Generalized Bezout's Theorem and Its Applications in Coding Theory

Gui-Liang Feng, T. R. N. Rao*, and Gene A. Berg†

July 17, 1996

Abstract

This paper presents a generalized Bezout theorem which can be used to determine a tighter lower bound of the number of distinct points of intersection of two or more curves for a large class of plane curves. A new approach to determine a lower bound on the minimum distance (and also the generalized Hamming weights) for algebraic-geometric codes defined from a class of plane curves is introduced, based on the generalized Bezout theorem. Examples of more efficient linear codes are constructed using the generalized Bezout theorem and the new approach. For $d = 4$, the linear codes constructed by the new construction are better than or equal to the known linear codes. For $d \geq 5$, these new codes are better than the known codes. The Klein code [22, 16, 6] over $GF(2^3)$ is also constructed.

Index Terms: algebraic-geometric codes, linear codes, minimum distance, generalized Hamming weights, Bezout's theorem.

1 Introduction

The minimum distance is one of most important parameters in error-correcting codes. It is a measure of the code's capacity to correct errors or detect errors or both [1]. The minimum distance d of a linear code C is defined by

$$d = \min_{\substack{\mathbf{u},\mathbf{v}\in C \\ \mathbf{u}\neq\mathbf{v}}} \{d(\mathbf{u},\mathbf{v})\},$$

where $d(\mathbf{u},\mathbf{v})$ expresses the Hamming distance between $\mathbf{u}$ and $\mathbf{v}$.

From this definition, we have the following lemma.

*Gui-Liang Feng, T. R. N. Rao are with the Center for Advanced Computer Studies, University of Southwestern Louisiana, Lafayette, LA. 70504, USA. email: glf@swamp.cacs.usl.edu and rao@swamp.cacs.usl.edu. This work was supported in part by the National Science Foundation under Grant NCR-9505619, Louisiana Education Quality Support Fund under Grant LEQSF-(1994-96)-RD-A-36, and NASA project under Grant NAG-W-4013.

†Gene A. Berg is with the National Security Agency, Fort George G. Meade, MD 20755-6000, email: geneberg@romulus.ncsc.mil.

Lemma 1.1 *Let C be an (n,k) linear code and $\mathbf{H}$ be a parity check matrix. If any (d^*-1) columns of $\mathbf{H}$ have rank (d^*-1), then the minimum distance d is at least d^*, i. e., $d \geq d^*$.*

Recently, another important sequence of parameters called the *generalized Hamming weight hierarchy* of linear codes was introduced by V. K. Wei [2]. Currently, at least four practical motivations for studying the weight hierarchies are known. The generalized Hamming weights are useful in two cryptographical applications (wire-tap channel and t-resilient function [2]), in trellis coding (lower bounding the number of trellis states [3-4]), and in truncating a linear block code [5]. Recently, G. D. Forney has shown that the dimension/length profile (DLP) of a linear code is equivalent to its generalized Hamming weight hierarchy [33]. Forney pointed out that these two concepts should be regarded as parts of the same subject and that these concepts can be used to easily prove some known results and further derive new results. Many interesting questions arise on weight hierarchies of linear codes. Hence, research on this topic is very active. Here we review some basic definitions. Let C be an (n,k) linear code and D be a subcode. The *support* of D, denoted $\chi(D)$, is the set of the not-always-zero bit positions of D, i.e., $\chi(D) := \{i$: there is a $(x_1, x_2, ..., x_n) \in D$, $x_i \neq 0\}$.

Following this definition, an (n,k) code is a binary linear code of rank or dimension k, and support size $\leq n$. The h-th generalized Hamming weight of C is then defined to be $d_h(C) := \min\{|\chi(D)| : D$ is an h-dimensional subcode of $C\}$. It is easy to see that $d_1(C)$ denotes the traditional minimum distance or minimum Hamming weight of the code. The weight hierarchy of C can then be defined to be the set of integers $\{d_h(C) : 1 \leq h \leq k\}$. From the definition, we have the following lemma.

Lemma 1.2 *([7]): Let C be an (n,k) linear code and $\mathbf{H}$ be a parity check matrix. If any (d^*-1) columns of $\mathbf{H}$ have rank (d^*-h), then the h-th generalized Hamming weight $d_h(C)$ is at least d^*, i. e., $d_h(C) \geq d^*$.*

When $h = 1$, Lemma 1.2 reduces to Lemma 1.1. It is very difficult to use Lemma 1.2 directly to determine the lower bound d^*. In the following we introduce a new concept and reduce Lemma 1.2 to another form, which allows d^* to be more easily determined.

Let $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_p$, and $\mathbf{u}$ be n-tuple vectors. If there are p coefficients c_i such that $\mathbf{u}+\sum_{i=1}^{p} c_i\mathbf{v}_i = \mathbf{0}$, where $\mathbf{0}$ is the zero vector, then we say that $\mathbf{u}$ is totally linearly dependent on vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_p$. Sometimes, $\mathbf{u}$ may be linearly dependent on the vectors for only some of the components (i.e., locations). Then $\mathbf{u}$ is said to be partially linearly dependent on the vectors $\mathbf{v}_i$ for $1 \leq i \leq p$. The maximal possible number of those components (i.e., locations) can be used to measure the linear dependence of the vector $\mathbf{u}$ on the vectors $\mathbf{v}_i$, for $1 \leq i \leq p$. The number of components, for which $\mathbf{u}$ is partially linearly dependent on the vectors, is called the *dependent-degree* of $\mathbf{u}$ on $\mathbf{v}_i$, for $1 \leq i \leq p$. Apparently, if the dependent-degree is equal to n, then $\mathbf{u}$ is totally linearly dependent on $\mathbf{v}_i$ for $1 \leq i \leq p$.

Example 1.1 *Let $\mathbf{u} = (1,1,1,0,0), \mathbf{v}_1 = (1,0,0,1,1)$ and $\mathbf{v}_2 = (0,1,1,1,0)$ over GF(2). Then $\mathbf{u}+\mathbf{v}_1+\mathbf{v}_2 = (0,0,0,0,1)$, $\mathbf{u}+\mathbf{v}_1 = (0,1,1,1,1)$, $\mathbf{u}+\mathbf{v}_2 = (1,0,0,1,0)$, and $\mathbf{u} = (1,1,1,0,0)$. From these four vectors, we see that the vector $(0,0,0,0,1)$ has the maximum number of zeros $(=4)$. Hence, the dependent degree of $\mathbf{u}$ on $\mathbf{v}_1$ and $\mathbf{v}_2$ is equal to 4.*

We generalize this concept to the case of a sequence of vectors $\mathbf{u}_i$. Let us consider two sequences of vectors $\mathbf{u}_i$ for $1 \leq i \leq p$, and vectors $\mathbf{v}_j$ for $1 \leq j \leq q$. Let there be some components, on which $\mathbf{u}_\mu$ $(1 \leq \mu \leq p)$ are partially linearly dependent on $\mathbf{v}_j$ for $1 \leq j \leq q$ and $\mathbf{u}_i$ for $1 \leq i < \mu$. The number of such components can be used to measure the consistent linear dependence of the vector $\mathbf{u}_1, ..., \mathbf{u}_p$ on vectors $\mathbf{v}_j$ for $1 \leq j \leq q$. The maximal possible number of such components is called the *consistent dependent-degree* of $\mathbf{u}_1$, ..., $\mathbf{u}_p$ on the vectors $\mathbf{v}_j$ for $1 \leq j \leq q$.

Example 1.2 *Let* $\mathbf{u}_1 = (1,1,1,0,0)$, $\mathbf{u}_2 = (0,1,1,0,1)$, $\mathbf{v}_1 = (1,0,0,1,1)$ *and* $\mathbf{v}_2 = (0,1,1,1,0)$ *over GF(2). Since* $\mathbf{u}_1+\mathbf{v}_2 = (1,0,0,1,0)$, $\mathbf{u}_2+\mathbf{v}_1+\mathbf{v}_2 = (1,0,0,0,0)$, *they have zeros at the second, third, and fifth components, and it can be easily checked that at most three components have zeros simultaneously. Therefore, the consistent dependent-degree of* $\mathbf{u}_1$, $\mathbf{u}_2$ *on* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *is equal to 3.*

For a sequence of linearly independent vectors $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r, ...\}$, let $\mathbf{v}_i^*$ express a linear combination $\mathbf{v}_i + \sum_{\mu=1}^{i-1} c_\mu \mathbf{v}_\mu$.

Definition 1.1 $D_{\{\mathbf{v}_{i_1}^*,\cdots,\mathbf{v}_{i_p}^*\}}$ *denotes the maximal consistent dependent-degree of a set of* $\{\mathbf{v}_{i_1}, \cdots, \mathbf{v}_{i_p}\}$ *on their previous vectors, respectively, i.e.,* $D_{\{\mathbf{v}_{i_1}^*,\cdots,\mathbf{v}_{i_p}^*\}}$ *denotes the maximal number of components (i.e., locations), on which* $\mathbf{v}_{i_\mu}^*$ *for* $1 \leq \mu \leq p$ *are all zero simultaneously.*

Definition 1.2 $D_p^{(r)} = \max\{D_{\{\mathbf{v}_{i_1}^*,\cdots,\mathbf{v}_{i_p}^*\}} | i_1 < \cdots < i_p \leq r\}$.

Remark (1): Let C_r be an $(n, n-r)$ linear code defined by a parity check matrix $\mathbf{H}_r = [\mathbf{h}_1, ..., \mathbf{h}_r]^T$, i.e. the parity check matrix has r rows. Then $D_p^{(r)} = n - d_p(C_r^\perp)$, where $C_r^\perp$ is the dual code of C_r, and $d_p(C_r^\perp)$ is the p-th generalized Hamming weight of $C_r^\perp$.

Using this concept, the determination of the generalized Hamming weights reduces to the calculation of $D_p^{(r)}$ for any given r vectors of a parity check matrix $\mathbf{H}$. We have the following theorem:

Theorem 1.1 *For a linear code* C_r *defined by* $\mathbf{H}_r$, *i.e., the parity check matrix has r rows, if the consistent dependent-degree of any* $(r - d^* + h + 1)$ *rows of* $\mathbf{H}_r$ *is always less than* $(d^* - 1)$, *i.e.,* $D_{r-d^*+h+1}^{(r)} < d^* - 1$, *then the h-th generalized Hamming weight* $d_h(C_r)$ *is at least* d^*, *i.e.,* $d_h(C_r) \geq d^*$.

