overline{S}_j(g; \mathfrak{G}) := \left\{ T \subseteq \mathfrak{G}^* \;\middle|\; |T| = j, \sum_{t \in T} t = g \right\}$$

and $\overline{s}_j(g; \mathfrak{G}) := |\overline{S}_j(g; \mathfrak{G})|$. Again, the following is a reformulation of Lemma 2:

Lemma 5 *Assume that the code $C(G, D)$ is not MDS. Then the number of minimum weight codewords of $C(G, D)$ equals $(q-1) \cdot \overline{s}_{\deg(G)}(P; \mathcal{E}(\mathbf{F}_q))$.* ∎

The next lemma gives the value of $\overline{s}_j(g; \mathfrak{G})$ for different j:

Lemma 6 *Let $\mathfrak{G}$ be an additive finite abelian group. Further let j be an integer such that $|\mathfrak{G}|$ is coprime to $j!$. Then one has*

$$\overline{s}_j(g; \mathfrak{G}) = \begin{cases} \dfrac{1}{|\mathfrak{G}|}\left(\binom{|\mathfrak{G}|-1}{j} + (-1)^j(|\mathfrak{G}|-1)\right) & \text{for } g = 0, \\ \dfrac{1}{|\mathfrak{G}|}\left(\binom{|\mathfrak{G}|-1}{j} + (-1)^j(|\mathfrak{G}|-1)\right) + (-1)^{j-1} & \text{for } g \neq 0. \end{cases}$$

Proof First observe that

$$s_j(g; \mathfrak{G}) = \overline{s}_j(g; \mathfrak{G}) + \overline{s}_{j-1}(g; \mathfrak{G}). \tag{8}$$

By Lemma 4 $s_\mu(g; \mathfrak{G}) = \frac{1}{|\mathfrak{G}|}\binom{|\mathfrak{G}|}{j}$ for all μ, $1 \leq \mu \leq j$, since $|\mathfrak{G}|$ and $j!$ are assumed to be coprime. Induction on j yields the assertion. Note that the reason why one gets different formulae for $g = 0$ and $g \neq 0$ is the fact that $\overline{s}_1(0; \mathfrak{G}) = 0$ while $\overline{s}_1(g; \mathfrak{G}) = 1$ if $g \neq 0$. ∎

Lemma 5 and Lemma 6 in combination with Theorem 2 yield the following weight distribution for the code $C(G, D)$:

Theorem 5 *Let $C(G,D)$ be defined as above. Suppose that $C(G,D)$ is not MDS. Let $j := \deg(G)$ and $n := |\mathcal{E}(\mathbf{F}_q)|$ and assume that $j!$ and n are coprime. Denote by A_w the number of codewords of weight w in $C(G,D)$. Then we have*

$$\begin{aligned} A_{n-j+r} &= \binom{n}{j-r}\sum_{i=0}^{r-1}(-1)^i\binom{n-j+r}{i}(q^{r-i}-1) \\ &\quad +\frac{(-1)^r}{n}\left(\binom{n-1}{j}+(-1)^j(n-1)\right)(q-1) \end{aligned}$$

for all $r \in \{0,\ldots,j\}$. ∎

Remark. The assumptions of the above theorem are satisfied if for example the class group of $\mathcal{E}$ is of prime order.

In [7] necessary and sufficient conditions for $C(G,D)$ to be MDS are given. We are not going to discuss the conditions in detail.

6 The General Problem

In this section we are going to outline briefly the general problem of determining the weight distribution of an elliptic code. The main result of this section is the assertion that if one knows the weight distribution of every elliptic code of Type I, then one can derive explicit formulae for the weight distribution of an arbitrary elliptic code.

In Section 3 we reduced the problem of determining the weight distribution of an elliptic code to a combinatorial problem in the theory of abelian groups. This Problem in mind, let us define for a subset M of the abelian group $\mathfrak{G}$ and an element $g \in \mathfrak{G}$

$$S_j(g;M) := \left\{T \subseteq M \;\middle|\; |T| = j, \sum_{t\in T} t = g\right\}$$

and $s_j(g;M) := |S_j(g;M)|$. Note that $\overline{s}_j(g;\mathfrak{G}) = s_j(g;\mathfrak{G}^*)$. It is easy to see that an elliptic code $C(G,D)$ is MDS if and only if $s_j(g;M) = 0$ where $j = \deg(G)$, $M = \{P_1,\ldots,P_n\}$ if $D = P_1 + \cdots + P_n$, $\mathfrak{G} = \mathcal{E}(\mathbf{F}_q)$ and $g = \mathrm{Tr}(G)$.

