Construction and decoding of a class of algebraic geometry codes

Justesen, Jørn; Larsen, Knud J.; Jensen, Helge Elbrønd; Havemose, Allan; Høholdt, Tom

Published in:
I E E E Transactions on Information Theory

Link to article, DOI:
10.1109/18.32157

Publication date:
1989

Document Version
Publisher's PDF, also known as Version of record

Construction and Decoding of a Class of Algebraic Geometry Codes

JØRN JUSTESEN, KNUD J. LARSEN, H. ELBRØND JENSEN,
ALLAN HAVEMOSE, AND TOM HØHOLDT

Abstract—**We construct a class of codes derived from algebraic plane curves. The concepts and results from algebraic geometry we use are explained in detail, and no further knowledge of algebraic geometry is needed. Parameters, generator and parity-check matrices are given. The main result is a decoding algorithm which turns out to be a generalization of the Peterson algorithm for decoding BCH codes.**

I. Introduction

In 1977 Goppa wrote a seminal paper [1] describing the connection between coding theory and algebraic geometry. This connection was further developed by Goppa in [2] and [3] and has led to remarkable results. In particular, the paper by Tsfasman *et al.* [4] contains a presentation of a sequence of codes over GF(q) (with $q = p^{2r}$), which exceeds the Gilbert–Varshamov bound for $q \geq 49$.

Since then, a number of papers dealing with algebraic geometry codes have appeared [5]–[9]. Most of these papers require extensive knowledge of algebraic geometry. One of the motivations for the present paper was to use the ideas of Goppa without the heavy machinery of algebraic geometry. The aim has been to construct codes, based on algebraic curves, in a rather elementary way and, further, to find possible simple decoding procedures for these codes.

The code construction uses only polynomials and points of a plane curve; many good codes are constructed in this way. Thus the construction is elementary, but we still need results from algebraic geometry to carry through some of the proofs (in Section II).

Moreover, for these codes it is possible to derive a simple decoding algorithm which conceptually involves only the solution of a system of linear equations. The algorithm contains as a special case the Peterson algorithm for decoding of Reed–Solomon codes.

The decoding algorithm is the main result of the paper. The idea behind it was recently generalized by Skorobogatov and Vladut [10] to cover codes from arbitrary algebraic curves.

Manuscript received April 19, 1988; revised September 8, 1988.

J. Justesen and K. J. Larsen are with the Institute of Circuit Theory and Telecommunication, the Technical University of Denmark, Building 343, DK-2800 Lyngby, Denmark.

H. E. Jensen, A. Havemose, and T. Høholdt are with the Mathematical Institute, the Technical University of Denmark, Building 303, DK-2800 Lyngby, Denmark.

IEEE Log Number 8929040.

The paper is organized as follows. Section II describes the code construction, giving the parameters and generator and parity-check matrices. Section III contains a number of examples of curves yielding good codes, while in Section IV we give the decoding algorithm and prove that it works.

In Section V we apply the decoding method to some of the codes from Section III. Finally, Section VI contains a discussion of the method.

II. The Code Constructions

Let $F = \mathrm{GF}(q)$ be the finite field with q elements, and let F_1 be the algebraic closure of F. A projective point is an equivalence class of nonzero 3-tuples (x, y, z) over F_1 under the relation $(x_1, y_1, z_1) \equiv (x_2, y_2, z_2)$ if and only if (iff) there exists a $\lambda \in F_1$ such that $\lambda(x_1, y_1, z_1) = (x_2, y_2, z_2)$. A projective point will be denoted by $\langle (x, y, z) \rangle$. The set of projective points is called the projective plane over F_1 and is denoted P^2. A *plane curve* C, or $C(x, y, z)$, in P^2 is a homogeneous polynomial in three variables with coefficients in F. A point of C is a projective point $\langle (x, y, z) \rangle \in P^2$, such that $C(x, y, z) = 0$. Since in the following all curves considered are plane curves, we will refer to them simply as *curves*.

The curve C is *irreducible* (resp. *absolutely irreducible*), if $C(x, y, z)$ cannot be factored in two nontrivial homogeneous polynomials with coefficients from F (resp. coefficients from an algebraic extension of F). A point $\langle (x, y, z) \rangle$ on the curve C is called *singular* if all the partial derivatives C'_x, C'_y, C'_z are zero at this point. The curve C is called *regular* if all points on the curve are nonsingular. It can be seen that a regular curve is absolutely irreducible. A *rational point* Q on the curve C is a point, which has a representant (x_1, y_1, z_1), for which $C(x_1, y_1, z_1) = 0$ and for which all three coordinates belong to F.

Let us now consider an *irreducible* curve C of *degree* m with rational points $P_1, P_2, \cdots, P_n$. For each of these points we choose a specific "representation," e.g. the one for which the first nonzero coordinate is 1. These points in F^3 are denoted by $P'_1, P'_2, \cdots, P'_n$. For any number $j < q$, let V_j denote the vector space over F consisting of the zero polynomial and all homogeneous polynomials with degree j and with coefficients in F.

0018-9448/89/0700-0811$01.00

We now define two linear codes $G_c(j)$ and $H_c(j)$ over GF(q) as follows:

$$G_c(j) = \{(f(P_1'), \cdots, f(P_n')) | f \in V_j\} \quad (2.1)$$

$$H_c(j) = G_c(j)^{\perp}. \quad (2.2)$$

Our first task is to find—or estimate—the parameters of these codes, that is, their length, dimension, and minimum distance. We shall do this in three steps where each step emphasizes the assumptions and the results from algebraic geometry needed to obtain the stated results. We first need the theorem of Bezout, which we quote from [11, p. 112.]

Bezout's Theorem: Let F and G be projective plane curves of degrees m and n, respectively. Assume that F and G have no common component. The number of points in $F \cap G$, counted with multiplicities, is mn.

Theorem 1: Let k be the dimension and d the minimum distance of the code $G_c(j)$. We have

$$d \geq n - mj. \quad (2.3)$$

Suppose that $n > mj$. If $j < m$, then

$$k = \binom{j+2}{2} \quad (2.4)$$

and if $j \geq m$, then

$$k = \binom{j+2}{2} - \binom{j-m+2}{2}. \quad (2.5)$$

Proof: Let $f \in V_j$. If C is a factor of f, then the codeword corresponding to f is $\mathbf{0}$. Otherwise, since C is irreducible, the Bezout theorem states that f and C have at most mj points in common. This gives (2.3).

The dimension of the vector space V_j is the number of different terms $x^{j_1}y^{j_2}z^{j_3}$, where $j_1 + j_2 + j_3 = j$. By the formula for combinations with repetitions (take j out of 3 with repetitions) we have

$$\dim V_j = \binom{j+2}{2}. \quad (2.6)$$

Consider the linear map from V_j to $G_c(j)$ given by

$$f \to (f(P_1'), \cdots, f(P_n')). \quad (2.7)$$

Since $n > mj$, it follows from the Bezout Theorem that an element $f \in V_j$ is mapped to $\mathbf{0}$ if and only if f has C as a factor. If $j < m$, the claim (2.4) now follows from (2.6). If $j \geq m$, the dimension of the kernel for (2.7) is $\binom{j-m+2}{2}$, and (2.5) follows from (2.6).