Proof: Assume that there is a submatrix consisting of $(d^* - 1)$ columns of $\mathbf{H}_r$ with rank $(d^* - \nu)$, where $\nu \geq h + 1$. Then there are $(r - d^* + \nu)$ rows, which are linearly dependent on their previous rows, i.e. the consistent dependent-degree of these $(r - d^* + \nu)$ rows is at least $(d^* - 1)$. Thus, there are $(r - d^* + h + 1)$ rows, the consistent dependent-degree of which is at least $(d^* - 1)$, i.e., $D_{r-d^*+h+1}^{(r)} \geq d^* - 1$. This is a contradiction. Therefore, the rank is at least $(d^* - h)$. Using Lemma 1.2, the h-th generalized Hamming weight of C_r is at least d^*. □

Corollary 1.1 *Consider a linear code C_r defined by $\mathbf{H}_r$, i.e., the parity check matrix has r rows. If the consistent dependent-degree of any $(r-d^*+2)$ rows of $\mathbf{H}_r$ is always less than (d^*-1), i.e., $D^{(r)}_{r-d^*+2} < d^*-1$, then the minimum distance d of C_r is at least d^*, i.e., $d \geq d^*$.*

Let LS be a set of distinct points in a plane or a set of distinct roots of a plane curve (i.e. a polynomial). Let $LS = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$ and let $h(x, y)$ be a polynomial (or monomial), then a vector $(h(x_1, y_1), h(x_2, y_2), ..., h(x_n, y_n))$ is called an evaluated vector of polynomial $h(x, y)$ on the set LS. Hereinafter, sometimes $\mathbf{v}_i$ expresses an evaluated vector and sometimes it expresses a polynomial or a curve if no confusion arises. Thus, from Definition 1.1, $D_{\{\mathbf{v}^*_{i_1},...,\mathbf{v}^*_{i_p}\}}$ denotes the number of distinct points of the intersection of curves $\mathbf{v}^*_{i_1} = 0$, ..., $\mathbf{v}^*_{i_p} = 0$ for the case of LS being the set of all points in a whole plane, or denotes the number of distinct points of the intersection of curves $\mathbf{v}^*_{i_1} = 0$, ..., $\mathbf{v}^*_{i_p} = 0$, and $f(x, y) = 0$ for the case of LS being the set of all points on the curve $f(x, y) = 0$. Similarly, $D^{(r)}_p$ for a given sequence of evaluated vectors expresses the maximal possible number of distinct points of the intersection of p curves among the first r curves of the given sequence of curves. Therefore, the calculation of $D^{(r)}_p$ reduces to the calculation of the number of distinct points of intersection of several curves. Bezout's theorem is the most well-known theorem that can be applied to solving the problem of determining the number of points of the intersection of two curves.

The organization of this paper is as follows. In the next section, we present a generalized Bezout theorem, which can be used to determine an upper bound of the number of points of intersection of several plane curves. We also derive some properties of $D_{\{\mathbf{v}^*_{i_1},...,\mathbf{v}^*_{i_p}\}}$ and $D^{(r)}_p$. In Section III, we use the generalized Bezout theorem to derive a lower bound for the generalized Hamming weights of the algebraic-geometric codes derived from a large class of plane curves. In Section IV, we introduce a new construction, by which many more efficient linear codes (include the algebraic-geometric codes) can be easily constructed. Some conclusions are included in Section V.

2 Generalized Bezout Theorem

In this section, we first introduce Bezout's theorem. Then we derive a generalized Bezout theorem and some properties of $D_{\{\mathbf{v}^*_{i_1},...,\mathbf{v}^*_{i_p}\}}$ and $D^{(r)}_p$, which are useful for the next sections.

Theorem 2.1 (Bezout) [15] *Let $F(X, Y, Z)$ and $G(X, Y, Z)$ be two plane curves without common components, i.e., F and G are homogeneous forms of degree n and m, respectively, with no common factors. Then the intersection of $F(X, Y, Z)$ and $G(X, Y, Z)$ has at most mn distinct common points (X, Y, Z).*

Definition 2.1 *The x-resultant matrix, denoted by $RM(f, g)$ (or RM) of two polynomials*

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$$

$$g(x) = b_0x^m + b_1x^{m-1} + \cdots + b_m$$

is given by the following $(m+n) \times (m+n)$ *matrix:*

$$\begin{bmatrix}
a_0 & a_1 & \dots & \dots & \dots & a_n & & & & \\
 & a_0 & a_1 & \dots & \dots & \dots & a_n & & & \\
 & & . & . & . & . & . & . & & \\
 & & & . & . & . & . & . & . & \\
 & & & & a_0 & a_1 & \dots & \dots & \dots & a_n \\
b_0 & b_1 & \dots & \dots & b_m & & & & & \\
 & b_0 & b_1 & \dots & \dots & b_m & & & & \\
 & & . & . & . & . & . & & & \\
 & & & . & . & . & . & . & & \\
 & & & & . & . & . & . & . & \\
 & & & & & b_0 & b_1 & \dots & \dots & b_m
\end{bmatrix},$$

and its determinant is called the x*-resultant of the two polynomials and denoted by* $Res_x(f,g)$ *(or* R*).*

Theorem 2.2 *$f(x)$ and $g(x)$ have a greatest common divisor polynomial with degree r if and only if rank$(RM(f,g)) = m+n-r$.*

Before giving a proof of Theorem 2.2, we introduce the following notations and lemmas.

Lemma 2.1 $RM \times [x^{n+m-1}, x^{n+m-2}, ..., x, 1]^T =$

$$[f(x)x^{m-1}, f(x)x^{m-2}, ..., f(x)x, f(x), g(x)x^{n-1}, g(x)x^{n-1}, ..., g(x)x, g(x)]^T.$$

Lemma 2.2 *gcd $(f(x), g(x)) = 1$, if and only if rank$(RM(f,g)) = m+n$.*

Proof: Since $\gcd(f(x), g(x)) = 1$, there exist no $P(x)$ and $Q(x)$ with $\deg P(x) \leq m-1$ and $\deg Q(x) \leq n-1$ such that $f(x)P(x) + g(x)Q(x) = 0$. And from Lemma 2.1, there exists no non-zero linear combination of the rows of $RM(f,g)$ to be zero vector. That implies that $RM(f,g)$ is a nonsingular matrix. □

Proof of Theorem 2.2: Assume that $f(x)$ and $g(x)$ have r common roots. Let $h(x) \equiv x^r + \sum_{i=1}^{r} c_i x^{r-i}$ be the greatest common factor of $f(x)$ and $g(x)$. Let $f(x) = (x^r + \sum_{i=1}^{r} c_i x^{r-i}) \times (a_0 x^{n-r} + \sum_{i=1}^{n-r} a_i^* x^{n-r-i})$ and $g(x) = (x^r + \sum_{i=1}^{r} c_i x^{r-i})(b_0 x^{m-r} + \sum_{i=1}^{m-r} b_i^* x^{m-r-i})$.

Let $f^*(x)$ and $g^*(x)$ denote $a_0 x^{n-r} + \sum_{i=1}^{n-r} a_i^* x^{n-r-i}$ and $b_0 x^{m-r} + \sum_{i=1}^{m-r} b_i^* x^{m-r-i}$, respectively. Then $f^*(x)$ and $g^*(x)$ have no common roots, i.e. $gcd(f^*(x), g^*(x)) = 1$. Without loss of generality, let $b_{m-r}^* \neq 0$. Thus, $RM(f,g)$ can be decomposed into a product of two matrices, i.e., $RM = Q \times P^{-1}$, where Q is as follows:

$$\begin{bmatrix}
a_0 & a_1^* & \dots & a_{n-r}^* & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 & a_0 & a_1^* & \dots & a_{n-r}^* & 0 & 0 & 0 & 0 & \dots & 0 \\
 & & . & . & . & . & & & & & \\
 & & & . & . & . & . & & & & \\
 & & & & a_0 & a_1^* & \dots & a_{n-r}^* & 0 & \dots & 0 \\
b_0 & b_1^* & \dots & b_{m-r}^* & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\
 & b_0 & b_1^* & \dots & b_{m-r}^* & 0 & 0 & 0 & 0 & \dots & 0 \\
 & & . & . & . & & & & & & \\
 & & & . & . & . & . & & & & \\
 & & & & b_0 & b_1^* & \dots & b_{m-r}^* & 0 & \dots & 0
\end{bmatrix},$$

and P^{-1} is as shown:

$$\begin{bmatrix}
1 & c_r & c_{r-1} & \cdots & c_2 & c_1 & & & & \\
 & 1 & c_r & c_{r-1} & \cdots & c_2 & c_1 & & & \\
 & & 1 & c_r & c_{r-1} & \cdots & c_2 & c_1 & & \\
 & & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\
 & & & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 & & & & & \cdot & \cdot & \cdot & \cdot & \cdot \\
 & & & & & & \cdot & \cdot & \cdot & \cdot \\
 & & & & & & & 1 & c_r & c_{r-1} & c_{r-2} \\
 & & & & & & & & 1 & c_r & c_{r-1} \\
 & & & & & & & & & 1 & c_r \\
 & & & & & & & & & & 1
\end{bmatrix}.$$

The matrix P^{-1} is a nonsingular matrix. The last r columns of the matrix Q are all zero. The other $n+m-r$ columns form a submatrix denoted by Q'. If we delete the last r rows from Q', the upper part consists of the first $m-r$ rows of Q', then we obtain the following matrix:

$$\begin{bmatrix}
 & & & & & & 0 & \cdots & 0 \\
 & & & & & & 0 & \cdots & 0 \\
 & & RM(f^*(x), g^*(x)) & & & & 0 & \cdots & 0 \\
 & & & & & & 0 & \cdots & 0 \\
 & & & & & & 0 & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & b^*_{m-r} & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & b^*_{m-r}
\end{bmatrix}.$$

Since gcd $(f^*(x), g^*(x)) = 1$, using Lemma 2.2, the left upper submatrix is a nonsingular $(m+n-2r) \times (m+n-2r)$ matrix. On the other hand, $b^*_{m-r} \neq 0$. Hence the matrix Q' is a full rank matrix. Therefore, rank$(Q) = n+m-r$, i.e., the rank(RM) is equal to $n+m-r$.

Conversely, if rank$(RM) = n+m-r$, then from Lemma 2.2, $f(x)$ and $g(x)$ have the greatest common divisor $h(x) = x^{r^*} + \sum_{\mu=1}^{r^*} d^*_\mu x^{r^*-\mu}$, where $r^* > 0$. Then, using the above proof, the rank$(RM) = n+m-r^*$. Thus, $r^* = r$. This completes the proof of Theorem 2.2. □

For convenience in the following discussion, we define

$$\vec{f}^{(0)} \equiv \vec{f} \equiv (a_0, a_1, ..., a_n, 0, ..., 0),$$

where on the rightmost side there are $(m-1)$ 0's, and

$$\vec{f}^{(i)} \equiv (0, ..., 0, a_0, a_1, ..., a_n, 0, ..., 0),$$

where on the leftmost side there are i 0's $(0 \leq i \leq m-1)$ and on the rightmost side there are $(m-i-1)$ 0's. Thus, the above matrix consists of the vectors $\vec{f}^{(\mu)}$ and $\vec{g}^{(\lambda)}$, for $0 \leq \mu \leq m-1$ and $0 \leq \lambda \leq n-1$.

Sometimes the x-resultant is called the Sylvester resultant because it was introduced by Sylvester [16]. In his paper, Sylvester showed that $Res_x(f,g) = 0$ if and only if either $a_0 = b_0 = 0$ or if f and g have a common root.

