Introduction

L'Enseignement Mathématique, t. 48 (2002), p. 3–21

THE COSET WEIGHT DISTRIBUTIONS OF CERTAIN BCH CODES AND A FAMILY OF CURVES

by G. VAN DER GEER and M. VAN DER VLUGT

INTRODUCTION

Many problems in coding theory are related to the problem of determining the distribution of the number of rational points in a family of algebraic curves defined over a finite field. Usually, these problems are very hard and a complete answer is often out of reach.

In the present paper we consider the problem of the weight distributions of the cosets of certain BCH codes. This problem turns out to be equivalent to the determination of the distribution of the number of points in a family of curves with a large symmetry group. The symmetry allows us to analyze closely the nature of these curves and in this way we are able to extend considerably our control over the coset weight distribution compared with earlier results.

For a binary linear code $\mathcal{C}$ of length n the weight distributions of the cosets of $\mathcal{C}$ in $\mathbf{F}_2^n$ are important invariants of the code. They determine for example the probability of a decoding error when using $\mathcal{C}$. However, the coset weight distribution problem is solved for very few types of codes.

In [C-Z] Charpin and Zinoviev study the weight distributions of the cosets of the binary 3-error-correcting BCH code of length $n = 2^m - 1$ with m odd. We denote this code by $BCH(3)$.

Let $\mathbf{F}_q$ be a finite field of cardinality $q = 2^m$ and let α be a generator of the multiplicative group $\mathbf{F}_q^*$. The matrix

$$H = \begin{pmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{n-1} \\ 1 & \alpha^3 & \alpha^6 & \dots & \alpha^{3(n-1)} \\ 1 & \alpha^5 & \alpha^{10} & \dots & \alpha^{5(n-1)} \end{pmatrix}$$

is a parity check matrix defined over $\mathbf{F}_q$ of $BCH(3)$. This means that

$$BCH(3) = \{c = (c_0, \ldots, c_{n-1}) \in \mathbf{F}_2^n : Hc^t = 0\}.$$

It was shown in [C-Z] that the coset weight distribution problem for $BCH(3)$ comes down to the same problem for the extended code $\widehat{BCH(3)}$ with parity check matrix

$$\widehat{H} = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 & 1 \\ 1 & \alpha & \alpha^2 & \ldots & \alpha^{n-1} & 0 \\ 1 & \alpha^3 & \alpha^6 & \ldots & \alpha^{3(n-1)} & 0 \\ 1 & \alpha^5 & \alpha^{10} & \ldots & \alpha^{5(n-1)} & 0 \end{pmatrix}$$

A coset $\widehat{D}$ of $\widehat{BCH(3)}$ in $\mathbf{F}_2^{n+1}$ is characterized by the syndrome $s(\widehat{D}) = \widehat{H}x^t \in \mathbf{F}_q^4$, where x is a representative of $\widehat{D}$. The weight of $\widehat{D}$ is the minimum weight of the vectors in $\widehat{D}$. Here the weight of a vector is the number of its non-zero entries.

Charpin and Zinoviev then show that the weight distribution problem for the cosets of $\widehat{BCH(3)}$ of length 2^m with m odd can be solved as soon as the weight distributions of the cosets $\widehat{D}_4$ of weight 4 with syndrome $s(\widehat{D}_4) = (0, 1, A, B)$ are determined.

From [C-Z] we recall: *The weight distribution of a coset $\widehat{D}_4$ is determined by the number $N(A, B)$ of vectors of weight 4 in $\widehat{D}_4$.*

Via the matrix $\widehat{H}$ this leads to the system of equations in four variables in $\mathbf{F}_{q=2^m}$:

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1, \\ x_1^3 + x_2^3 + x_3^3 + x_4^3 &= A, \\ x_1^5 + x_2^5 + x_3^5 + x_4^5 &= B, \end{aligned} \tag{1}$$

and $N(A, B)$ is the number of S_4-orbits of solutions of (1) with distinct $x_i \in \mathbf{F}_q$. In particular the number of values of $N(A, B) > 0$ equals the number of different coset weight distributions of cosets of type $\widehat{D}_4$. Note that since the set of solutions of (1) is invariant under translation over $(1, 1, 1, 1)$ the quantity $N(A, B)$ is even.

In this paper we shall show that by analyzing carefully the curves defined by (1) we can determine good upper and lower bounds for the pivotal quantity $N(A, B)$. The bounds are obtained by dissecting the Jacobian variety of the curves in our family in isogeny factors of dimension 1 and 2. This yields restrictions on the traces of Frobenius. The splitting of the Jacobian is a corollary from a very effective description of the curves defined by (1) as fibre products over $\mathbf{P}^1$ of three elliptic curves.

We show that for odd m the $N(A,B)$ lie in an explicit interval of length $\sim 1.57\sqrt{q}$, cf. [C-Z], where the interval is $\sim q/4$. Moreover, we argue that on statistical grounds one may expect that almost all $N(A,B)$ lie in an explicit interval of length $\sim 0.9\sqrt{q}$. We then give numerical results that confirm strongly these heuristics and extend the table of $BCH(3)$ codes with known coset weight distribution.

For an introduction to the theory of codes we refer to [vL] and for a general introduction to curves over finite fields to [S]. The reader can find basic facts about Jacobians in the survey paper [Mi] and a general introduction to curves and their Jacobians in [Mu].

§1. A FAMILY OF CURVES

We consider the algebraic curve $C' = C'_{A,B}$ in $\mathbf{P}^4$ given by the equations

$$s_1 = x_0\,, \quad s_3 = Ax_0^3\,, \quad s_5 = Bx_0^5\,, \tag{2}$$

where s_j is the j-th power sum $\sum_{i=1}^4 x_i^j$ in the variables $x_1,\ldots,x_4$. Let σ_j denote the j-th elementary symmetric function in $x_1,\ldots,x_4$. If we apply Newton's formulas for power sums we find

$$\begin{aligned} s_1 + x_0 &= \sigma_1 + x_0 = 0\,,\\ s_3 + Ax_0^3 &= (A+1)x_0^3 + \sigma_2 x_0 + \sigma_3 = 0\,,\\ s_5 + Bx_0^5 &= x_0\left((B+A)x_0^4 + (A+1)\sigma_2 x_0^2 + \sigma_4\right) = 0\,. \end{aligned}$$

This implies that the curve C' consists of the three lines in the hyperplane $x_0 = 0$ given by

$$x_i + x_j = x_k + x_l = 0\,, \quad \text{with} \quad \{i,j,k,l\} = \{1,2,3,4\}\,, \tag{3}$$

and a curve $C = C_{A,B}$ given by

$$\begin{aligned} \sigma_1 &= x_0\,,\\ \sigma_3 &= (A+1)x_0^3 + \sigma_2 x_0\,,\\ \sigma_4 &= (B+A)x_0^4 + (A+1)\sigma_2 x_0^2\,. \end{aligned} \tag{4}$$

The symmetric group S_4 operates on C' and on C by permuting the coordinates $x_1,\ldots,x_4$. Moreover, there is an involution τ acting on C via

$$(x_0 : x_1 : \ldots : x_4) \mapsto (x_0 : x_1 + x_0 : \ldots : x_4 + x_0)\,.$$