As we shall later see, the length of the considered codes can in many cases be determined by direct calculations, either by hand or by computer. On the theoretical level we emphasize the following well-known result [12].

Serre's Improvement of the Weil Bound: Let n denote the length of the codes $G_c(j)$ and $H_c(j)$. If the curve C is absolutely irreducible, then

$$n \leq q + 1 + g\lfloor 2\sqrt{q} \rfloor \quad (2.8)$$

where g is the genus of the curve C.

The genus g of a curve is a number which can be calculated, and if C is regular, this number is

$$g = (1/2)(m-1)(m-2). \quad (2.9)$$

This number also appears in (2.5). Calculations give

$$\binom{j+2}{2} - \binom{j-m+2}{2}$$
$$= (1/2)(j+2)(j+1) - (1/2)(j-m+2)(j-m+1)$$
$$= mj - (1/2)(m-1)(m-2) + 1.$$

For a regular curve and $j \geq m$, the dimension for $G_c(j)$ is therefore

$$k = mj - g + 1. \quad (2.10)$$

Direct calculations show that (2.10) also holds in the cases $j = m-1$ and $j = m-2$.

What is left to consider is the minimum distance of the code $H_c(j)$. Again the genus g of the curve is an essential term in the result, in this case as a consequence of the proof, which uses the Riemann–Roch theorem stated next [13, ch. II]. The approach in the proof of Theorem 2 is from [8], which also explains the concepts involved.

Riemann–Roch Theorem: Let C be an algebraic curve of genus g, and let D be any divisor. If degree$(D) > 2g-2$, then the dimension $l(D)$ of the vector space $L(D)$ associated with D is

$$l(D) = \text{degree}(D) - g + 1.$$

Theorem 2: Suppose that the curve C is regular and that $n > mj$ and $j \geq m-2$. For the minimum distance of the code $H_c(j)$ we then have

$$d_{\min} \geq mj - 2g + 2. \quad (2.11)$$

Proof: Consider the subspace $I = \{f \in V_j | C$ is a factor of $f\}$. The vector space V_j/I has dimension $k = mj - g + 1$, cf. the proof of Theorem 1 and (2.10). If $[f_0], \cdots, [f_{k-1}]$ is a basis for this vector space, then

$$\boldsymbol{G} = \begin{bmatrix} f_0(P_1') & \cdots & f_0(P_n') \\ f_1(P_1') & \cdots & f_1(P_n') \\ f_{k-1}(P_1') & \cdots & f_{k-1}(P_n') \end{bmatrix} \quad (2.12)$$

is a generator matrix for the code $G_c(j)$. Let $s = mj - 2g + 1$. We will prove that any s columns of $\boldsymbol{G}$ are linearly independent.

Let f_0 be chosen so that $O_{P_i}(f_0) \leq 1$ for $i = 1, \cdots, n$, where $O_{P_i}(f_0)$ is the order on the curve f_0 of the point P_i. (For the existence of such an f_0 see below.[1]) The curve f_0 cuts curve C in a divisor G with degree mj. Let $L(G)$ be the vector space associated with G. The dimension of $L(G)$ is $l(G) = mj - g + 1$ according to the Riemann–Roch theorem. This is the same as the dimension of the vector space V_R with basis $\{[f_i]/[f_0] | i = 0, \cdots, k-1\}$, and since clearly $V_R \subseteq L(G)$, we have $V_R = L(G)$.

[1]Let l be a line in PG(2, F) not tangent to C in any of the points P_i'. Since there are $q^2 + q + 1$ lines in PG(2, q), the existence of l is guaranteed by Lemma 1. Now f_0 can be obtained by raising the equation of l to the power j.

Now, take any s rational points $P_{i_1}, \cdots, P_{i_s}$, and consider the divisor

$$G_1 = G - P_{i_1} - \cdots - P_{i_s}.$$

Since degree$(G_1) = mj - s = 2g - 1 > 2g - 2$, it follows from Riemann–Roch that the dimension of $L(G_1)$ is $l(G_1) = l(G) - s = g$. An element $[f_i]/[f_0] \in V_R = L(G)$ belongs to $L(G_1)$ if and only if $f_i(P_{i_j}) = 0$, $j = 1, \cdots, s$. This follows by considering, from each point P_{i_j}, the two cases $f_0(P_{i_j}) = 0$ and $f_0(P_{i_j}) \neq 0$ and using the fact that $f_0(P_{i_j}) = 0$ iff ord $P_{i_j}(f_0) = 1$. Therefore, the solutions $(\lambda_0, \cdots, \lambda_{k-1})$ of the equations

$$\lambda_0 f_0(P_{i_j}) + \cdots + \lambda_{k-1} f_{k-1}(P_{i_j}) = 0, \qquad j = 1, \cdots, s \tag{2.13}$$

form a vector space of dimension g. This however, means that the rank of the matrix for the system (2.13) is $k - g = s$. Consequently, any s columns of $\boldsymbol{G}$ are linearly independent, and the minimum distance for the code $H_c(j)$ is at least $s + 1$.

We will briefly consider the question concerning the conditions under which the codes $H_c(j)$ are of the form $G_c(j')$ for some j', that is, the condition under which the duals of the codes we consider are of the same kind.

The reason for this question is that the decoding method works for the codes $H_c(j)$, and it would be nice to have the generator matrix of this code. In general, this seems to be a difficult question [9], but based on formula (2.10) for the dimensions it is fairly easy to derive a necessary condition.

Proposition 1: Let $j, j' \geq m - 2$, $n > mj$ and $n > mj'$ If $H_c(j) = G_c(j')$, then m divides n, and $j + j' = m - 3 + (n/m)$.

Proof: If $H_c(j) = G_c(j')$, then $n - k(j) = k(j')$ and therefore,

$$n - (mj - g + 1) = mj' - g + 1$$

so

$$n = m(j' + j) - 2g + 2 = m(j' + j) - m(m - 3)$$

from which the proposition follows.

We shall return to this in connection with the codes treated in the next section. We conclude this section with the following example.

Example 1: Let C be the Klein quartic with the equation

$$x^3y + y^3z + z^3x = 0$$

and let F be GF(8). This curve is considered in [7] and it is known that C is regular with genus $g = 3$. If n is the number of rational points on C, Serre's improvement of the Weil bound gives $n \leq 24$.

It is actually easy to find 24 points in GF(8) on the curve, namely, the points $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ and the 21 points of the form $(1, \beta, \gamma)$ where β is any nonzero element of GF(8), and γ is any one of the three different solutions to $z^3 + \beta^3 z + \beta = 0$. Putting $z = \beta^5 w$, the equation becomes $\beta(w^3 + w + 1) = 0$, and hence the given equation has three different solutions.