According to the results in Section 3 the number of minimum weight codewords of an arbitrary non-MDS elliptic code over $\mathbf{F}_q$ equals $(q-1)s_j(g;M)$ for appropriate j, M, g and $\mathfrak{G}$. In view of Theorem 2 the weight distribution of an elliptic code is known, once one knows the numbers $s_j(g;M)$. In order to relate the numbers $s_j(g;M)$ with the numbers $s_k(a;\mathfrak{G})$ we need the following

Lemma 7 *Let $\mathfrak{G}$ be a finite abelian group, $M \subseteq \mathfrak{G}$, j a natural number with $1 \leq j < |M|$, $g \in \mathfrak{G}$ and $b \in \mathfrak{G} \setminus M$. Then we have*

$$s_j(g; M \cup \{b\}) = s_j(g;M) + s_{j-1}(g-b;M). \tag{9}$$

Proof Denote by S the set

$$S := \Big\{ T \in S_j(g; M \cup \{b\}) \;\Big|\; b \in T \Big\}.$$

The mapping $\xi: S \to S_{j-1}(g-b;M)$ defined by $\xi(T) := T \setminus \{b\}$ is clearly a well defined bijection, since $b \notin M$ is assumed. Hence we have $|S| = s_{j-1}(g-b;M)$. On the other hand, the set $S_j(g;M \cup \{b\})$ is the disjoint union of S and $S_j(g;M)$. Hence the assertion follows. ∎

Having Lemma 7 at hand, it is possible to give a formula for $s_j(g;M)$ in terms of the numbers $s_\mu(a; M \cup \{b\})$.

Lemma 8 *With the assumptions of Lemma 7 we have*

$$\begin{aligned} s_j(g;M) \;=\; & \sum_{\mu=2}^{j} s_\mu(g-(j-\mu)b; M \cup \{b\})(-1)^{j-\mu} \\ & +(-1)^{j-1}\chi_M(g-(j-1)b) \end{aligned} \tag{10}$$

where χ_M is the characteristic function of M.

Proof The formula follows from (9) using induction on j. Note that $s_1(a;M) = \chi_M(a)$. ∎

Formula (10) is the basis of a recursion procedure which at the end yields a formula for $s_j(g;M)$ as an integral linear combination of numbers $s_\mu(a;\mathfrak{G})$ for $\mu \leq j$ and appropriate $a \in \mathfrak{G}$ and numbers $\chi_L(a')$ for appropriate subsets $M \subseteq L \subseteq \mathfrak{G}$ and $a' \in \mathfrak{G}$. Since the formula is very lengthy, we are not going to discuss it in detail.

Theorem 6 *With the above assumptions, the number $s_j(g;M)$ can be expressed as an integral linear combination of the numbers $s_\mu(a_\mu;\mathfrak{G})$, $2 \leq \mu \leq j$ and $\chi_L(b)$ for appropriate $a_\mu \in \mathfrak{G}$, $b \in \mathfrak{G}$ and $M \subseteq L \subseteq \mathfrak{G}$.*

7 Examples

This section contains several examples how to compute the weight distribution of elliptic codes with the aid of the theorems presented in this paper.

Example 1 Consider the elliptic curve $\mathcal{E}$ over $\mathbf{F}_5$ given by the Weierstraß-equation

$$y^2 = x^3 + 2x + 1.$$

$\mathcal{E}$ has seven $\mathbf{F}_5$-rational points. Denoting the point with the projective coordinates $(0:1:0)$ by ∞, the other six points have the affine coordinates

$$\begin{array}{ll} P_1 = (1,2), & P_2 = (1,-2), \\ P_3 = (-2,2), & P_4 = (-2,-2), \\ P_5 = (0,1), & P_6 = (0,-1). \end{array}$$

