

Using Quadratic Forms to find Curves with Many Points

Pippa Charters

1 INTRODUCTION

In both algebraic geometry and coding theory, there is a great deal of interest in finding curves with many rational points. In particular, the correspondence between trace codes and Artin-Schreier curves gives a relation between the weights of codewords and the number of rational points on such curves, low weight codewords yielding curves with a large number of rational points. Further, subcodes of these codes correspond to fiber products of Artin-Schreier curves. In this report, I will be following [3].

2 CODES

Some important coding terminology is as follows. To start with, we will be looking at Reed-Muller codes (a subset of trace codes) which are *block codes*. That is, they take a string of letters in some alphabet (in our case a finite field) of set length, and output a codeword in another (possibly different) alphabet, with said codeword also having fixed length (greater than or equal to the original word).

Definition. For a vector $v = (v_1, \dots, v_n) \in \mathcal{C}$ a codeword, the **Hamming weight** of v is defined by

$$w(v) := \#\{i \mid 1 \leq i \leq n, v_i \neq 0\}.$$

That is, the Hamming weight is the number of nonzero places in our codeword v .

Definition. Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a k -dimensional code (also called a rank k code). Then n is called the **length** of \mathcal{C} and

$$d := \min\{w(v) \mid v \in \mathcal{C} \setminus 0\}$$

is called the **minimum distance** of \mathcal{C} .

We will be looking for codewords of small minimum distance (also called small minimum weight) as these will correspond to Artin-Schreier curves with many rational points.

The specific correspondence we will be working with is that between binary second-order Reed-Muller codes [which in turn correspond to quadratic forms] and curves with many points. It turns out that binary

second-order RM codes are *trace codes*, [that is, codewords are gotten as the trace of some function, where $Tr : \mathbb{F}_{q^m} \rightarrow \mathbb{F}_q$ is given by $Tr(x) = \sum_{i=0}^{m-1} x^{q^i}$] with a special relationship to quadratic forms.

2.1 SECOND ORDER REED-MULLER CODES

Definition. The 2^{nd} order Reed-Muller code $R(2, m)$ of length $n = 2^m$ [$0 \leq r \leq m$], is the set of all vectors f , where $f(v_1, \dots, v_m)$ is a Boolean [taking values in $\{0, 1\}$] function which is polynomial of degree at most 2.

Example 1. The second order RM code of length 8 consists of all linear combinations of the vectors corresponding to the products

$$1, v_1, v_2, v_3, v_1v_2, v_1v_3, v_2v_3$$

I.e., $f = v_1 + v_1v_2 + v_2v_3$ gives $f(000) = 0, f(001) = 0, f(010) = 0, f(011) = 1, f(100) = 1, f(101) = 1, f(110) = 0, f(111) = 1$ which corresponds to the codeword:

$$(0\ 0\ 0\ 1\ 1\ 1\ 0\ 1)$$

We can write a typical codeword in the second order RM codes as

$$\begin{aligned} S(\vec{v}) &= \sum_{1 \leq i \leq j \leq m} q_{ij} v_i v_j + \sum_{1 \leq i \leq m} l_i v_i + \epsilon \\ &= Q(v) + L(v) + \epsilon \end{aligned}$$

where $Q(v)$ is a quadratic form and $L(v)$ is a linear form over \mathbb{F}_2 .

Definition. For $2 \leq m - 1$, the *punctured RM code* $\mathcal{R}(2, m)^*$ is obtained by puncturing (or deleting) the coordinate corresponding to $v_1 = v_2 = \dots = v_m = 0$ from all the codewords of $\mathcal{R}(2, m)$.

3 CONSTRUCTION OF SUBCODE

In [3], van der Geer and van der Vlugt construct a subcode of the punctured binary 2^{nd} order RM code, and recall a previously proven relationship between the number of rational points on fiber products of curves associated to trace codes and the Hamming weights of such codes in order to construct curves with many rational points. This is done as follows.