The coefficients of f and g could be polynomials in y. We could have:

$$f(x,y) = a_0(y)x^m + a_1(y)x^{m-1} + \cdots + a_m(y),$$

$$g(x,y) = b_0(y)x^n + b_1(y)x^{n-1} + \cdots + b_n(y).$$

Then for $R(y) = Res_x(f,g)$ and for any value β of y we would have $R(\beta) = 0$ if and only if either $a_0(\beta) = b_0(\beta) = 0$ or $f(\alpha,\beta) = g(\alpha,\beta) = 0$ for some α. Thus, the roots of $R(y) = 0$ are the projections of the points of intersection of f and g. In fact, the resultant gives more precise information. Namely, if the order of the zero β of $R(y)$ is e, i.e., if $R(y) = (y-\beta)^e D(y)$ with $D(\beta) \neq 0$, then counting properly, there are exactly e points of intersection of $f(x,y)$ and $g(x,y)$ lying on $y = \beta$. If it is not counted properly, there are at most e distinct points of intersection of $f(x,y)$ and $g(x,y)$ lying on $y = \beta$. Therefore, we have the following theorem.

Theorem 2.3 *The number of distinct points of intersection of two polynomials $f(x,y)$ and $g(x,y)$ without common components is at most equal to the degree of their resultant $R(y)$.*

Proof: If β is a root of $R(y) = 0$ and its order is equal to r, then the $x-resultant$ matrix $RM(\beta)$ has rank $n+m-r$. From Theorem 2.2, there are at most r distinct values (denoted by α) of x such that (α,β) are the points in the intersection of $f(x,y) = 0$ and $g(x,y) = 0$, i.e., there are at most r common roots with $y = \beta$ of $f(x,y) = 0$ and $g(x,y) = 0$. Thus, for each root β with an order r of $R(y) = 0$, there are at most r distinct points (α,β) in the intersection of $f(x,y) = 0$ and $g(x,y) = 0$. On the other hand, the summation of all root orders of equation $R(y) = 0$ is equal to the degree of $R(y)$. Therefore, the number of distinct points in the intersection of $f(x,y) = 0$ and $g(x,y) = 0$ is at most the degree of $R(y)$. The proof of Theorem 2.3 is completed. □

From Theorem 2.3, the Bezout theorem can be derived. If $f(x,y)$ and $g(x,y)$ have no common components, then their resultant $R(y)$ is not identical with zero. Thus, deg $R(y) > -\infty$, we define deg $0 = -\infty$.

Now we will generalize Theorem 2.2. Let us consider p curves in affine plane curves without common components, i.e., $f_\mu(x,y) = 0$ for $\mu = 1, 2, ..., p$. Without loss of generality, $deg_x f_1 \geq deg_x f_2 \geq \cdots \geq deg_x f_p$, and let $deg_x f_1 = m$ and $deg_x f_2 = n$, where $deg_x f_\mu$ indicates the maximal i such that the monomial $x^i y^j$ is a term in f_μ. We define the x-resultant matrix of these p curves or polynomials as the following $\Sigma \times (m+n)$ matrix,

where $\Sigma = \sum_{\mu=1}^{p} (m + n - deg_x f_\mu)$ and $s = deg_x f_p$:

$$\begin{bmatrix}
a_0^{(1)} & a_1^{(1)} & . & . & . & a_m^{(1)} & 0 & . & . & 0 \\
0 & a_0^{(1)} & a_1^{(1)} & . & . & . & a_m^{(1)} & 0 & . & 0 \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
0 & 0 & . & 0 & a_0^{(1)} & a_1^{(1)} & . & . & . & a_m^{(1)} \\
a_0^{(2)} & a_1^{(2)} & . & . & a_n^{(2)} & 0 & 0 & . & . & 0 \\
0 & a_0^{(2)} & a_1^{(2)} & . & . & a_n^{(2)} & 0 & . & . & 0 \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
0 & 0 & . & . & . & 0 & a_0^{(2)} & a_1^{(2)} & . & a_n^{(2)} \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
a_0^{(p)} & a_1^{(p)} & . & . & a_s^{(p)} & 0 & 0 & . & . & 0 \\
0 & a_0^{(p)} & a_1^{(p)} & . & . & a_s^{(p)} & 0 & . & . & 0 \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
0 & 0 & . & . & . & 0 & a_0^{(p)} & a_1^{(p)} & . & a_s^{(p)}
\end{bmatrix}.$$

Let $R(y) = Res_x(f_1, f_2, \cdots, f_p)$ be the non-zero determinant of the nonsingular submatrix with the smallest degree of y of the x-resultant matrix. Similar to the proof of Theorem 2.3, we have the general theorem as follows.

Theorem 2.4 *The number of distinct points of the intersection of $f_\mu(x, y)$ without common components, for $\mu = 1, 2, ..., p$, is at most equal to the degree of their resultant $R(y)$, i.e., $degR(y)$.*

In order to get an upper bound of $degR(y)$, we introduce a new concept. Among the f's with the same degree of x, we select one. Thus, we can select f_{λ_μ}, for $\mu = 1, 2, ..., q(\leq p)$, such that $deg_x f_{\lambda_i} > deg_x f_{\lambda_{i+1}}$, $\{ deg_x f_{\lambda_\sigma} | \sigma = 1, 2, ..., q\} = \{deg_x f_\mu | \mu = 1, 2, ..., p\}$, and f_{λ_σ} have no common components. We define the x-partial resultant matrix of these p curves or polynomials as the following $(m + n) \times (m + n)$ matrix:

$$[\bar{f}_{\lambda_1}^{(0)}, ..., \bar{f}_{\lambda_1}^{(d_1+d_2-d_1-1)}, \bar{f}_{\lambda_2}^{(d_1+d_2-d_1)}, ..., \bar{f}_{\lambda_2}^{(d_1+d_2-d_2-1)}, ..., \bar{f}_{\lambda_q}^{(d_1+d_2-d_{q-1})}, ..., \bar{f}_{\lambda_q}^{(d_1+d_2-d_q-1)}]^T,$$

namely, $[\bar{f}_{\lambda_1}^{(0)}, ..., \bar{f}_{\lambda_1}^{(d_2-1)}, \bar{f}_{\lambda_2}^{(d_2)}, ..., \bar{f}_{\lambda_2}^{(d_1-1)}, ..., \bar{f}_{\lambda_q}^{(d_1+d_2-d_{q-1})}, ..., \bar{f}_{\lambda_q}^{(d_1+d_2-d_q-1)}]^T$, where d_σ denotes $deg_x f_{\lambda_\sigma}$.

Obviously, this matrix is an upper triangle matrix when $d_q = 0$. The determinant of this matrix can be easily calculated for the special case, i.e., the determinant is equal to the product of all elements on main diagonal of this matrix. This determinant is called a partial resultant and denoted by $PR(y)$.

Corollary 2.1 *The number of distinct points of the intersection of* $f_\mu(x,y)$, *for* $\mu = 1,2,...,p$, *is at most equal to the degree of their partial resultant* $PR(y)$.

Example 2.1 *Let us consider the number of common points on the following four curves over* $GF(2^4)$:

$$\begin{cases} x^5+y^4+y=0 \\ x^3+a(y)x^2+b(y)x+c(y)=0 \\ xy+e(y)=0 \\ y^2+fy+g=0 \end{cases},$$

We have the following matrix:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & y^4+y & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & y^4+y & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & y^4+y \\ 0 & 0 & 0 & 1 & a(y) & b(y) & c(y) & 0 \\ 0 & 0 & 0 & 0 & 1 & a(y) & b(y) & c(y) \\ 0 & 0 & 0 & 0 & 0 & y & e(y) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & y & e(y) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & y^2+fy+g \end{bmatrix}.$$

Thus, $PR(y) = y^2(y^2+fy+g)$. Obviously, $degPR(y) = 4$. Therefore, the number of distinct points of the intersection of the four curves is at most 4.

Remark (2): Here we regard $f_\mu(x,y)$ as a polynomial of x and the coefficients are polynomials in y. We also can regard $f_\mu(x,y)$ as a polynomial of y and the coefficients are polynomials in x. The number of the distinct points of intersection of $f_\mu(x,y)$'s is the same. The distinct points of intersection of $f_\mu(x,y)$'s obtained by the two approachs are also the same.

Remark (3): It is sufficient and necessary that f_μ, for $\mu = 1,2,...,p$, have no common components.

Definition 2.2 $D_{\{f_1,f_2,\cdots,f_p\}}$ *denotes the number of distinct points of the intersection of curves* $f_\mu(x,y)=0$, *for* $\mu = 1,2,...,p$.

Definition 2.3 *Given a sequence of polynomials* $\{f_\mu(x,y)|\mu = 1,2,...,r\}$.

$$D_p^{(r)} = \max\{D_{\{f^*_{\lambda_1},f^*_{\lambda_2},\cdots,f^*_{\lambda_p}\}}|\lambda_1,\cdots,\lambda_p \le r\},$$

where $f^*_{\lambda_\mu}$ *expresses a linear combination of* f_i *for* i =1, 2, ..., λ_μ, *and the coefficient of* f_{λ_μ} *is 1, i.e.,* $f^*_{\lambda_\mu} = f_{\lambda_\mu} + \sum_{i=1}^{\lambda_\mu - 1} c_i f_i$.

We have the following results:

Proposition 2.1 $D_{\{\cdots,f(x,y)g(x,y),\cdots\}} \le D_{\{\cdots,f(x,y),\cdots\}} + D_{\{\cdots,g(x,y),\cdots,\}}$.

Proof: The set of all roots of $f(x,y)g(x,y)=0$ is a union of the set of all roots of $f(x,y)=0$ and the set of all roots of $g(x,y)=0$. We have Proposition 2.1. □

Proposition 2.2 $D_{\{f_1,...,f_p\}} \leq \min\{D_{\{f_\mu\}} | \mu = 1,2,...,p\}$.

Proof: All the points of intersection of $f_\mu(x,y) = 0$, for $\mu = 1,2,...,p$, are the points of $f_\mu(x,y) = 0$, respectively. Therefore, we have Proposition 2.2. □

From Proposition 2.1 and Proposition 2.2, we have:

Proposition 2.3 $D_{\{gf_1,...,gf_p\}} \leq D_{\{g\}} + D_{\{f_1,...,f_p\}}$.

Proposition 2.4 $D_{\{gf_1,f_1,...,f_p\}} = D_{\{f_1,...,f_p\}}$.

Proposition 2.5 $D_p^{(r)} \geq D_{p+1}^{(r)} + 1$.