Using Theorem 1, we have thus constructed codes G_c corresponding to $j = 2,3,4,5$, with parameters

$$(24,6,16) \quad (24,10,12) \quad (24,14,8) \quad (24,18,4).$$

From (2.2) and Theorem 2 it follows that the codes H_c for $j = 2,3,4,5$, have parameters

$$(24,18,4) \quad (24,14,8) \quad (24,10,12) \quad (24,6,16).$$

Theorem 1 and Theorem 2 give lower bounds on the minimum distance, but for the codes considered here it turn out to be easy to find codewords of minimum weight; thus the theorem actually gives the true distance. It is not a coincidence that length and dimension for the two series of codes above are identical because it can be shown that for these codes we have

$$G_c(j) = H_c(7 - j), \qquad j = 2,3,4,5. \tag{2.14}$$

It is clear that the necessary condition in Proposition 1 is satisfied for these codes, and (2.14) then follows either by direct calculation of the matrices or by arguing in the same manner as in the Appendix for another class of codes.

III. Classes of Codes

To construct good codes of the type characterized in Section II, we need irreducible curves of low genus and with a large number of rational points. We shall be interested in obtaining infinite classes of good codes based on such curves.

For any particular field F and q elements, the number of rational points is, of course, bounded above by the number of F-points in P^2,

$$n \leq q^2 + q + 1.$$

Thus an infinite class of codes must use a sequence of alphabets of increasing size.

Example 2: The projective line is described by the equation

$$x + y + z = 0.$$

It has degree $m = 1$ and genus $g = 0$, and $n = q + 1$ points: $(0,1,-1)$ and $(1, \alpha, -1 - \alpha)$, $\alpha \in F$. Using (2.10) and Theorem 2, we see that the code $H_c(j)$ has dimension $n - j - 1$ and minimum distance $j + 2$, and thus the code is maximum distance separable (MDS). Using the polynomials $x^j, x^{j-1}y, \cdots, y^j$ as a basis, we get the parity-check matrix of the doubly extended Reed–Solomon codes.

For codes on a curve with genus g, we noted Serre's improvement of Weil's upper bound on the number of rational points (2.8). Actually, for $q = p^s$, the number of rational points on an irreducible smooth curve is [14]

$$n_s = 1 + q - \sum_{i=1}^{2g} \alpha_i^s. \tag{3.1}$$

Here the α_i are pairs of complex conjugates with $|\alpha_i| = p^{s/2}$. When n_s is known for sufficiently many values of s, it is possible to find the α_i and obtain the number of points in

other fields. Sometimes $\alpha_i^s = -p^{s/2}$ for all i, and n_s attains the upper bound.

Example 3: For the Hermitian curve described by the equation

$$x^{p+1} + y^{p+1} + z^{p+1} = 0$$

the genus is $g = (1/2)p(p-1)$ and the degree is $m = p+1$. We know that $n_2 = 1 + p^3$ [15], and it follows that

$$n_s = 1 + q - g\left(i\sqrt{p}\right)^s - g\left(-i\sqrt{p}\right)^s$$

$$n_s = \begin{cases} 1+q, & \text{for } s \equiv 1,3 \pmod 4 \\ 1+q+2g\sqrt{q}, & \text{for } s \equiv 2 \pmod 4 \\ 1+q-2g\sqrt{q}, & \text{for } s \equiv 0 \pmod 4. \end{cases}$$

For $s \equiv 2 \pmod 4$ the parameters of the codes $H_c(j)$ are

$$n = 1 + q + p(p-1)\sqrt{q}$$
$$n - k = j(p+1) - p(p-1)/2 + 1$$
$$d \geq j(p+1) - p(p-1) + 2$$

for

$$p - 1 \leq j < n/(p+1).$$

As demonstrated by Example 3, a family of codes derived from a fixed curve will satisfy

$$\lim_{q \to \infty} n/q = 1$$

and thus the codes are not significantly longer than Reed–Solomon codes. However, we may use the Hermitian curves as a basis for a more interesting class of codes, as suggested by van Lint [8].

Example 4: Let C be the curve given by the equation

$$x^{r+1} + y^{r+1} + z^{r+1} = 0$$

where $q = r^2$. The length of the codes is now $n = q^{3/2} + 1$.

For two fields of particular interest we get $q = 16$: $m = 5$, $g = 6$ and for $3 \leq j \leq 11$, the parameters of the codes are $n = 65$, $n - k = 5j - 5$, $d = 5j - 10$. For $q = 256$: $m = 17$, $g = 120$, and for $15 \leq j \leq 240$, the parameters of the codes are $n = 4097$, $n - k = 17j - 119$, $d = 17j - 238$.

In a later section, we treat the decoding of these particular codes in an example, and in the Appendix we show that for this class we have $G_c'(j) = H_c'(q - 2 - j)$, $\sqrt{q} - 2 \leq j \leq q - \sqrt{q}$. The codes of Example 4 attain the upper bound (2.8) and thus have the most favorable rates g/n. When a code is derived from a plane curve, we note an additional restriction.

Lemma 1: The number of rational points on a plane curve satisfies

$$n \leq qm + m - q. \tag{3.2}$$

Proof: If there is no rational point on the curve, the lemma is true. If there is a rational point on the curve, consider all the lines in PG(2, F) through this point. All other rational points have to be on these lines. From the theorem of Bezout, every such line has m points in common with the curve. The number of these lines is $|F| + 1 = q + 1$. In conclusion, $n \leq (q+1)(m-1) + 1 = qm + m - q$. The factor $m - 1$ in the last expression comes from the fact that every line contains the fixed point.

Combining (3.2) and (2.8), we see that the number of points can satisfy Weil's bound only when $n \leq q^{3/2}$. Longer codes must have relatively large genus and are inferior to codes constructed from curves in spaces of higher dimensions.

For q a square, it is known [4] that there exist curves for which

$$g/n \simeq \left(\sqrt{q} - 1\right)^{-1}. \tag{3.3}$$

We shall consider classes of codes to be good if

$$\lim_{q \to \infty} g/n = 0. \tag{3.4}$$

From (2.10) and (2.11) we find that for large j (3.4) implies

$$\lim_{q \to \infty} d/(n-k) = 1 \tag{3.5}$$

and thus the codes are close to being MDS for large q. It may be noted that none of the well-known classes of codes constructed by other methods (apart from Reed–Solomon codes) satisfy (3.5). In particular, second-order BCH codes of length q^2 do not have this property.

Unfortunately, the literature gives only a few results on specific plane curves over finite fields. The following examples demonstrate that, in some cases, good curves can be constructed from polynomials in a single variable.

For $m \simeq q^{2/3}$, it follows from Lemma 1 that the best plane curves have $n \simeq q^{5/3}$. Thus g/n will be of the order $q^{-1/3}$, and the distances are inferior to the asymptotic result (3.3) but still good in the sense that (3.4) is satisfied. The following example gives a class of codes with approximately these parameters. The advantage of this construction is that it can be applied in cases where q is not a square.