Definition. Let $q = 2^m$. Then for $0 < h \leq \lfloor m/2 \rfloor$, we define the \mathbb{F}_q - vector space of 2-linearized polynomials by

$$R_h = \{R(x) = \sum_{i=0}^h a_i x^{2^i} \mid a_i \in \mathbb{F}_q\}$$

From this vector space, we can derive a binary subcode \mathcal{C}_h of the punctured second-order RM code given by $\mathcal{C}_h = \{c_R = (Tr(xR(x)))_{x \in \mathbb{F}_q^*} \mid R \in R_h\}$. To each codeword $c_R \in \mathcal{C}_h$, we can define a non-singular projective curve C_R associated to it via the affine equation $y^2 + y = xR(x)$. It turns out that this is an Artin-Schreier cover of \mathbb{P}_1 . That is, it is an abelian covering of degree p for some prime p or equivalently, the field of rational functions on this curve is a $\mathbb{Z}/p\mathbb{Z}$ extension of the field of rational functions on \mathbb{P}_1 , not a constant field extension.

4 THE RELATION BETWEEN CODES AND QUADRATIC FORMS

In order to make what we are doing clear, I will try and split the process up into a series of steps.

1. Recall from the previous section that to each non-zero word $c_R \in \mathcal{C}_h$ we can associate a non-singular projective curve. It turns out that we can also associate to \mathcal{C}_h the quadratic form

$$Q(x) = Q_R(x) = Tr(xR(x))$$

a quadratic form over \mathbb{F}_2 in m variables. This in turn has an associated symmetric (and also symplectic) bilinear form

$$B(x, y) = Tr(xR(y) + yR(x))$$

Define

$$W = \{x \in \mathbb{F}_q \mid B(x, y) = 0 \forall y \in \mathbb{F}_q\}, \quad W_0 = \{x \in W \mid Q(x) = 0\}$$

Let $w = \dim W$ as an \mathbb{F}_2 -vector space, and note that since B is symplectic, $m - w \equiv 0 \pmod{2}$, and moreover $\dim W_0 = w$ or $\dim W_0 = w - 1$.

2. We can now classify these quadratic forms in terms of the number of zeros of $Q(x)$ in the following manner.

- $W_0 = W$. In this case, $Q(x)$ has rank $m - w$, and either

$$Q(x) \sim X_1 X_2 + \cdots + X_{m-w-1} X_{m-w}$$

and $Q(x)$ has $\frac{q+\sqrt{q2^w}}{2}$ zeros or

$$Q(x) \sim X_1X_2 + \cdots + X_{m-w-1}X_{m-w} + X_{m-w-1}^2 + X_{m-w}^2$$

and $Q(x)$ has $\frac{q-\sqrt{q2^w}}{2}$ zeros.

- $W_0 \neq W$. In this case, $Q(x)$ has rank $m - w + 1$, and

$$Q(x) \sim X_1X_2 + \cdots + X_{m-w-1}X_{m-w} + X_{m-w+1}^2$$

with $Q(x)$ having $q/2$ zeros.

We can then use these classifications to come up with a simple expression for both the weight of our codeword and the number of points on the corresponding curve, namely

$$w(c_R) = q - \#\text{zeros of } Q(x)$$

and

$$\#C_R(\mathbb{F}_q) = 2(\#\text{zeros of } Q(x)) + 1$$

We can use these definitions combined with facts about quadratic forms to create low-weight subcodes of \mathcal{C}_h for odd m .

4.1 THE CONSTRUCTION

In order to construct low weight subcodes, first we define the quadratic form

$$Q(a, b) = Q(a_1, \dots, a_{(m-w)/2}, b_1, \dots, b_{(m-w)/2}) = \sum_{i=1}^{(m-w)/2} \text{Tr}(a_i x) \text{Tr}(b_i x)$$

equivalent to $X_1X_2 + \cdots + X_{m-w-1}X_{m-w}$ and hence having $\frac{q+\sqrt{q2^w}}{2}$ zeros in \mathbb{F}_q . It turns out that the word induced by evaluating $Q(a, b)$ is an element of \mathcal{C}_h if the elements $a_i, b_i \in \mathbb{F}_q$, $1 \leq i \leq (m-w)/2$ satisfy the system of equations

$$\sum_{i=1}^{(m-w)/2} (a_i^{2^j} b_i + a_i b_i^{2^j}) = 0 \tag{1}$$

for j between $h + 1$ and $(m - 1)/2$, inclusive.