Proof: Assume $D_{p+1}^{(r)} = D_{\{f^*_{\lambda_1},f^*_{\lambda_2},...,f^*_{\lambda_p},f^*_{\lambda_{p+1}}\}}$, where $\lambda_{p+1} \leq r$. Let (x',y') not be in the intersection of the $p+1$ curves, i.e., $f_{\lambda_\mu}(x',y')$ are not all equal to zero, for $\mu = 1,2,...,p$, $p+1$. Without loss of the generality, let $f^*_{\lambda_1}(x',y') \neq 0$. We denote $f^*_{\lambda_\mu}(x',y') = v_\mu$ for $\mu = 1,2,...,p,p+1$. Thus, $v_1 \neq 0$. Now we define $f'_{\lambda_\mu} = f^*_{\lambda_\mu} - \frac{v_\mu}{v_1} f^*_{\lambda_1}$, for $\mu = 2,3,...,p,p+1$. Thus, we have $f'_{\lambda_\mu}(x',y') = 0$ for $\mu = 2,3,...,p,\ p+1$. It is easily seen that if $f^*_{\lambda_\mu}(x^*,y^*) = 0$ for $\mu = 1,2,3,...,p,p+1$, then $f'_{\lambda_\mu}(x^*,y^*) = 0$ for $\mu = 2,3,...,p,p+1$. Therefore, $D_{\{f'_{\lambda_2},...,f'_{\lambda_p},f'_{\lambda_{p+1}}\}} \geq D_{p+1}^{(r)} + 1$. From the definition of $D_p^{(r)}$, we have $D_p^{(r)} \geq D_{p+1}^{(r)} + 1$. The proof is completed. □

Remark (4): Proposition 2.5 corresponds to the monotonicity of the generalized Hamming weights (See Theorem 1 in [2]).

3 The Generalized Hamming Weights of AG Codes from a Large Class of Plane Curves

We are now interested in the following irreducible curves [17-18]:

$$x^a + y^b + f(x,y) = 0, \tag{1}$$

where $(a,b) = 1$, and $bi + aj < ab$ for any x^iy^j being a term in $f(x,y)$. Miura-Kamiya curves are special cases of (1) [19]. Since they are irreducible, any set containing one of these polynomials has no common non-constant factor. The results of this section can be generalized to the curves of (1), but for convenience of exposition we derive them using the following Hermitian curve over $GF(2^4)$ as an example:

$$x^5 + y^4 + y = 0. \tag{2}$$

For (2), we define the weights of monomials as follows: $w(x) = 4$, $w(y) = 5$ and $w(x^iy^j) = 4i + 5j$. We have the following sequence of monomials:

$H = \{\ 1,\ x,\ y,\ x^2,\ xy,\ y^2,\ x^3,\ x^2y,\ xy^2,\ y^3,\ x^4,\ x^3y,\ x^2y^2,\ xy^3,\ x^5,\ x^4y,\ x^3y^2,\ x^2y^3,\ x^6,\ x^5y,\ x^4y^2,\ ...\ \} = \{\ x^iy^j | 0 \leq i \leq 15, 0 \leq j \leq 3\} = \{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3, ..., \mathbf{h}_r, ..., \mathbf{h}_{64}\ \}$.

It can be checked that the weights of monomials in H form an ascending sequence: $W = \{0,4,5,8,9,10,12,13,14,15,...,62,63,65,66,67,70,71,75\}$.

Let $L(r)$ be the linear space spanned by the first r monomials of H. Obviously $\mathbf{h}_r \in L(r) - L(r-1)$. If polynomials $f(x,y)$, $g(x,y) \in L(r) - L(r-1)$, we say $f(x,y)$ and $g(x,y)$ are *consistent* and write $f(x,y) \sim g(x,y)$. In this paper, $[x^iy^j]$ (or $\mathbf{h}_r^*$) denotes all polynomials that are linear combinations of x^iy^j (or $\mathbf{h}_r$) and its previous monomials in which the coefficient of x^iy^j (or $\mathbf{h}_r$) is 1, i.e., $\mathbf{h}_r^* \equiv \mathbf{h}_r + \sum_{\mu=1}^{r-1} c_\nu \mathbf{h}_\mu$. Hence we have $[x^iy^j] \sim x^iy^j$ and $\mathbf{h}_r^* \sim \mathbf{h}_r$. For convenience, let $\mathbf{h}_0 = x^5 + y^4 + y$. Sometimes, if no confusion arises, $D_{\{\mathbf{h}_{\lambda_1}^*, \mathbf{h}_{\lambda_2}^*, \cdots, \mathbf{h}_{\lambda_p}^*\}}$ is represented as $D_{\{\lambda_1, \lambda_2, \cdots, \lambda_p\}}$. From these definitions and the results in Section II, we have the following lemmas.

Lemma 3.1 $D_{\{[x^iy^j]\}} \leq 4i + 5j$.

Proof: Let $\mathbf{h}_r = x^iy^j$ and consider any linear combination of the form $\mathbf{h}_r^* = x^iy^j + \sum_{\mu=1}^{r-1} c_\mu \mathbf{h}_\mu$. Each monomial $\mathbf{h}_\mu$, $1 \leq \mu < r$, has a y-exponent at most 3. Thus, $x^5 + y^4 + y$ is not a factor of $\mathbf{h}_r^*$. Since $x^5 + y^4 + y$ is irreducible, $\mathbf{h}_r^*$ and $x^5 + y^4 + y$ have no common factors. So Theorem 2.3 applies.

The x – *resultant* $R(y)$ of $x^5 + y^4 + y = 0$ and $x^iy^j + \cdots = 0$ is the determinant of the following matrix:

$$\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & y^4+y & 0 & 0 & \ldots & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & y^4+y & 0 & \ldots & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & y^4+y & \ldots & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & \ldots & 1 & 0 & \ldots & y^4+y \\
y^j & a(y) & b(y) & \ldots & c(y) & 0 & 0 & \ldots & 0 & 0 \\
0 & y^j & a(y) & b(y) & \ldots & c(y) & 0 & \ldots & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & \ldots & y^j & a(y) & \ldots & c(y)
\end{bmatrix},$$

where $deg\ a(y), deg\ b(y), \ldots, deg\ e(y)$ are all less than 4. Thus, $R(y) = (y^j)^5(y^4+y)^i + \cdots$, and $degR(y) = 4i + 5j$. The proof is completed. □

Lemma 3.2 *Let* $gcd(\mathbf{h}_{\lambda_1}, \ldots, \mathbf{h}_{\lambda_p}) = \mathbf{h}$. *Then* $D_{\{\mathbf{h}_{\lambda_1}^*, \ldots, \mathbf{h}_{\lambda_p}^*\}} \leq D_{\{\mathbf{h}\}} + D_{\{[x^{i_1}y^{j_1}], \ldots, [x^{i_t}y^{j_t}]\}}$, *where* $t \leq 4, 4 \geq i_1 > i_2 > \cdots > i_t = 0$, *and* $0 = j_1 < j_2 < \cdots < j_t \leq 3$.

Proof: Since $y^4 = x^5 + y$, and applying Proposition 2.3 and Proposition 2.4, we have Lemma 3.2. □

Example 3.1 *Let* $\mathbf{h}_{\lambda_\mu}$, *for* $\mu = 1, 2, \ldots, p$, *be* x^6, x^5, x^3y, x^4y^2, x^2y^2, xy^2. *Thus,* $gcd(x^6, x^5, x^3y, x^4y^2, x^2y^2, xy^2) = x$, *i.e.,* $\mathbf{h} = x$. *From Proposition 2.4,* x^6, x^4y^2, *and* x^2y^2 *can be deleted. Thus, from Lemma 3.2, we have*

$$D_{\{[x^6],[x^5],[x^3y],[x^4y^2],[x^2y^2],[xy^2]\}} \leq D_{\{[x]\}} + D_{\{[x^4],[x^2y],[y^2]\}}.$$

Therefore, $t = 3, i_\mu = 4, 2, 0$, *and* $j_\mu = 0, 1, 2$.

Theorem 3.1 $D_{\{[x^{i_1}y^{j_1}],...,[x^{i_t}y^{j_t}]\}} \leq \sum_{\mu=1}^{t-1}(i_\mu - i_{\mu+1})(j_{\mu+1} - j_1)$, *where* $t \leq 4, 4 \geq i_1 > i_2 > \cdots > i_t = 0$, *and* $0 = j_1 < j_2 < \cdots < j_t \leq 3$.

Proof: Since $deg_x(x^5 + y^4 + y) > deg_x[x^{i_1}y^{j_1}] > \cdots > deg_x[x^{i_t}y^{j_t}]$ and $i_t = 0$, we can construct $PR(y)$ and know that $degPR(y) = \sum_{\mu=0}^{t-1}(i_\mu - i_{\mu+1})j_{\mu+1} = \sum_{\mu=1}^{t-1}(i_\mu - i_{\mu+1})(j_{\mu+1} - j_1)$, where $j_1 = 0, i_0 = 5$. Thus, the proof is completed. □

Example 3.2 $D_{\{[x^4],[x^2y],[y^2]\}} \leq (4-2)(1-0) + (2-0)(2-0) = 6$.

Lemma 3.3 *If* $D_p^{(r)} = D_{\{s_1,s_2,...,s_p\}}$ *and* $\mathbf{h}_{t_\lambda}$ *is deleted, i.e.,* $t_\lambda \in \{1,2,...,r\} - \{s_1, s_2, ..., s_p\}$, *then all factors of* $\mathbf{h}_{t_\lambda}$ *should be deleted, i.e., it is not in the set* $\{\mathbf{h}_{s_1}, ..., \mathbf{h}_{s_p}\}$.

Proof: Suppose that

$$D_p^{(r)} = D_{\{s_1,s_2,...,s_p\}}.$$

Let $\{t_1, t_2, ..., t_{r-p}\} = \{1,2,...,r\} - \{s_1, s_2, ..., s_p\}$. If $\mathbf{h}_{s_\mu}$ is a factor of $\mathbf{h}_{t_\lambda}$, then from Proposition 2.4 and the definitions, we have

$$D_p^{(r)} = D_{\{s_1,s_2,...,s_p\}} = D_{\{s_1,s_2,...,s_p,t_\lambda\}} \leq D_{p+1}^{(r)}.$$

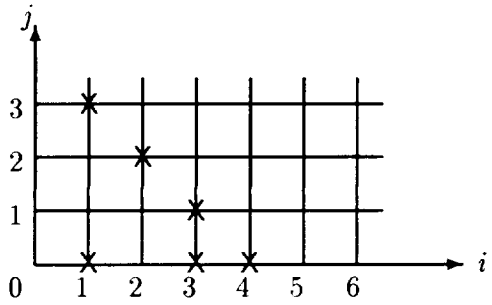
However, from Proposition 2.5, we have $D_p^{(r)} \geq D_{p+1}^{(r)} + 1$. Thus, we have a contradiction. □

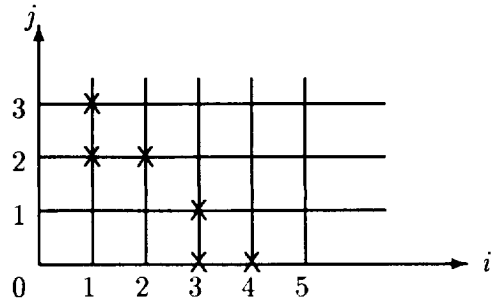
Definition 3.1 *A set* S *of non-negative integer points* (i,j) *(i.e.,* i *and* j *are non-negative integers) is called a regular set if for* $(i,j) \in S$, *we have* $(i',j') \in S$, *for all* $0 \leq i' \leq i$ *and* $0 \leq j' \leq j$.