Example 5: Let p be a prime power, and

$$u(w) = w^{p^r+1} - w + 1$$

and let γ be a root of u. Thus

$$\gamma^{p^r+1} = \gamma - 1$$

and direct calculation shows that

$$\gamma^{p^{3r}-1} = \left[\gamma\left(\gamma^{p^r+1}\right)^{p^r}\right]^{p^r-1} = 1.$$

Therefore, γ is in GF(p^{3r}), and u has $p^r + 1$ distinct roots in GF(p^{3r}). Making the substitution $w = u^{p^r-1}$, we get

$$u^{p^{2r}-1} - u^{p^r-1} + 1$$

with $p^{2r} - 1$ distinct roots in GF(p^{3r}) since $p^r - 1$ divides $p^{3r} - 1$. Now, with the substitution $u = zy^{p^r}$, we obtain

$$z^{p^{2r}-1}y^{p^{3r}-p^r} - z^{p^r-1}y^{p^{2r}-p^r} + 1$$

$$= z^{p^{2r}-1}y^{1-p^r} - z^{p^r-1}y^{p^{2r}-p^r} + 1$$

and the curve

$$x^{p^r-1}z^{p^{2r}-1} - z^{p^r-1}y^{p^{2r}-1} + x^{p^{2r}-1}y^{p^r-1} = 0$$

which has $(p^{3r}-1)(p^{2r}-1)$ rational points with all coordinates nonzero and the additional points $(1,0,0),(0,1,0)$, $(0,0,1)$. Klein's quartic may be considered as the first member of this sequence for $p=2$ and $r=1$.

The approach taken in Example 5 may be extended to other polynomials. However, we do not have families of curves that yield good codes over all fields. To get a good code over a particular field, GF(p^r), we take a polynomial $u(w)$ with coefficients in GF(p) and roots in GF(p^r). By substituting $y^s z$ for w, multiplying by a power of y, and homogenizing with x, we obtain an equation for a curve. Usually, the curve has $(p^r-1)\deg(u)$ rational points with nonzero coordinates and a few more with one coordinate equal to zero. It is necessary to check that the curve is irreducible. As discussed in Section II, it is sufficient that the curve be regular. If there are singular points, it is necessary to ensure that these are not intersections of components of lower degree. Bezout's theorem may be used to predict the degree of such factors. The following example is the result of a computer search for curves that yield codes with low g/n over some fields that were not covered by previous constructions.

Example 6: For the field GF(32) a complete search was done for binary polynomials and substitutions to find the best ratio g/n. Using

$$u(w) = w^5 + w^2 + 1$$

and $w = y^{11}z$, we get

$$x^2y^2z^5 + x^7z^2 + y^9 = 0$$

with 157 rational points. The points $(0,0,1)$ and $(1,0,0)$ are singular, but the curve is irreducible. Because of the two singular points, the genus is reduced by two, relative to (2.9):

$$g = (9-1)(9-2)/2 - 2 = 26.$$

Thus g/n becomes 0.166. A similar search in GF(128) gave $g/n = 0.085$ for the curve

$$x^3yz^{10} + x^{13}z + y^{14} = 0$$

which has 891 rational points. This curve was constructed from the polynomial

$$u(w) = w^{10} + w + 1$$

and the substitution $w = y^{113}z$.

As noted earlier, a good code based on a plane curve cannot have length much greater than $q^{3/2}$. To obtain longer codes, we must consider curves in spaces of higher dimension. If we restrict the curves considered to be intersections of hypersurfaces, it is possible to generalize our approach to such curves. We briefly discuss a particular class of codes in the following example.

Example 7: Let $q = p^{2r}$, and consider the curve

$$\begin{Bmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ & & \cdots & \\ a_{(s-2)1} & a_{(s-2)2} & \cdots & a_{(s-2)s} \end{Bmatrix} \begin{Bmatrix} x_1^{p^r+1} \\ x_2^{p^r+1} \\ \cdots \\ x_s^{p^r+1} \end{Bmatrix} = \mathbf{0}.$$

The curve is regular if all $(s-2)\times(s-2)$ submatrices are nonsingular. For $s \le \sqrt{q}+1$, such a matrix may be obtained as the parity-check matrix of an MDS code with alphabet GF(p^r). The p^r+1st powers of the coordinates then form codewords in the MDS code. From the weight distribution of an MDS code we could find the exact number of nonzero coordinates. However, since most codewords in such a code have weight s, we may use the approximation

$$n \simeq p^r(p^r+1)^{s-1} \simeq \sqrt{q}^s$$

where p^r is the approximate number of codewords with first coordinate equal to one, and the remaining $s-1$ coordinates may assume p^r+1 values in GF(p^{2r}). The genus of the curve is approximately

$$g \simeq (1/2)(s-2)p^{r(s-1)}.$$

Thus for fixed s, g/n vanishes for $q \to \infty$. In three-dimensional space $s=4$, there are codes of length $n=(q-1)^2$ and genus $g = \sqrt{q}(q+\sqrt{q}-1)$. In GF(64) these codes have parameters $(3969, 81k, 81(42-k))$ for $11 \le k \le 38$.

IV. Decoding Procedures

In Section II we constructed, for a given irreducible curve C, two kinds of codes, $G_c(j)$ and $H_c(j)$. For the codes $H_c(j)$—or shortened versions thereof—it follows from Theorem 2 that the number

$$D = mj - 2g + 2$$

is a lower bound on the minimum distance. This number D is called the *designed distance* for the code in question. In the following, we shall describe a decoding procedure for these codes which turns out to be a generalization of the well-known Peterson algorithm for decoding Reed–Solomon codes [16, sec. 9.4]. In the general setup, this procedure does not correct errors up to half the designed distance but only up to a smaller number. For all the examples we have considered it has been possible to extend the general method in such a way that any number of errors up to half the designed distance can be corrected. However, we have not yet been able to describe and prove such extensions for the general case.

Let us consider an absolutely irreducible plane curve $C(x, y, z)$. Among the rational points on this curve we only take points of the form $(1, y_i, z_i)$. These points are denoted $Q_1, \cdots, Q_n$.

Let $H_c'(j)$ be the code whose parity-check matrix is obtained from the parity-check matrix of $H_c(j)$ by deleting the columns corresponding to the points $(0, y_i, z_i)$.

More precisely, if

$$G_c'(j) = \{(f(Q_1), \cdots, f(Q_n)) \mid f \in V_j\} \quad (4.1)$$

then

$$H_c'(j) = G_c'(j)^{\perp}. \quad (4.2)$$

As before, m is the degree of C and g the genus of C, and in accordance with (2.10) we define the number $k(j)$, $(j \geq m-2)$ as

$$k(j) = mj - g + 1. \quad (4.3)$$

From Theorem 1 and its proof, and Theorem 2, we have the following.