We now fix $a = (a_1, \dots, a_{(m-w)/2})$. It turns out that not only do the words $Q(a, b)$ induced by the solutions $b = (b_1, \dots, b_{(m-w)/2})$ to equation 1 form a subcode of \mathcal{C}_h , but moreover for such fixed a there are at least q such solutions b .

Now let $M = (m - w)/2$. Recall that to obtain curves with many points, we need to find words of minimum weight in \mathcal{C}_h . In order to do this, we will be looking at solutions b of $Q(a, b)$ such that the rank over \mathbb{F}_2 of $\{a, b\}$ is $2M$.

Proposition 4.1. *Let $b = (b_1, \dots, b_M)$ be a solution to equation 1. Then if*

$$rk_{\mathbb{F}_2}(\{a_1, \dots, a_M, b_1\}) = M + 1,$$

$$\text{then } rk_{\mathbb{F}_2}(\{a_1, \dots, a_M, b_1, \dots, b_M\}) = 2M$$

.

That is, if one of the b_i is linearly independent, then they all are. Thus codewords of minimum weight do exist. What we now want to know is whether or not there exists a minimum weight subcode of \mathcal{C}_h , that is a subcode made up of words which all have minimum weight.

Theorem 4.2. *Let S be the \mathbb{F}_2 -vector space of solutions to $\sum_{i=1}^M (a_i^{2^j} b_i + a_i b_i^{2^j}) = 0$ for some fixed a . Let V be the image in \mathbb{F}_q of the projection of S onto the first coordinate b_1 . That is, V is the set of first coordinates to the solutions of the equation. Then if $r = \dim_{\mathbb{F}_2}(V) - M > 0$, then there exists a minimum weight subcode of \mathcal{C}_h of dimension r .*

So it is possible that such a code does exist, which is part of what we are looking for. But how can we figure out $\dim_{\mathbb{F}_2}(V)$? What we need is the following theorem:

Theorem 4.3. *Let $b = (b_1, \dots, b_M) \in S$ [so b is a solution to the above equation]. Then $b' = (b_1, b'_2, \dots, b'_M) \in S$ iff there exists a symmetric $(M - 1) \times (M - 1)$ matrix A over \mathbb{F}_2 such that*

$$(b'_2, \dots, b'_M) = (b_2, \dots, b_M) + (a_2, \dots, a_M)A$$

where recall that $a = (a_1, \dots, a_M)$ is fixed.

It follows that since we can choose $\frac{M(M-1)}{2}$ entries in A independently [the entries on the upper triangle], we have

Corollary 4.4. *The dimension of V over \mathbb{F}_2 satisfies*

$$\dim_{\mathbb{F}_2}(V) = \dim_{\mathbb{F}_2}(S) - \binom{M}{2} \geq m - \binom{M}{2}.$$

Thus we have a lower bound on $\dim_{\mathbb{F}_2}(V)$, which can be applied to Theorem 4.2 to see that there exists a minimum weight subcode of \mathcal{C}_h of dimension r whenever

$$\begin{aligned} r &= \dim_{\mathbb{F}_2}(V) - M \\ &\geq m - \binom{M}{2} \\ &= m - \frac{[(m-w)/2][(m-w)/2-1]}{2} - \frac{m-w}{2} \\ &= m - \frac{(m-w)^2 - 2m - 2w}{8} \\ &= \frac{-(m-w)^2 + 6m + 2w}{8} > 0. \end{aligned}$$

It follows that codes of low weight, hence curves with many points, can be constructed using this method.

References

- [1] MacWilliams, F.J. and N.J.A. Sloane. The Theory of Error-Correcting Codes. North-Holland, 1977.
- [2] Shabat, Vasily, *Curves with Many Points*. Thesis, 2001.
- [3] Van der Geer, Gerard and Marcel van der Vlugt, *Quadratic forms, generalized Hamming weights of codes and curves with many points*, J. Number Theory, **59** (1996), no. 1, 20-36.