Using the definition, we have the following result:

Corollary 3.1 *For* $D_{\{k_1,k_2,...,k_p\}}$, *if set* $\{(i,j)|x^iy^j \in \{\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_r\} - \{\mathbf{h}_{k_1}, \mathbf{h}_{k_2}, ..., \mathbf{h}_{k_p}\}\}$ *is not a regular set, then there exists at least one set of* $\{s_1, s_2, ..., s_p\}$ *with* $s_p \leq k_p$, *such that* $D_{\{s_1,s_2,...,s_p\}} \geq 1 + D_{\{k_1,k_2,...,k_p\}}$.

Example 3.3 *Let* $r = 14$ *and* $p = 6$. *The first 14 monomials are* $\{$ *1,* x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3, x^4, x^3y, x^2y^2, xy^3 $\}$. *If* $\{$ k_1, ..., k_6 $\}$ $=$ $\{$ *2, 7, 11, 12, 13, 14* $\}$, *then* $\{1,2,...,r\} - \{k_1,...,k_p\} = \{1,3,4,5,6,8,9,10\}$. *Obseve the set* $\{(i,j)|x^iy^j \in \{\mathbf{h}_1, \mathbf{h}_3, \mathbf{h}_4, \mathbf{h}_5, \mathbf{h}_6, \mathbf{h}_8, \mathbf{h}_9, \mathbf{h}_{10}\}\} = \{(0,0),(0,1),(2,0),(1,1),(0,2),(2,1),(1,2),(0,3)\}$ *does not form a regular set, because* $(2,0)$ *belongs to this set but (1,0) does not. If we choose* $\{s_1,...,s_6\} = \{7,9,11,12,13,14\}$, *then* $\{1,2,...,r\} - \{s_1,...,s_p\} = \{1,2,3,4,5,6,8,10\}$. *The corresponding monomials are* $\{1,x,y,x^2,xy,y^2,x^2y,y^3\}$ *and form a regular set. Obviously,* $D_{\{7,9,11,12,13,14\}} \geq 1 + D_{\{2,7,11,12,13,14\}}$. *The corresponding monomial points are shown in the following figures.*





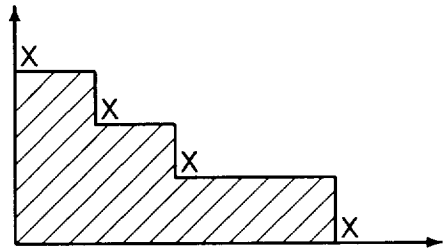
Theorem 3.2 *Let a set of points corresponding to $k_1, k_2, ..., k_p$ be obtained by deleting a regular set, and let their greatest common divisor be 1, i. e., $gcd(\mathbf{h}_{k_1}, \mathbf{h}_{k_2}, ..., \mathbf{h}_{k_p}) = 1$. Then*

$$D_{\{k_1,k_2,...,k_p\}} \leq k_p - p.$$

Proof: Since $gcd(\mathbf{h}_{k_1}, \mathbf{h}_{k_2}, ..., \mathbf{h}_{k_p}) = 1$, these monomials contain the points $(0, t_1), (s_2, t_2), ..., (s_{\lambda-1}, t_{\lambda-1}), (s_\lambda, 0)$ and their multiples, where $\lambda \leq \min\{5, p\}$ and

$$3 \geq t_1 > t_2 > \cdots > t_{\lambda-1} > t_\lambda = 0 \text{ and } 0 = s_1 < s_2 < \cdots < s_{\lambda-1} < s_\lambda \leq 4.$$

Using Theorem 3.1 and the definition of regular set, $D_{\{k_1,k_2,...,k_p\}} \leq$ the number of points contained in the area obtained by these points as sentinels of the regular set (see the figure below), that is denoted by $D_{\{k_1,k_2,...,k_p\}}$.



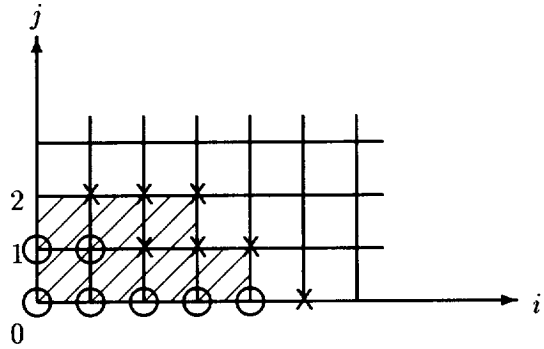
This area contains at most $k_p - p$ points. This completes the proof. □

The following corollary can be easily seen.

Corollary 3.2 $D_p^{(r)} \leq \max\{D_{\{\mathbf{h}_\mu\}} + k_p - p\}$, *where* $\{\mathbf{h}_\mu \mathbf{h}_{k_1}, \cdots, \mathbf{h}_\mu \mathbf{h}_{k_p}\} \subset \{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_r\}$ *and* $k_1 < k_2 < \cdots < k_p$.

We show two examples to illustrate this lemma.

Example 3.4 *Let us consider* $D_{\{[y^2],[x^2y],[xy^2],[y^3],[x^3y],[x^2y^2],[xy^3],[x^5],[x^4y],[x^3y^2],[x^2y^3]\}}, p = 11$ *and* $k_p = 18$. *From Theorem 3.2,* $D_{\{[y^2],[x^2y],[xy^2],[y^3],[x^3y],[x^2y^2],[xy^3],[x^5],[x^4y],[x^3y^2],[x^2y^3]\}} \leq 7 = k_p - p$.



 # of ○ *is 7.*

Example 3.5 *Let us consider* $D_{\{[y^3],[x^3y],[y^4],[x^4y],[x^3y^2],[x^2y^3],[x^6],[x^5y],[x^4y^2]\}}$, *where* $p = 9$ *and* $k_p = 21$. *From Theorem 3.2,* $D_{\{[y^3],[x^3y],[y^4],[x^4y],[x^3y^2],[x^2y^3],[x^6],[x^5y],[x^4y^2]\}} \leq 12 = k_p - p$.

Theorem 3.3 $D_p^{(r)} \leq w(\mathbf{h}_r) - w(\mathbf{h}_p)$.

Proof: Suppose $D_p^{(r)} = D_{\{\mathbf{h}_{s_1}^*, \mathbf{h}_{s_2}^*, ..., \mathbf{h}_{s_p}^*\}}$. Let $gcd(\mathbf{h}_{s_1}, \mathbf{h}_{s_2}, ..., \mathbf{h}_{s_p}) = \mathbf{h}_q$, i.e.,

$$\{\mathbf{h}_{s_1}, \mathbf{h}_{s_2}, ..., \mathbf{h}_{s_p}\} = \mathbf{h}_q \times \{\mathbf{h}_{k_1}, \mathbf{h}_{k_2}, ..., \mathbf{h}_{k_p}\},$$

where

$$gcd(\mathbf{h}_{k_1}, \mathbf{h}_{k_2}, ..., \mathbf{h}_{k_p}) = 1.$$

From Theorem 3.2 and Proposition 2.3, we have

$$D_{\{\mathbf{h}_{s_1}^*, \mathbf{h}_{s_2}^*, ..., \mathbf{h}_{s_p}^*\}} \leq D_{\{\mathbf{h}_q^*\}} + D_{\{\mathbf{h}_{k_1}^*, \mathbf{h}_{k_2}^*, ..., \mathbf{h}_{k_p}^*\}} \leq w(\mathbf{h}_q) + k_p - p.$$

Since $k_p \geq p$, we have $w(\mathbf{h}_p) - p \leq w(\mathbf{h}_{k_p}) - k_p$. On the other hand, $\mathbf{h}_{k_p} \cdot \mathbf{h}_q = \mathbf{h}_{s_p}$ and $s_p \leq r$. This means $w(\mathbf{h}_{k_p}) + w(\mathbf{h}_q) \leq w(\mathbf{h}_r)$. Combining the above two equations, we have

$$w(\mathbf{h}_q) + k_p - p \leq w(\mathbf{h}_q) + w(\mathbf{h}_{k_p}) - w(\mathbf{h}_p) \leq w(\mathbf{h}_r) - w(\mathbf{h}_p).$$

Thus, the proof is completed. □

Corollary 3.3 *If* $\mathbf{h}_r \sim \mathbf{h}_p \cdot \mathbf{h}_\mu$ *for some* $1 \leq \mu < r$, *and* $D_{\{\mathbf{h}_\mu^*\}} = w(\mathbf{h}_\mu)$, *then*

$$D_p^{(r)} = w(\mathbf{h}_\mu).$$

Proof: Since $\mathbf{h}_r \sim \mathbf{h}_p \cdot \mathbf{h}_\mu$,

$$\mathbf{h}_\mu \times \{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_p\} \subseteq \{\mathbf{h}_1, \cdots, \mathbf{h}_p, \cdots, \mathbf{h}_r\}.$$

Thus, by Proposition 2.3, $D_p^{(r)} \geq D_{\{\mathbf{h}_\mu^*\mathbf{h}_1^*, \mathbf{h}_\mu^*\mathbf{h}_2^*, \cdots, \mathbf{h}_\mu^*\mathbf{h}_p^*\}} \geq D_{\{\mathbf{h}_\mu^*\}}$.

On the other hand, $D_{\{\mathbf{h}_\mu^*\}} = w(\mathbf{h}_\mu) = w(\mathbf{h}_r) - w(\mathbf{h}_p)$. From Theorem 3.3, $D_p^{(r)} \leq w(\mathbf{h}_r) - w(\mathbf{h}_p) = D_{\{\mathbf{h}_\mu^*\}}$. Therefore, $D_p^{(r)} = D_{\{\mathbf{h}_\mu^*\}} = w(\mathbf{h}_\mu)$.

Lemma 3.4 *If there is no* $1 \leq \mu < r$ *such that* $\mathbf{h}_r \sim \mathbf{h}_p \cdot \mathbf{h}_\mu$, *and* $r - p \geq w(\mathbf{h}_\nu)\}$, *for any* $1 \leq \nu < r$ *with that* $\mathbf{h}_{r'} \sim \mathbf{h}^p \cdot \mathbf{h}_\nu$ *and* $r' \leq r$, *then*

$$D_p^{(r)} \leq r - p. \tag{3}$$

Proof: If $D_p^{(r)} = D_{\{\mathbf{h}_{k_1}^*, \mathbf{h}_{k_2}^*, ..., \mathbf{h}_{k_p}^*\}}$, where $gcd(\mathbf{h}_{k_1}, \mathbf{h}_{k_2}, ..., \mathbf{h}_{k_p}) = 1$, then from Lemma 3.4, $D_p^{(r)} \leq r - p$. Otherwise, $gcd(\mathbf{h}_{k_1}, \mathbf{h}_{k_2}, ..., \mathbf{h}_{k_p}) = \mathbf{h}_\nu$, $\mathbf{h}_r' \sim \mathbf{h}_p \cdot \mathbf{h}_\nu$, and $r' \leq r$. Since $r - p \geq w(\mathbf{h}_\nu)$ and Theorem 3.3, $D_p^{(r)} \leq w(\mathbf{h}_r') - w(\mathbf{h}_p) \leq r - p$. □

Example 3.6 *For the Hermitian curve (2), let us consider* $D_6^{(16)}$. *Since* $D_{\{\mathbf{h}_\mu\}} = w(\mathbf{h}_\mu)$, *using Theorem 3.3, it can be easily checked that* $D_6^{(16)} \leq r - p \leq 10$.