Theorem 3: Suppose that $n > mj$ and $j \geq m-2$. Then for the code $H_c'(j)$ with length n, the dimension is $n - k(j)$, and the minimum distance satisfies

$$d_{\min} \geq mj - 2g + 2 = k(j) - g + 1. \quad (4.4)$$

We now consider a code $H_c'(j)$ for which the assumptions in Theorem 3 are satisfied. Let $f_0, \cdots, f_{s-1}$ be the polynomials $x^{j_1}y^{j_2}z^{j_3}$, $j_1 + j_2 + j_3 = j$, where $s = \binom{j+2}{2}$. Then

$$\boldsymbol{H} = \{f_i(Q_l)\}, \qquad i = 0, \cdots, s-1, \quad l = 1, \cdots, n \quad (4.5)$$

is a parity-check matrix for $H_c'(j)$. Since we only insert points Q_l with first coordinate equal to 1, it is sufficient instead of the f_i to consider the polynomials in two variables

$$\varphi_i(y, z) = \varphi_{i_1, i_2}(y, z) = y^{i_1}z^{i_2}, \qquad i_1 + i_2 \leq j. \quad (4.6)$$

The notation in (4.6) implicitly involves an arrangement of the pairs (i_1, i_2), $i_1 + i_2 \leq j$, in a sequence. We shall use the ordering known from Cauchy multiplication of series, that is, the pairs are arranged in the sequence $(0,0)$, $(1,0),(0,1),(2,0),(1,1)\cdots$, etc. By τ we denote the map for which $\tau(i_1, i_2) = i$ in (4.6). It can be seen that $\tau(i_1, i_2) = (1/2)(i_1 + i_2 + 1)(i_1 + i_2) + i_2$, and to find $\tau^{-1}(i)$, first solve $\binom{j}{2} \leq i < \binom{j+1}{2}$. Then $i_1 + i_2 = j - 1$ and $i_2 = i - \binom{j}{2}$.

If we put

$$Q_l' = (y_l, z_l), \qquad l = 1, \cdots, n \quad (4.7)$$

we can write the matrix (4.5) as

$$\overline{\boldsymbol{H}} = \{\varphi_i(Q_l')\}, \qquad i = 0, \cdots, s-1, \quad l = 1, \cdots, n. \quad (4.8)$$

Recall that the designed distance is defined as

$$D = mj - 2g + 2. \quad (4.9)$$

In the following analysis, t denotes an (unknown) number, $t \leq (D-1)/2$, and we suppose that the actual number of errors is less than or equal to t.

We then develop a decoding algorithm, and to make this algorithm work, we have to put restrictions on t. The final value of t is implicit in the inequalities in Theorem 4.

Now let us consider a vector $\boldsymbol{v}$, for which $\boldsymbol{v} - \boldsymbol{e} \in H_c'(j)$, where the error vector $\boldsymbol{e}$ is nonzero for at most t positions. We then calculate the vector $\boldsymbol{Hv} = \boldsymbol{He}$. It will be convenient to look at exactly t positions in $\boldsymbol{e}$. Let I denote the set of these indices, and let e_i denote the value at position $i \in I$. If $\boldsymbol{S} = \boldsymbol{He}$ and $\boldsymbol{S} = (S_0, S_1, \cdots, S_{s-1})$, we then have

$$S_{l_1, l_2} = S_l = \sum_{i \in I} e_i y_i^{l_1} z_i^{l_2}, \quad \tau(l_1, l_2) = l, \qquad l = 0, \cdots, s-1. \quad (4.10)$$

In the decoding situation, we know the vector $\boldsymbol{S}$ and, assuming that at most t errors have occurred, we know that equation (4.10) has a unique solution. Some of the values e_i for this solution may be zero, namely, when the actual number of errors is less than t. The decoding problem is then to determine from (4.10) the error positions I and the error values e_i.

The key idea in the decoding algorithm we shall describe is the following: we first determine a *locator polynomial* $\sigma(y, z)$ which among its roots has the points Q_i', $i \in I$. By intersecting the curve given by $\sigma(y, z) = 0$ and the given curve (where $x = 1$), we find possible error positions. Inserting these points in an equation corresponding to (4.10), but possibly with more than t points, we obtain a system of equations from which the values e_i are calculated. Before describing the details, we illustrate the idea in the following example.

Example 8: We consider the code obtained from Example 1 by shortening by two positions for which $x = 0$. It is a $(22, 12, 8)$-code. In this example we shall demonstrate how at most three errors may be corrected.

The ten syndrome equations (4.10) may be written

$$\sum_{i=1}^{3} e_i y_i^{l_1} z_i^{l_2} = S_{l_1 l_2}, \qquad l_1 + l_2 \leq 3.$$

We now assume that we have received a word such that the values of the syndromes are given by $S_{00} = \alpha^5$, $S_{10} = \alpha$, $S_{01} = \alpha^6$, $S_{20} = \alpha^6$, $S_{11} = \alpha$, $S_{02} = \alpha^5$, $S_{30} = \alpha^5$, $S_{21} = 0$, $S_{12} = \alpha^6$, and $S_{30} = \alpha^2$. If the errors were located on a straight line, we could find a nonzero error locator $\sigma_{00} + \sigma_{10}y + \sigma_{01}z$ which would be a linear recursion among the syndromes; namely, since

$$(\sigma_{00} + \sigma_{10}y + \sigma_{01}z)S_{l_1, l_2} = 0,$$

we get

$$\sigma_{00}S_{l_1 l_2} + \sigma_{01}S_{l_1+1, l_2} + \sigma_{01}S_{l_1, l_2+1} = 0.$$

However, the syndromes do not satisfy such a relation. By adding a second-order term, we can always find a locator polynomial

$$\sigma(y, z) = y^2 + \sigma_{10}y + \sigma_{01}z + \sigma_{00}.$$

By the same reasoning as before, the coefficients may be determined by requiring that σ be a recursion among the known syndromes:

$$\begin{pmatrix} S_{00} & S_{10} & S_{01} \\ S_{10} & S_{20} & S_{11} \\ S_{01} & S_{11} & S_{02} \end{pmatrix} \begin{pmatrix} \sigma_{00} \\ \sigma_{10} \\ \sigma_{01} \end{pmatrix} = \begin{pmatrix} S_{20} \\ S_{30} \\ S_{21} \end{pmatrix}$$

which implies $\sigma(y, z) = y^2 + \alpha^2 y + \alpha^5 z$, where α is a root

of x^3+x+1. Intersecting C and σ, we obtain the five points $(1,0,0)$, $(1,1,\alpha)$, $(1,\alpha,1)$, $(1,\alpha^3,\alpha^3)$, and $(1,\alpha^4,1)$. If we expand the syndrome equations to include these five points, we get a unique solution: $e_{i_1}=1$, $e_{i_2}=\alpha^2$, $e_{i_3}=0$, $e_{i_4}=\alpha$, $e_{i_5}=0$.