Let the divisor G be given by $G = (f)_0$ where $f = (x+1)(x-2)/(x^2+x+1)$. It is easy to check that $\mathrm{ord}_\infty(G) = \mathrm{ord}_{P_1}(G) = \cdots = \mathrm{ord}_{P_6}(G) = 0$. Further $\deg(G) = 4$. Let $D := \infty + P_1 + \cdots + P_6$. According to [7, Theorem 3] the code $C(G,D)$ is not MDS. It is a $(7,4,3)$ code. By Lemma 3 and Lemma 4 the number of codewords of weight 3 in $C(G,D)$ equals

$$(5-1)\frac{1}{7}\binom{7}{4} = 20.$$

Using Theorem 3 we get the following weight enumerator for $C(G,D)$

$$140z^7 + 200z^6 + 204z^5 + 60z^4 + 20z^3 + 1.$$

Example 2 With the notations of Example 1 let now $G := 4\infty$ and $D := P_1 + \cdots + P_6$. By [7, Theorem 1] the code $C(G,D)$ is not MDS. Thus $C(G,D)$ is a $(6,4,2)$-code. Taking ∞ as the neutral element of $\mathcal{E}(\mathbf{F}_5)$, we have $\mathrm{Tr}(G) = \infty$. So, by Lemma 5 the number of minimum weight codewords of $C(G,D)$ equals $4\overline{s}_4(0;C_7)$ where C_7 denotes the cyclic group of order 7. By Lemma 6 the number of minimum weight codewords of $C(G,D)$ thus equals

$$4 \cdot \frac{1}{7}\left(\binom{6}{3} + 6\right) = 12.$$

Theorem 5 now yields the following weight enumerator for $C(G,D)$:

$$172z^6 + 216z^5 + 192z^4 + 32z^3 + 12z^2 + 1.$$

Example 3 The following example will demonstrate how to compute the numbers $s_j(g;M)$ from the numbers $s_\mu(a;\mathfrak{G})$ for $\mathfrak{G} = C_7 = \mathbf{Z}/7\mathbf{Z}$, $j = 4$, $M = \{2,3,4,5,6\}$ and $g = 0$. The result will be then used to compute the weight distribution of the code $C(G,D)$ where $G = 4\infty$ and $D = P_2 + P_3 + \cdots + P_6$, where we again have used the notations of the foregoing two examples.

By Lemma 8 we have

$$s_4(0;M) = \sum_{\mu=2}^{4} \overline{s}_\mu(0-(4-\mu);C_7)(-1)^{4-\mu} - \chi_M(-3).$$

Using Lemma 6 we compute $s_4(0;M) = 1$. The number of minimum weight codewords of $C(G,D)$ thus equals 4. $C(G,D)$ is a $(5,4,1)$-code. Its weight enumerator, which can be computed using Theorem 2, equals

$$208z^5 + 244z^4 + 144z^3 + 24z^2 + 4z + 1.$$

Example 4 We begin this example by computing the values $s_j(a;C_9)$ for all values $1 \leq j \leq 8$ and $a \in C_9$ where C_9 is the cyclic group of order 9.

First of all, we remark that $s_j(a;C_9) = s_{9-j}(-a;C_9)$. This can be seen easily by looking at the complements of the sets in $S_j(a;C_9)$ and taking into account that the sum of the elements in C_9 equals 0. Secondly, by Lemma 4 we have $s_j(a;C_9) = \frac{1}{9}\binom{9}{j}$ if j is coprime to 9. Hence we only have to compute $s_3(a;C_9)$ for all $a \in C_9$. By considering the sets $\{t+1 \mid t \in T\}$ for $T \in S_3(a;C_9)$ we can show similar to the proof of Lemma 4 that $s_j(a;C_9)$ is constant (as a function of a) on the residue classes of C_9/C_3. Hence, identifying C_9 with $\mathbf{Z}/9\mathbf{Z}$ and taking for the elements the residue classes of the numbers from 0 to 8, we see that it only remains to compute $s_3(0;C_9)$, $s_3(1;C_9)$ and $s_3(2;C_9)$. This computation can be done by hand or by a small pocket calculator. The result is

$$s_3(0;C_9) = 10, \quad s_3(1;C_9) = 9, \quad s_3(2;C_9) = 9.$$

The following is a list of all the values of $s_j(a; C_9)$ for $1 \leq j \leq 4$:

$$\forall\, a \in C_9\colon s_1(a; C_9) = 1$$

$$\forall\, a \in C_9\colon s_2(a; C_9) = 4$$

$$s_3(0; C_9) = s_3(3; C_9) = s_3(6; C_9) = 10$$

$$s_3(1; C_9) = s_3(4; C_9) = s_3(7; C_9) = 9$$

$$s_3(2; C_9) = s_3(5; C_9) = s_3(8; C_9) = 9$$

$$\forall\, a \in C_9\colon s_4(a; C_9) = 14.$$

Now let $\mathcal{E}$ be the elliptic curve over $\mathbf{F}_5$ defined by

$$y^2 = x^3 + x + 1.$$

This curve has 9 rational points over $\mathbf{F}_5$. Using the explicit equations for adding the points on the curve [10] one computes the group $\mathcal{E}(\mathbf{F}_5)$ to be of type C_9. Let ∞ denote the point with the projective coordinates $(0 : 1 : 0)$. Further let the other 8 points be denoted by $P_1, \ldots, P_8$ and let $G := 4\infty$ and $D := P_1 + \cdots + P_8$. Then the trace of G is ∞ which is the neutral element of $\mathcal{E}(\mathbf{F}_5)$. So by Lemma 5 the number of minimum weight codewords of $C(G, D)$ equals $4\overline{s}_4(0; C_9)$. In order to compute $\overline{s}_4(0; C_9)$, we use the formula (10) which yields

$$\overline{s}_4(0; C_9) = \sum_{\mu=2}^{4} s_\mu(0; C_9)(-1)^{4-\mu}.$$

We thus get

$$\overline{s}_4(0; C_9) = 4 - 10 + 14 = 8.$$

The number of minimum weight codewords of the $(8, 4, 4)$-code $C(G, D)$ is 32. Hence we get by Theorem 2 the following weight enumerator for $C(G, D)$

$$112z^8 + 192z^7 + 192z^6 + 96z^5 + 32z^4 + 1.$$

Example 5 Let the elliptic curve $\mathcal{E}$ and the point ∞ be defined as in the previous example. $P = (2, 1)$ is a point of order 3 of $\mathcal{E}(\mathbf{F}_5)$. We can identify it with the element 3 in C_9. Let $G = 3P$ and $D = \sum_{Q \in \mathcal{E}(\mathbf{F}_5), Q \neq P} Q$. The code

$C(G,D)$ is an $(8,3,5)$-code. The number of minimum weight codewords of $C(G,D)$ equals $4s_3(3;M)$ where $M = C_9 \setminus \{3\}$. By (10) we have

$$s_3(3;M) = \sum_{\mu=2}^{3} s_\mu(3-(3-\mu)3;C_9)(-1)^{3-\mu} + 1.$$

We thus have

$$s_3(3;M) = -4 + 10 + 1 = 7.$$

The number of minimum weight codewords of $C(G,D)$ is equal to 28. We thus get the following weight enumerator for $C(G,D)$

$$16z^8 + 52z^7 + 28z^6 + 28z^5 + 1.$$

References

[1] E.R. Berlekamp: *Algebraic Coding Theory.* McGraw-Hill, New York, 1968.

[2] E.R. Berlekamp, R. J. McEliece, H. C. A. van Tilborg: On the Inherent Intractibility of Certain Coding Problems. *IEEE Transactions on Information Theory*, **24**, 384–386, 1978.

[3] V. D. Goppa: Codes on Algebraic Curves. *Soviet Math. Doklady,* **24**, 170–172 (1981).

[4] D. Husemöller: *Elliptic Curves.* Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1987.

[5] F. J. MacWilliams, N. J. A. Sloane: *The Theory of Error Correcting Codes.* North Holland, Amsterdam, New York, Oxford, 1977.

[6] H. G. Rück: A Note on Elliptic Curves over Finite Fields. *Mathematics of Computation*, **49**, 301–304, (1987).

[7] M. A. Shokrollahi: Minimum Distance of Elliptic Codes. *Research Report, Universität Bonn* (1989).

[8] J. H. Silverman: *The Arithmetic of Elliptic Curves.* Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1986.

[9] H. Stichtenoth: Self Dual Goppa Codes. *Journal of Pure and Applied Algebra,* **55**, 199–211 (1988).

[10] J. Tate: The Arithmetic of Elliptic Curves. *Inventiones Math.,* **123**, 179–206 (1974).

[11] F. Vatan: The Complexity of Computing the Packing Radius and the Minimum Weight of a Code. *Preprint*, 1989.