In the following we show how to find all generalized Hamming weights for an algebraic-geometric code by one example.

Example 3.7 *Consider $r = 15$ and $p = 7$, i.e. $\mathbf{h}_r = 50 = x^5 \sim y^4, \mathbf{h}_p = 30 = x^3$. We have $\mu = 4$, because $\mathbf{h}_4 \cdot \mathbf{h}_7 \sim \mathbf{h}_{15}$, i.e., $w(\mathbf{h}_{15}) = w(\mathbf{h}_4) + w(\mathbf{h}_7)(20 = 12+8)$. On the other hand, for Hermitian curves, $D_{\mathbf{h}_\mu} = w(\mathbf{h}_\mu)$. Using Corollary 3.3, we have $D_7^{(15)} = 8$.*

Example 3.8 *Let us consider a Hermitian code C_{16} over $GF(2^4)$, a parity check matrix of which is:*

$$\mathbf{H}_{16} = [1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3, x^4, x^3y, x^2y^2, xy^3, x^5, x^4y]^T.$$

We have the weight sequence: $W = \{0, 4, 5, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21\}$ and Table 1

μ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$w(\mathbf{h}_\mu)$	0	4	5	8	9	10	12	13	14	15	16	17	18	19	20	21

Table 1:

With Table 1 and Theorem 3.3, we compute

$D_1^{(16)} = w(\mathbf{h}_{16}) = 21$, $D_2^{(16)} = w(\mathbf{h}_{12}) = 17$, $D_3^{(16)} = w(\mathbf{h}_{11}) = 16$, $D_4^{(16)} = w(\mathbf{h}_8) = 13$,
$D_5^{(16)} = w(\mathbf{h}_7) = 12$, $D_7^{(16)} = w(\mathbf{h}_5) = 9$, $D_8^{(16)} = w(\mathbf{h}_4) = 8$, $D_{11}^{(16)} = w(\mathbf{h}_3) = 5$,
$D_{12}^{(16)} = w(\mathbf{h}_2) = 4$, $D_{16}^{(16)} = w(\mathbf{h}_1) = 0$.

Now using the monotonicity proposition, i.e. Proposition 2.5, we obtain

$$D_9^{(16)} = 7, D_{10}^{(16)} = 6, D_{13}^{(16)} = 3, D_{14}^{(16)} = 2, D_{15}^{(16)} = 1.$$

For the last remaining value, $D_6^{(16)}$, if using Proposition 2.5, then we have only have $10 = D_7^{(16)} + 1 \leq D_6^{(16)} \leq D_5^{(16)} - 1 = 11$. Using Corollary 3.2, we have $D_6^{(16)} = 10$ (see Example 3.6). Thus, we have the following values.

$D_1^{(16)} = 21$, $D_2^{(16)} = 17$, $D_3^{(16)} = 16$, $D_4^{(16)} = 13$, $D_5^{(16)} = 12$, $D_6^{(16)} = 10$, $D_7^{(16)} = 9$, $D_8^{(16)} = 8$, $D_9^{(16)} = 7$, $D_{10}^{(16)} = 6$, $D_{11}^{(16)} = 5$, $D_{12}^{(16)} = 4$, $D_{13}^{(16)} = 3$, $D_{14}^{(16)} = 2$, $D_{15}^{(16)} = 1$, $D_{16}^{(16)} = 0$.

These terms appear in the rightmost column of Table 2.

This table allows us to compute the generalized Hamming weights, $d_i(C_{16})$, as follows. From the table, for each column $h = i$ ($i = 1, 2, 3, 4, \ldots$), we consider the first entry that is greater than the entry at the same row and the last column. According to Theorem 1.1, this entry plus 1 gives a lower bound of $d_i(C_{16})$. In the above table, these entries are marked by an '*'. For example, in the column for $h = 1$, $16(< 21), 15(< 17), 14(< 16), 13(= 13), 12(= 12), 11(> 10), 10(> 9), 9(> 8), \ldots, 1(> 0)$. Therefore, 11 is the first entry that is greater than 10 (which is at the same row and in the last column). Hence, $d_1(C_{16}) \geq 12$. Similarly we have $d_2(C_{16}) \geq 15, d_3(C_{16}) \geq 16, d_4(C_{16}) \geq 19, d_5(C_{16}) \geq 20, d_6(C_{16}) \geq 21, d_7(C_{16}) \geq 23$, and $d_h(C_{16}) \geq h + 16$, for $h = 8, 9, 10, 11, \ldots, 48$.

For the Hermitian code, for each x, since there are exactly 4 distinct rational points (x, y), and for each $y \neq 0$, since there are exactly 5 distinct rational points (x, y), the generalized Hamming weight lower bounds derived by the above approach are identical to their generalized Hamming weights.

	d - 1 = 16 - p + h												
p	h=1	h=2	h=3	h=4	h=5	h=6	h=7	h=8	h=9	h=10	h=11	h=12	$D_p^{(16)}$
1	16	17	18	19	20	21	22*	23*	24*	25*	26*	27*	21
2	15	16	17	18*	19*	20*	21	22	23	24	25	26	17
3	14	15	16	17	18	19	20	21	22	23	24	25	16
4	13	14*	15*	16	17	18	19	20	21	22	23	24	13
5	12	13	14	15	16	17	18	19	20	21	22	23	12
6	11*	12	13	14	15	16	17	18	19	20	21	22	10
7	10	11	12	13	14	15	16	17	18	19	20	21	9
8	9	10	11	12	13	14	15	16	17	18	19	20	8
9	8	9	10	11	12	13	14	15	16	17	18	19	7
10	7	8	9	10	11	12	13	14	15	16	17	18	6
11	6	7	8	9	10	11	12	13	14	15	16	17	5
12	5	6	7	8	9	10	11	12	13	14	15	16	4
13	4	5	6	7	8	9	10	11	12	13	14	15	3
14	3	4	5	6	7	8	9	10	11	12	13	14	2
15	2	3	4	5	6	7	8	9	10	11	12	13	1
16	1	2	3	4	5	6	7	8	9	10	11	12	0

Table 2:

4 Construction of Efficient Linear Codes with Small Minimum Distance

In this section, three examples illustrate how the generalized Bezout theorem can be used to construct more efficient linear codes including the algebraic geometric codes from plane curves.

4.1 Efficient Linear Codes with Minimum Distance 4

Let us consider $D_2^{(4)}$ over $GF(2^m)$, where the first four polynomials are $\{1, x, y, x^2+\beta xy+y^2\}$, where β is any element with $\mathrm{tr}(\beta^{-1}) \equiv \sum_{i=0}^{m-1} \beta^{-2^i} = 1$ over $GF(2^m)$.

In the following, we will prove an important result:

Theorem 4.1 $D_2^{(4)} \leq 2$.

Proof: Since the four polynomials are $1, x, y, x^2+\beta xy+y^2$, we have the following selections: $\{1,x\},\{1,y\},\{1,x^2+\beta xy+y^2\};\{x,y\},\{x,x^2+\beta xy+y^2\}$, and $\{y,x^2+\beta xy+y^2\}$.

From $D_{\{[1]\}} = 0$ and Lemma 2.3, $D_{\{[1],[x]\}} = D_{\{[1],[y]\}} = D_{\{[1],[x^2+\beta xy+y^2]\}} = 0$. Also, from Theorem 3.4, $D_{\{[x],[y]\}} = 1$. Now we calculate the fifth selection, i.e. $D_{\{[x],[x^2+\beta xy+y^2]\}}$. For any constants b, c, d, e, we calculate the number of distinct points of intersection of the following curves:

$$\begin{cases} x+b=0, \\ x^2+\beta xy+y^2+cy+dx+e=0. \end{cases}$$

We have

$$R(y) = \begin{vmatrix} 1 & \beta y + d & y^2 + cy + e \\ 1 & b & 0 \\ 0 & 1 & b \end{vmatrix},$$

i.e., $R(y) = y^2 + cy + e + \cdots$. Thus, $deg R(y) = 2$. From Theorem 2.3, $D_{\{[x],[x^2+\beta xy+y^2]\}} \leq 2$.

Now let us consider the last selection. For any constants a, b, c, d, e, we calculate the number of distinct points of intersection of the following curves:

$$\begin{cases} y + ax + b = 0, \\ x^2 + \beta xy + y^2 + cy + dx + e = 0. \end{cases}$$

We have

$$R(x) = \begin{vmatrix} 1 & \beta x + c & x^2 + dx + e \\ 1 & ax + b & 0 \\ 0 & 1 & ax + b \end{vmatrix},$$

i.e., $R(x) = a^2x^2 + b^2 + x^2 + dx + e - \beta ax^2 - acx - \beta bx - bc = (a^2 + \beta a + 1)x^2 + (d + \beta b + ac)x + bc + b^2$. Since $\mathrm{tr}(\beta^{-1}) = 1, a^2 + \beta a + 1 \neq 0$ for all $a \in GF(2^m)$. Thus, deg $R(x) = 2$. Therefore, $D_{\{[y],[x^2+\beta xy+y^2]\}} \leq 2$.

Combining the above results, we have $D_2^{(4)} \leq 2$. □

Remark (5): For the maximal number of distinct points of intersection, the following forms are equivalent:

$$\begin{cases} y + ax + b = 0, \\ x^2 + \beta xy + y^2 + cy + dx + e = 0. \end{cases}$$

and

$$\begin{cases} y + ax + b = 0, \\ x^2 + \beta xy + y^2 + dx + e = 0. \end{cases}$$

Hereinafter, we use the second form for simplification.

Remark (6): Observe that $D_{\{\mathbf{h}^*_{\lambda_1},\ldots,\mathbf{h}^*_{\lambda_p}\}} \leq l$, if and only if $-\infty <$ deg of $R(x) \leq l$. Further, if the degree of $R(x)$ is equal to $-\infty$, i.e., $R(x)$ is a zero, then $R(x) = 0$ has all x for solutions.

Theorem 4.2 *Let C be a linear code with length $n = 2^{2m}$ over $GF(2^m)$, which is defined by a parity check matrix $\mathbf{H} \equiv [1, x, y, x^2 + \beta xy + y^2]^T$. Then the minimum distance of C is at least 4.*

Proof: From Theorem 4.1, $D_2^{(4)} = D_{\{[y],[x^2+\beta xy+y^2]\}} \leq 2 < 4 - 1$. From Corollary 1.1, the minimum distance of C is at least 4. □

Example 4.1 *Let $m = 2$, and let α be a primitive element of $GF(2^2)$, $\beta = \alpha^{-1}$. It can be easily checked that $\alpha + \alpha^2 = 1 \neq 0$. From Theorem 4.2, the following matrix defines a linear code C with length 16 and minimum distance at least 4 over $GF(2^2)$:*

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & \alpha & \alpha & \alpha & \alpha & \alpha^2 & \alpha^2 & \alpha^2 & \alpha^2 \\ 0 & 1 & \alpha & \alpha^2 & 0 & 1 & \alpha & \alpha^2 & 0 & 1 & \alpha & \alpha^2 & 0 & 1 & \alpha & \alpha^2 \\ 0 & 1 & \alpha^2 & \alpha & 1 & \alpha^2 & \alpha^2 & 1 & \alpha^2 & \alpha^2 & \alpha & \alpha & \alpha & 1 & \alpha & 1 \end{bmatrix}.$$

Construction 4.1: Let $n = 2^{2km}$, The following parity check matrix $\mathbf{H}$ defines a linear code C with length $n = 2^{2km}$ and minimum distance at least 4 over $GF(2^m)$, i.e., C is a $(2^{2km}, 2^{2km} - 1 - 3k, \geq 4)$ code:

$$\mathbf{H} = [1, x_1, y_1, x_1^2 + \beta x_1 y_1 + y_1^2, \cdots, x_k, y_k, x_k^2 + \beta x_k y_k + y_k^2]^T,$$

where $\mathrm{tr}(\beta^{-1}) = 1$.