The important term in the decoding is a certain matrix $\bar{S}$, whose construction we will now describe. First, however, we assume the existence of a number h such that

$$t+1\le k(h)\le D-g-t. \tag{4.11}$$

Clearly, (4.11) implies that

$$t\le\frac{D-g-1}{2}. \tag{4.12}$$

If $h'=j-h$, we have

$$k(h')=k(j-h)=mj-2g+2-k(h)$$

provided that $h\ge m-2$ and $j-h\ge m-2$. In this case, it follows from (4.11) that

$$k(h')\ge t+g. \tag{4.13}$$

To summarize, we assume the existence of a number h such that $m-2\le h\le j-m+2$ and (4.11) is satisfied. Putting $h_1=(1/2)(h+1)(h+2)$ and $h_1'=(1/2)(h'+1)(h'+2)$, we define $\bar{S}$ by

$$\bar{S}=\{a_{l,r}\},\qquad l=0,\cdots,h_1'-1,\quad r=0,\cdots,h_1-1$$

where

$$a_{l,r}=S_{l_1+r_1,l_2+r_2}$$
$$\tau(l_1,l_2)=l\qquad\tau(r_1,r_2)=r. \tag{4.14}$$

Lemma 2: For any l and r in (4.14), we have

$$l_1+r_1+l_2+r_2\le j.$$

Proof: From (4.14) and the definition of τ, we have $l_1+l_2\le h'$ and $r_1+r_2\le h$.

This means that all terms $S_{a,b}$ is $\bar{S}$ are known. Let us denote the indices in I in (4.10) by $i_1,i_2,\cdots,i_t$, and consider the following three matrices:

$$\boldsymbol{F}=\{b_{l,u}\},\qquad l=0,\cdots,h_1'-1,\quad u=1,\cdots,t$$
$$b_{l,u}=y_{i_u}^{l_1}z_{i_u}^{l_2}\qquad\tau(l_1,l_2)=l \tag{4.15}$$
$$\boldsymbol{E}=\{c_{u,v}\},\qquad u=1,\cdots,t,\qquad v=1,\cdots,t$$
$$c_{u,u}=e_{i_u}\qquad c_{u,v}=0\qquad u\ne v \tag{4.16}$$
$$\boldsymbol{P}=\{d_{v,r}\},\qquad v=1,\cdots,t,\qquad r=0,\cdots,h_1-1$$
$$d_{v,r}=y_{i_v}^{r_1}z_{i_v}^{r_2}\qquad\tau(r_1,r_2)=r. \tag{4.17}$$

We then have the following lemma.

Lemma 3: The following decomposition holds:

$$\bar{S}=\boldsymbol{FEP}. \tag{4.18}$$

Proof: The rth column in $\bar{E}\bar{P}$ is the vector $(e_{i_v}y_{i_v}^{r_1}z_{i_v}^{r_2})$, $v=1,\cdots,t$. Hence the element at position (l,r) in the product on the right side of (4.18) is

$$\sum_{v=1}^{t}y_{i_v}^{l_1}z_{i_v}^{l_2}e_{i_v}y_{i_v}^{r_1}z_{i_v}^{r_2}=\sum_{v=1}^{t}e_{i_v}y_{i_v}^{l_1+r_1}z_{i_v}^{l_2+r_2}=S_{l_1+r_1,l_2+r_2}.$$

Lemma 4: The matrix $\boldsymbol{F}$ has rank t.

Proof: Consider the code $H_c(h')$. By (4.15) $\bar{F}$ consists of t columns of a parity-check matrix for this code, namely, the t columns corresponding to the points $(1,y_{i_v},z_{i_v})$, $v=1,\cdots,t$. The minimum distance of $H_c(h')$ is by Theorem 2 greater than or equal to $mh'-2g+2$. However, by (4.13) we have

$$t<k(h')-g+1=mh'-2g+2$$

and the lemma follows.

Next, let us consider the following system of equations:

$$\bar{S}\boldsymbol{\sigma}=\boldsymbol{0}. \tag{4.19}$$

For any solution $\boldsymbol{\sigma}=(\sigma_0,\sigma_1,\cdots,\sigma_{h_1-1})$ to (4.19) let $\sigma(y,z)$ be the polynomial in two variables

$$\sigma(y,z)=\sum\sigma_{ab}y^az^b,\qquad\tau(a,b)\in\{0,\cdots,h_1-1\}. \tag{4.20}$$

Note that the degree of $\sigma(y,z)$ is at most h, and $\sigma_{ab}=\sigma_{\tau(a,b)}$.

Lemma 5: Let $\boldsymbol{\sigma}$ be any solution to (4.19), and let (y_{i_v},z_{i_v}) be a point where $i_v\in I$ and $e_{i_v}\ne 0$. Then

$$\sigma(y_{i_v},z_{i_v})=0.$$

Proof: Since by Lemma 3

$$\boldsymbol{S\sigma}=\boldsymbol{F}(\boldsymbol{EP\sigma})$$

and since by Lemma 4 the matrix $\boldsymbol{F}$ has rank t, any solution to (4.19) satisfies

$$\boldsymbol{EP\sigma}=\boldsymbol{0}. \tag{4.21}$$

The component at position number v of the vector on the left side of (4.21) is

$$\sum_{s=0}^{q-1}e_{i_v}d_{v,s}\sigma_s=e_{i_v}\sum_s y_{i_v}^{s_1}z_{i_v}^{s_2}\sigma_s,\qquad\tau(s_1,s_2)=s$$
$$=e_{i_v}\sigma(y_{i_v},z_{i_v}).$$

Since by (4.12) this quantity is zero, the lemma follows.

For any polynomial $\sigma(y,z)$ in (4.20), we put

$$\sigma(x,y,z)=\sum\sigma_{ab}x^{h-a-b}y^az^b,$$
$$\tau(a,b)\in\{0,\cdots,h_1-1\}.$$

We shall use the phrase "the curve $C(x,y,z)$ is not a divisor of $\sigma(y,z)$" to mean that $C(x,y,z)$ is not a divisor of the above polynomial $\sigma(x,y,z)$.

Lemma 6: The system of equations (4.19) always has a solution $\boldsymbol{\sigma}$ such that the corresponding polynomial $\sigma(y,z)$ does not have the curve C as a divisor.

Proof: The system (4.19) has h_1 unknowns and by Lemma 3 the rank of $\bar{S}$ is at most t. Hence for any fixed $h_1-(t+1)$ positions, there is a proper solution to (4.19) which is zero at these positions.

Consider a fixed term $x^{m-a-b}y^az^b$ in the equation for the curve C for which $a+b$ is maximal. Suppose that $C(x, y, z)$ is a divisor of $\sigma(x, y, z)$. Then $\sigma(x, y, z)$ contains terms of the form $P(x, y, z)x^{m-a-b}y^az^b$, where the degree of the polynomial $P(x, y, z)$ is $h-m$. These terms correspond to at most

$$\binom{h-m+2}{2} \tag{4.22}$$

terms in $\sigma(y, z)$. Therefore, if the coefficients for these terms are all zero, then $\sigma(y, z)$ cannot have C as a divisor.

The lemma now follows if

$$\binom{h-m+2}{2} \le h_1-(t+1).$$

This, however, is the same as

$$\binom{n+2}{2}-\binom{h-m+2}{2} \ge t+1;$$

that is, $k(h) \ge t+1$, which is satisfied by the definition of h.

We now describe the decoding algorithm.