These codes are better than the Kaneda codes [21], and same as the Chen codes [21]. However, the derivations are different. The codes obtained by Construction 4.1 can be more easily decoded than the Chen codes.

4.2 Efficient Linear Codes with Minimum Distance ≥ 5

Let us consider $D_5^{(7)}$ over $GF(2^m)$, where the polynomials are $\{1, x, y, x^2, xy, y^2, x^3 + \gamma x^2 y + \beta xy^2 + y^3\}$, where $\gamma, \beta \in GF(2^m)$ and $x^3 + \gamma x^2 y + \beta xy^2 + y^3$ is irreducible. We have the following theorem.

Theorem 4.3 $D_4^{(7)} \leq 3$.

Proof: Let $D_4^{(7)} = D_{\{\lambda_1,\lambda_2,\lambda_3,\lambda_4\}}$. If $\lambda_1 = 1$, then from Proposition 2, $D_{\{\lambda_1,\lambda_2,\lambda_3,\lambda_4\}} \leq D_{\{1\}} = 0$. If $\lambda_1 = 2$ or 3, then it can be easily checked that $D_{\{\lambda_1,\lambda_2,\lambda_3,\lambda_4\}} \leq 3$. Now we need only to consider the case of $\lambda_1 = 4$, i.e., $D_{\{[x^2],[xy],[y^2],[x^3+\gamma x^2 y+\beta xy^2+y^3]\}}$. Let us consider the following curves:

$$\begin{cases} x^2 + ay + bx + c = 0 \\ xy + dy + ex + f = 0 \\ y^2 + gy + hx + i = 0 \\ x^3 + \gamma x^2 y + \beta xy^2 + y^3 + jy + kx + l = 0 \end{cases},$$

where $a, b, ..., l$ are any constants of $GF(2^m)$. We have the following partial x-resultant matrix:

$$\begin{bmatrix} 1 & \gamma y & \beta y^2 + k & y^3 + jy + l & 0 \\ 0 & 1 & \gamma y & \beta y^2 + k & y^3 + jy + l \\ 0 & 0 & 1 & b & ay + c \\ 0 & 0 & 0 & y + e & dy + f \\ 0 & 0 & 0 & h & y^2 + gy + i \end{bmatrix}.$$

$PR(y) = (y + e)(y^2 + gy + i) - h(dy + f)$. $degPR(y) = 3$. Combining the above results, the proof is completed. □

From Corollary 1.1 and Theorem 4.3, we have the following theorem.

Theorem 4.4 *Let C be a linear code with length $n = 2^{2m}$ over $GF(2^m)$, which is defined by a parity check matrix $\mathbf{H} \equiv [1, x, y, x^2, xy, y^2, x^3 + \gamma x^2 y + \beta xy^2 + y^3]^T$, where $x^3 + \gamma x^2 y + \beta xy^2 + y^3$ is a irreducible polynomial over $GF(2^m)$. Then the minimum distance of C is at least 5.*

Example 4.2 *Let $m = 2$, and let α be a primitive element of $GF(2^2)$. It can checked that $x^3 + \alpha x^2 y + y^3$ is an irreducible polynomial. From Theorem 3.3.2, the following matrix defines a linear code $C(16, 9, \geq 5)$ over $GF(2^2)$:*

$$\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & \alpha & \alpha & \alpha & \alpha & \alpha^2 & \alpha^2 & \alpha^2 & \alpha^2 \\
0 & 1 & \alpha & \alpha^2 & 0 & 1 & \alpha & \alpha^2 & 0 & 1 & \alpha & \alpha^2 & 0 & 1 & \alpha & \alpha^2 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & \alpha^2 & \alpha^2 & \alpha^2 & \alpha^2 & \alpha & \alpha & \alpha & \alpha \\
0 & 0 & 0 & 0 & 0 & 1 & \alpha & \alpha^2 & 0 & \alpha & \alpha^2 & 1 & 0 & \alpha^2 & 1 & \alpha \\
0 & 1 & \alpha^2 & \alpha & 0 & 1 & \alpha^2 & \alpha & 0 & 1 & \alpha^2 & \alpha & 0 & 1 & \alpha^2 & \alpha \\
0 & 1 & 1 & 1 & 1 & \alpha & \alpha^2 & 1 & 1 & 1 & \alpha & \alpha^2 & 1 & \alpha^2 & 1 & \alpha
\end{bmatrix}.$$

Construction 4.2: Let $n = 2^{2km}$, The following parity check matrix $\mathbf{H}$ defines a linear code C with length $n = 2^{2km}$ and minimum distance at least 5 over $GF(2^m)$, i.e., C is a $(2^{2km}, 2^{2km} - 1 - 6k, \geq 5)$ code:

$$\mathbf{H} = [1, \cdots, x_i, y_i, x_i^2, x_i y_i, y_i^2, x_i^3 + \gamma x_i^2 y_i + \beta x_i y_i^2 + y_i^3, \cdots]^T,$$

where $x_i^3 + \gamma x_i^2 y_i + \beta x_i y_i^2 + y_i^3$ are irreducible over $GF(2^m)$ for $i = 1, 2, ..., k$.
Remark (7): The result can be generalized to the cases of any minimum distance $d \geq 6$. The code dimension can be increased by at least one, while keeping the same code length $n = 2^{2m}$ and the minimum distance $d \geq 6$ over $GF(2^m)$ [18, 22].

4.3 Improved Klein Codes

Let us consider the Klein curve over $GF(2^3)$:

$$x^3 y + y^3 + x = 0. \tag{4}$$

There are 22 points on the curve. Let LS be the set of all points on the curve. The LS has the following points:

$x =$	$y =$	$x =$	$y =$	$x =$	$y =$
0	0;				
1	$\alpha, \alpha^2, \alpha^4$;				
α	$1, \alpha^2, \alpha^6$;	α^2	$1, \alpha^4, \alpha^5$;	α^4	$1, \alpha, \alpha^3$;
α^3	$\alpha^2, \alpha^3, \alpha^5$;	α^6	$\alpha^4, \alpha^6, \alpha^3$;	α^5	$\alpha, \alpha^5, \alpha^6$;

where α is a primitive element of $GF(2^3)$.

Let us consider a linear code C defined by the parity check matrix

$$\mathbf{H} \equiv [1, x, y, x^2, xy, x^3, y^2]^T.$$

Theorem 4.5 *The minimum distance of the code C is at least 6, i.e., C is a linear code $(22, 15, \geq 6)$ over $GF(2^3)$.*

Proof: From Corollary 1.1, it is sufficient to prove $D_3^{(7)} \leq 4$. From the definition,

$$D_3^{(7)} = \max\{D_{\{[1],*,*\}}, D_{\{[x],*,*\}}, D_{\{[y],*,*\}}, D_{\{[x^2],[xy],[y^2]\}}, D_{\{[x^2],[x^3],[y^2]\}}, D_{\{[x^2],[xy],[x^3]\}}, D_{\{[xy],[x^3],[y^2]\}}\}.$$

It is easily seen that the first number is equal to zero, the second and third numbers are at most 3, and the 4-th number is also at most 3. In the following we prove that the 7-th number is at most 4. The proof of that the 5-th number and 6-th number are at most 4 is similar to the case of the 7-th number.

From the definition, $D_{\{[x^2],[xy],[x^3]\}}$ expresses the maximal number of distinct points of the intersection of the following four curves:

$$\begin{cases} x^3y + y^3 + x = 0 \\ x^2 + iy + jx + k = 0 \\ xy + ax^2 + by + cx + d = 0 \\ x^3 + fy + gx + h = 0 \end{cases}, \tag{5}$$

where $i, j, k, a, b, c, d, f, g, h$ are any constants over $GF(2^3)$. From these equations, we have the following $PR(y)$:

$$\begin{vmatrix} 1 & 0 & g & fy+h & 0 & 0 \\ 0 & 1 & 0 & g & fy+h & 0 \\ 0 & 0 & 1 & 0 & g & fy+h \\ 0 & 0 & y & 0 & 1 & y^3 \\ 0 & 0 & 0 & 1 & j & iy+k \\ 0 & 0 & 0 & a & y+c & by+d \end{vmatrix}.$$

It is equal to

$$\begin{vmatrix} 1 & 0 & g & fy+h & 0 & 0 \\ 0 & 1 & 0 & g & fy+h & 0 \\ 0 & 0 & 1 & 0 & g & fy+h \\ 0 & 0 & y & 0 & 1 & y^3 \\ 0 & 0 & 0 & 1 & j & iy+k \\ 0 & 0 & 0 & 0 & y+c' & b'y+d' \end{vmatrix}.$$

It is equal to

$$\begin{vmatrix} 1 & g & fy+h \\ y & 1 & y^3 \\ 0 & y+c' & b'y+d' \end{vmatrix}.$$

It equal to $y^4 + \cdots$. Therefore, the degree is 4. Hence, $D_{\{[x^2],[xy],[x^3]\}} \leq 4$. Combining the other numbers, we have $D_3^{(7)} \leq 4$. Thus, C is a $(22, 15, \geq 6)$ linear code over $GF(2^3)$. □

Using the result in [17,18], we can only obtain a linear code $(22, 14, \geq 6)$ over $GF(2^3)$. For the current Klein code, using the Riemann-Roch theorem, we also obtain a $(22, 14, 6)$ Klein code.

Let us consider the another linear code C^* defined by the parity check matrix

$$\mathbf{H} \equiv [1, x, y, x^2, xy, x^3 + y^2]^T.$$

Theorem 4.6 *The minimum distance of the code C is at least 5, i.e., C is a linear code $(22, 16, \geq 5)$ over $GF(2^3)$.*

Proof: From Corollary 1.1, it is sufficient to prove $D_3^{(6)} \leq 3$. From the definition,

$$D_3^{(6)} = \max\{D_{\{[1],*,*\}}, D_{\{[x],*,*\}}, D_{\{[y],*,*\}}, D_{\{[x^2],[xy],[x^3+y^2]\}}\}.$$

It is easily seen that the first number is equal to zero, the second and third number is at most 3. In the following we prove that the last number is at most 3.

From the definition, $D_{\{[x^2],[xy],[x^3+y^2]\}}$ expresses the maximal number of distinct points of the intersection of the following four curves:

$$\begin{cases} x^3y + y^3 + x = 0 \\ x^3 + y^2 + ax + by + c = 0 \\ x^2 + dx + ey + f = 0 \\ xy + gx + hy + i = 0 \end{cases}, \tag{6}$$

where $a, b, c, d, e, f, g, h, i$ are any constants over $GF(2^3)$. From these equations, we have the following $PR(y)$:

$$\begin{vmatrix} 1 & 0 & a & y^2+by+c \\ 1 & d & ey+f & 0 \\ 0 & 1 & d & ey+f \\ 0 & 0 & y+g & hy+i \end{vmatrix}.$$

It is equal to

$$\begin{vmatrix} 1 & 0 & a & y^2+by+c \\ 0 & d & ey+f* & y^2+by+c \\ 0 & 1 & d & ey+f \\ 0 & 0 & y+g & hy+i \end{vmatrix}.$$

It is equal to

$$\begin{vmatrix} 0 & ey+f' & y^2+b'y+c' \\ 1 & d & ey+f \\ 0 & y+g & hy+i \end{vmatrix} = \begin{vmatrix} ey+f' & y^2+b'y+c' \\ y+g & hy+i \end{vmatrix}.$$

The determinant is equal to $y^3 + \cdots$. Thus, $D_3^{(6)} \leq 3$. □

4.4 Improved Hermitian Codes

Let us consider the Hermitian curve over $GF(2^4)$:

$$x^5 + y^4 + y = 0.$$

Let us consider the Hermitian code over $GF(2^4)$ defined by the following parity check matrix:

$$\mathbf{H} \equiv [1, x, y, x^2, xy, y^2, x^3, y^3 + x^4]^T.$$

We have the following theorem.

Theorem 4.7 *The minimum distance of the defined above Hermitian codes is at least 6, i.e., the Hermitain code is a linear code* $(64, 56, \geq 6)$ *over* $GF(2^4)$.

Proof: From Corollary 1.1, it is sufficient to prove $D_4^{(8)} \leq 4$. From the definition,

$$D_4^{(8)} = \max\{D_{\{[1],*,*,*\}}, D_{\{[x],*,*,*\}}, D_{\{[y],*,*,*\}}, D_{\{[x^2],[y^2],*,*\}}, D_{\{[xy],[y^2],[x^3],[y^3+x^4]\}}, D_{\{[x^2],[xy],[x^3],[y^3+x^4]\}}\}.$$

It is easily seen that the first number is equal to zero, the second and third number is at most 3. The 4th number is at most 4. In the following we prove that the last two numbers are at most 4.

From Proposition 2.3, $D_{\{[xy],[y^2],[x^3],[y^3+x^4]\}} \leq D_{\{[xy],[y^2],[x^3]\}}$. Using Theorem 3.1, the right side is at most 4. Thus, the 5-th number is also at most 4.

Now let us consider the 6-th number. From Proposition 2.4,

$$D_{\{[x^2],[xy],[x^3],[y^3+x^4]\}} = D_{\{[x^2],[xy],[y^3+x^4]\}}.$$

From the definition, $D_{\{[x^2],[xy],[y^3+x^4]\}}$ expresses the maximal number of distinct points of the intersection of the following four curves:

$$\begin{cases} x^5 + y^4 + y = 0 \\ x^4 + y^3 + ay^2 + by + cx + d = 0 \\ x^2 + a'y^2 + b'y + c'x + d' = 0 \\ xy + a''y^2 + b''y + c''x + d'' = 0 \end{cases}, \tag{7}$$

where $a, b, c, d, a', b', c', d', a'', b'', c'', d''$ are any constants over $GF(2^4)$. From these equations, we have the following $PR(y)$:

$$\begin{vmatrix} 1 & 0 & 0 & c & y^3 + ay^2 + by + d \\ 1 & c' & a'y^2 + b'y + d' & 0 & 0 \\ 0 & 1 & c' & a'y^2 + b'y + d' & 0 \\ 0 & 0 & 1 & c' & a'y^2 + b'y + d' \\ 0 & 0 & 0 & y + c'' & a''y^2 + b''y + d'' \end{vmatrix}.$$

It is equal to

$$\begin{vmatrix} 1 & 0 & 0 & c & y^3 + ay^2 + by + d \\ 0 & c' & a'y^2 + b'y + d' & c & y^3 + ay^2 + by + d \\ 0 & 1 & c' & a'y^2 + b'y + d' & 0 \\ 0 & 0 & 1 & c' & a'y^2 + b'y + d' \\ 0 & 0 & 0 & y + c'' & a''y^2 + b''y + d'' \end{vmatrix}.$$

It can be easily seen that when $a' = 0$ and $a'' = 0$, the determinant is equal to $y^4 + \cdots$. The degree is equal to 4. When $a' \neq 0$, the 6-th number reduces to the 5-th number, i.e. $D_{\{[x^2],[xy],[x^3],[y^3+x^4]\}}$ reduces to $D_{\{[xy],[y^2],[x^3],[y^3+x^4]\}}$. When $a'' \neq 0$, the the 6-th number reduces to the 4-th number. Thus, the 6-th number is also at most 4. Therefore, the 6-th number is at most 4. □

Using Riemann-Roch Theorem, the current AG code with $d \geq 6$, r should be $d+g-1 = 11$. That means the linear code $(64, 53, \geq 6)$ defined by $\mathbf{H}' = [1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3, x^4]^T$ has the minimum distance $d \geq 6$. Using the construction in [18], an improved Hermitian code defined by $\mathbf{H}' = [1, x, y, x^2, xy, y^2, x^3, y^3, x^4]^T$ is a linear code $(64, 55, \geq 6)$. The new code is more efficient.

5 Conclusions

It is well known that Bezout's theorem can be used to determine an upper bound of the number of common points of two plane curves. In this paper, we first reduce the determination of a lower bound of a linear code's minimum distance and generalized Hamming weights to the determination of the number of distinct points of the intersection of several curves. Then, we present a generalized Bezout theorem and a new approach to determine the minimum distance and the generalized Hamming weights. We also present a new construction of linear codes with any length n over GF(2^b) with minimum distance 4 and ≥ 5. The codes with $d = 4$ have been used in computer memory systems. The new codes have applications not only in computer memory systems but also in distributed systems [24, 25], CD audio, Video disk, and CD ROM.

In this paper, we discuss only the case in two-dimensional affine spaces. Our results should generalize to high-dimensional affine spaces.

References

[1] R. E. Blahut, *Theory and Practice of Error Control Codes*, Reading, Ma: Addison-Wesley, 1985.

[2] V. K. Wei, "Generalized Hamming weights for linear codes", *IEEE Trans. on Information Theory*, Vol. IT-37, pp. 1412-1428, Sept., 1991.

[3] T. Kasami, T. Tanaka, T. Fujiwara, and S. Lin, "On the optimum bit orders with respect to the state complexity of trellis diagrams for binary codes", *IEEE Trans. on Information Theory*, Vol. IT-39, pp. 242-245, Jan., 1993.

[4] T. Kasami, T. Tanaka, T. Fujiwara, and S. Lin, "On complexity of trellis structure of linear block codes", *IEEE Trans. on Information Theory*, Vol. IT-39, pp. 1057-1064, May 1993.

[5] T. Helleseth and P. V. Kumar, "The weight hierarchy of the Kasami codes," *Discrete Math.* to appear.

[6] G. D. Forney, Jr., "Density/Length Profiles and Trellis Complexity of Linear Block Codes and Lattices," *IEEE Trans. on Information Theory*, Vol. IT-40, pp. 1741-1752, Nov., 1994.

[7] G. L. Feng, K. K. Tzeng, and V. K. Wei, "On the generalized Hamming weights of several classes of cyclic codes," *IEEE Trans. on Information Theory*, Vol. IT-38, pp.1125-1130, May, 1992.

[8] T. Helleseth, T. Klove, and O. Ytrehus, "Generalized Hamming Weights of Linear Codes," *IEEE Trans. on Information Theory*, vol. 38, pp. 1133-1140, May 1992.

[9] T. Klove, "Support weight distribution of linear codes," *IEEE Trans. on Information Theory*, Vol. IT-39, pp. 648-654, May 1993.

[10] V. K. Wei and K. Yang, "On the generalized Hamming weights of product codes," *IEEE Trans. on Information Theory*, Vol. IT-39, pp. 1709-1713, Sept. 1993.

[11] J. W. P. Hirschfeld, M. A. Tsfasman, and S. G. Vladut, "The Weight Hierarchy of Higher Dimensional Hermitian Codes," *IEEE Trans. on Information Theory*, vol. 40, pp. 275-278, January 1994.

[12] K. Yang, P. V. Kumar, and H. Stichtenoth, "On the Weight Hierarchy of Geometric Goppa Codes," *IEEE Trans. on Information Theory*, Vol. IT-40, pp. 913-920, May 1994.

[13] H. Stichtenoth and C. Voβ , "Generalized Hamming weights of trace codes," manuscript, 1993.

[14] T. Helleseth, T. Klove, V. I. Levenshtein, and O. Ytrehus, "Bounds on the Minimum Support Weights," *IEEE Trans. on Information Theory*, Vol. IT-41, pp. 432-440, March 1995.

[15] S. S. Abhyankar, *Algebraic Geometry for Scientists and Engineers*, The American Mathematical Society, 1990.

[16] J.J. Sylvester, "On a general method of determining by mere inspection the derivations from two equations of any degree", *Philosophical Magazine*, 16(1840), 132-135.

[17] G. L. Feng and T. R. N. Rao, "A Simple Approach for Construction of Algebraic Geometric Codes from Affine Plane Curves," *IEEE Trans. on Information Theory*, Vol. IT-40, No.4, pp. 1003-1012, July 1994.

[18] G. L. Feng and T. R. N. Rao, "Improved Geometric Goppa Codes, Part I: Basic Theory" to appear in *IEEE Trans. on Information Theory.*

[19] S. Miura and N. Kamiya, "Geometric-Goppa Codes on Some Maximal Curves and Their Minimum Distance," *Proceedings of 1993 IEEE Information Theory Workshop*, June 4-8, 1993 at Shizuoka, Japan, pp. 85-86.

[20] S. Kaneda and E. Fujiwara, "Single byte error correcting-double byte error detecting codes for memory systems," *IEEE Trans. on Computers*, Vol. C-31, pp.596-602, July, 1982.

[21] C. L. Chen, "Byte-oriented error-correcting codes for semiconductor memory systems," *IEEE Trans. on Computers*, Vol. C-35, pp.646-648, July, 1986.

[22] C. Heegard and K. Saints, "Cascaded Reed-Solomon codes and Grobner Bases," *Proceedings of the 1993 IEEE Workshop on Information Theory*, pp.5.2.1-5.2.2, June 4-8, 1993, Japan.

[23] R. E. Blahut,"On Codes Containing Hermitian Codes", *Proceedings of the 1995 IEEE International Symposium on Information Theory.*

[24] M. O. Rabin, "Efficient dispersal of information for security, load balancing and fault-tolerance," Harvard University, Cambridge, MA. TR-02-87, Apr. 1987.

[25] G. Agrawal and P. Jalote, "Coding-based replication schemes for distributed systems," *IEEE Trans. on Parallel and Distributed Systems*, Vol. 6, pp. 240-251, Mar. 1995.