1. For the received vector $\boldsymbol{v}$, calculate the syndrome vector $\boldsymbol{S} = \boldsymbol{H}\boldsymbol{v}$.
2. Consider the matrix $\bar{S}$ defined in (4.14). Find the solution σ to the system (4.19) for which the corresponding polynomial $\sigma(y, z)$ has minimal degree and does not have the curve C as divisor.
3. Find, among the points $Q_1', \cdots, Q_n'$, those points Q_i', for which $\sigma(Q_i') = 0$. Denote these points by $Q_1'', Q_2'', \cdots, Q_t''$.
4. Insert the coordinates for $Q_1'', \cdots, Q_t''$ in the system of equations (4.10), and calculate the values e_i.

Theorem 4: The algorithm just described corrects t errors if there exists a number h such that

$$t+1 \le k(h) \le D-g-t$$

where $m-2 \le h \le j-m+2$.

Proof: Most of the proof follows directly from the preceding lemmas. By Lemma 5, the polynomial $\sigma(y, z)$ determined in step 2 has among its roots all the error positions. It may have other roots among the points $Q_1', \cdots, Q_n'$. Whatever the case, since $\sigma(y, z)$ does not have the curve C as a divisor, it follows from the Theorem of Bezout that for the number t' in step 3, we have $t' \le mh$. From (4.13) it follows that $k(j-h) \ge t+g$, which gives $mj-2g+1 \ge mh+t$, and therefore, $mh \le mj-2g+2 = D$. This implies that the system considered in step 4 has a unique solution. If there were two different solutions, we would be able to construct a codeword of weight greater than zero and less than D by adding these two solutions. Indeed, the sum of the two solutions has weight less than D since the nonzero components of this word have to be among the positions $Q_1'', \cdots, Q_t''$. The positions for which $e_i \neq 0$ are then the error positions, and e_i the corresponding error values.

If we consider the Reed–Solomon code coming from the projective line as in Example 2 and wish to correct t errors, this is possible for $t \le (D-1)/2$, since $h = (j+1)/2$ satisfies the conditions in Theorem 4. It then follows from the parity-check matrix that the matrix $\bar{S}$ used in the decoding algorithm is the well-known syndrome matrix in the Peterson algorithm and that the σ-polynomial is the ordinary error locator polynomial in one variable (no z's). Of course, in this case the error values can be found by the standard method.

V. Decoding Examples

In this section we apply the decoding algorithm derived in Section IV to some specific codes, and we discuss details and extensions of the algorithm. In Example 8 we demonstrated how a small code can be decoded using the technique of a two-variable error locator polynomial. In such small examples the restrictions and difficulties of the general case do not arise, and in fact, all of the codes based on Klein's quartic can be decoded for $t < d/2$ in a similar way. For simplicity, we shall specifically consider codes on Hermitian curves, although the decoding of other codes discussed in Section III differs only in certain minor details.

In the definition of $\bar{S}$ (4.14) we have included all syndromes of sufficiently low degree to simplify the notation. The problem of finding independent equations and independent terms in σ is then treated separately. In the case of the Hermitian curves we shall reduce the size of $\bar{S}$ from the start by considering only syndromes S_{ab} with $b < m$. Similarly, we exclude multiples of C from σ by allowing only terms $\sigma_{ab}y^az^b$ with $b < m$. As before, h is the maximal degree of σ, and we may now write the number of coefficients as

$$h_1 = k(h).$$

Similarly, h' is the maximal degree of the syndromes in any equation in the system (4.10), and the number of equations may be expressed as

$$s = h_1' = k(h').$$

The codes over GF(16) provide a simple example where the general form of Theorem 4 applies.

Example 9: In Example 4 we found the parameters of codes over GF(16) based on the Hermitian curve with $m = 5$. We reduce the length of the code by the five positions with $x = 0$ and obtain the codes (60, 30, 25), (60, 20, 35), and (60, 10, 45). For the first of these codes, Theorem 4 says that taking $h = 3$ the decoding algorithm will correct nine errors. Thus from the 30 independent syndromes $S_{a,b}$ with $b < 5$ we calculate the remaining six of degree ≤ 7, using the fact that $z^5 = y^5+1$. We get 15

equations and must determine σ of degree ≤ 3 with ten coefficients. In this case, all terms of degree 3 are used since $b < m$. In what follows, we assume that we have calculated the syndromes $S_{a,b}$ and that they are given in the following matrices; we write the integer i for the field element α^i, where α is a primitive element in GF(16); $*$ indicates the zero element:

$$\begin{bmatrix} 10 & * & 1 & 8 & 0 & * & 9 & 2 & 0 & 4 \\ * & 8 & 0 & 9 & 2 & 0 & 1 & 14 & 9 & 4 \\ 1 & 0 & * & 2 & 0 & 4 & 14 & 9 & 4 & 12 \\ 8 & 9 & 2 & 1 & 14 & 9 & 10 & 14 & 8 & 1 \\ 0 & 2 & 0 & 14 & 9 & 4 & 14 & 8 & 1 & 5 \\ * & 0 & 4 & 9 & 4 & 12 & 8 & 1 & 5 & * \\ 9 & 1 & 14 & 10 & 14 & 8 & 2 & 3 & 2 & 14 \\ 2 & 14 & 9 & 14 & 8 & 1 & 3 & 2 & 14 & 7 \\ 0 & 9 & 4 & 8 & 1 & 5 & 2 & 14 & 7 & 2 \\ 4 & 4 & 12 & 1 & 5 & * & 14 & 7 & 2 & 9 \\ 1 & 10 & 14 & 2 & 3 & 2 & 3 & 11 & 2 & 9 \\ 14 & 14 & 8 & 3 & 2 & 14 & 11 & 2 & 9 & * \\ 9 & 8 & 1 & 2 & 14 & 7 & 2 & 9 & * & 13 \\ 4 & 1 & 5 & 14 & 7 & 2 & 9 & * & 13 & 12 \\ 12 & 5 & * & 7 & 7 & 9 & * & 13 & 12 & 2 \end{bmatrix} \cdot \begin{bmatrix} \sigma_{00} \\ \sigma_{10} \\ \sigma_{01} \\ \sigma_{20} \\ \sigma_{11} \\ \sigma_{02} \\ \sigma_{30} \\ \sigma_{21} \\ \sigma_{12} \\ \sigma_{03} \end{bmatrix} = \mathbf{0}.$$

We may solve this system to obtain the unique solution

$$\sigma = 1 + \alpha^4 y + a^5 z + \alpha^{13} y^2 + \alpha^7 zy + \alpha^5 z^2 + \alpha^{12} y^3 + \alpha^3 y^2 z + \alpha^8 y z^2 + z^3.$$

Since σ has degree 3, there may be 15 intersections with C. In this case 14 of the points have $x \neq 0$. We list (y, z):

$$(0,1),(0,\alpha^3),(0,\alpha^{12}),(1,0),(\alpha,\alpha^{14}),(\alpha^2,\alpha^{10}),$$
$$(\alpha^2,\alpha^7),(\alpha^2,\alpha^{13}),(\alpha^5,\alpha^{10}),(\alpha^{10},\alpha^{14}),$$
$$(\alpha^7,\alpha^5),(\alpha^6,0),(\alpha^{13},\alpha^2),(\alpha^{12},0).$$

We now solve the 30 independent syndrome equations with these points inserted to find the error values, of which at most nine are assumed to be nonzero. Taken in the same order as the locations, the e_j are

$$\alpha^5, \alpha^8, 0, 1, 0, \alpha^{13}, 0, \alpha^3, \alpha^{10}, 0, \alpha^2, \alpha^8, 0, \alpha^5.$$

Since a curve of degree 2 may intersect C in 10 points, it may occur that there is an error locator of degree 2. Consider the following example:

$$\begin{bmatrix} 10 & 0 & 5 & 5 & 0 & 3 & 10 & 10 & 2 & 4 \\ 0 & 5 & 0 & 10 & 10 & 2 & 0 & 5 & 6 & 10 \\ 5 & 0 & 3 & 10 & 2 & 4 & 5 & 6 & 10 & 9 \\ 5 & 10 & 10 & 0 & 5 & 6 & 5 & 0 & 3 & 2 \\ 0 & 10 & 2 & 5 & 6 & 10 & 0 & 3 & 2 & 13 \\ 3 & 2 & 4 & 6 & 10 & 9 & 3 & 2 & 13 & 0 \\ 10 & 0 & 5 & 5 & 0 & 3 & 10 & 10 & 2 & 4 \\ 10 & 5 & 6 & 0 & 3 & 2 & 10 & 2 & 4 & 10 \\ 2 & 6 & 10 & 3 & 2 & 13 & 2 & 4 & 10 & 5 \\ 4 & 10 & 9 & 2 & 13 & 0 & 4 & 10 & 5 & 0 \\ 0 & 5 & 0 & 10 & 10 & 2 & 0 & 5 & 6 & 10 \\ 5 & 0 & 3 & 10 & 2 & 4 & 5 & 6 & 10 & 9 \\ 6 & 3 & 2 & 2 & 4 & 10 & 6 & 10 & 9 & 10 \\ 10 & 2 & 13 & 4 & 10 & 5 & 10 & 9 & 10 & 10 \\ 9 & 13 & 0 & 10 & 5 & 0 & 9 & 10 & 10 & 2 \end{bmatrix} \cdot \begin{bmatrix} \sigma_{00} \\ \sigma_{10} \\ \sigma_{01} \\ \sigma_{20} \\ \sigma_{11} \\ \sigma_{02} \\ \sigma_{30} \\ \sigma_{21} \\ \sigma_{12} \\ \sigma_{03} \end{bmatrix} = \mathbf{0}.$$

This system has rank only 7, but we get a unique solution of degree 2

$$\sigma = 1 + y + y^2.$$

There are 10 points on the curve with $y = \alpha^5$ or $y = \alpha^{10}$.

$$(\alpha^5,\alpha) \quad (\alpha^5,\alpha^4) \quad (\alpha^5,\alpha^7) \quad (\alpha^5,\alpha^{10}) \quad (\alpha^5,\alpha^{13})$$
$$(\alpha^{10},\alpha^2) \quad (\alpha^{10},\alpha^5) \quad (\alpha^{10},\alpha^8) \quad (\alpha^{10},\alpha^{11}) \quad (\alpha^{10},\alpha^{14}).$$

The error values are found as before:

$$\alpha^5, \alpha^8, \alpha^{13}, 1, \alpha^3, \alpha^{10}, \alpha^8, \alpha^2, 0, \alpha^5.$$

As mentioned in the beginning of Section IV, we have found it possible to decode all errors of weight $t < d/2$ in the cases we have investigated. If $j = 6$, then the code parameters are (60, 35, 20). Again we shall correct nine errors, and consider the first 25 syndromes in the two numerical cases treated earlier. In the first case we get ten equations and find the same unique solution for a σ of degree 3. In the second case the ten equations have low rank, and we do not immediately get a solution. However, if we assume that σ has degree 2, the last five equations involve only the known syndromes, and we again find the same solution as before. If σ has degree 3, but some of the leading coefficients are zero, it is always possible to obtain additional equations. However, we have not found a proof that the system has sufficiently high rank, and we shall not discuss these cases in further detail.

For practical applications, the codes over GF(256) are particularly interesting. We shall discuss such codes in the following example and concentrate on a code with rate

close to 7/8, which is the rate used in some systems using concatenated codes.

Example 10: In Example 4 we found the parameters of codes over GF(256) derived from Hermitian curves. If we exclude points with $x = 0$, the length becomes $n = 4080$. Theorem 4 is satisfied by 102 codes with $15 \le h \le 116$. For this sequence of codes the dimension decreases and the minimum distance increases in steps of 34 from (4080, 3570, 391) to (4080, 136, 3825). The first of these codes will correct 135 errors if the decoding algorithm of Section IV is used. As indicated in the beginning of this section, the matrix $\bar{S}$ is reduced to 255×136, and we solve for a locator polynomial with 136 coefficients. If there are 135 errors, we have $8 \le \text{degree}(\sigma) \le 15$. For a locator polynomial of degree 15, there may be 255 intersections with C. From the 510 syndrome equations, we obtain the error values, of which at most 135 are assumed to be nonzero.

It would be interesting to extend the decoding algorithm in several directions. We lack a general proof that all errors of weight $< d/2$ can be corrected. Further, it would be interesting to decode the high-rate codes which are excluded by the assumptions in Theorem 4, and it would be of interest to include the points with $x = 0$. It appears to be a straightforward generalization to apply the same algorithm to the codes in higher dimensional space described in Example 7.

As mentioned in earlier sections, the decoding algorithm may be seen as a generalization of Peterson's algorithm for BCH codes. For long codes the algorithm involves very time-consuming solutions of systems of linear equations. It would be a major improvement if more efficient algorithms similar to those available for BCH codes could be found.

VI. Discussion

We have treated algebraic geometry codes in terms which should be familiar to readers with a background in standard coding theory. Many details in the construction of generator and parity-check matrices, and in the decoding algorithm, are closely related to methods with a long tradition in coding theory.

The use of projective geometries in the construction of error-correcting codes is not new, either. The essentially new aspect is the restriction of the space to the points on a curve and the results directly related to the properties of the curve.

The new codes are better than any previously known when the alphabet is large enough. This is true even for codes of moderate length. Unfortunately, our knowledge of good curves over finite fields is still rather limited, and more extensive work in this direction is needed.

Acknowledgment

The authors wish to thank the referees for their remarks and suggestions. Several useful discussions with Johan P. Hansen and Anders B. Sørensen are also gratefully acknowledged.

Appendix

We will prove that for the Hermitian curves given by the equation (Example 4) $x^{r+1} + y^{r+1} + z^{r+1} = 0$ over $F = \text{GF}(q)$, where $q = r^2$, we have

$$H_c'(j) = G_c'(q-2-j) \qquad \sqrt{q} - 2 \le j \le q - \sqrt{q}.$$

The curve has $q^{3/2} - q^{1/2} = r^3 - r$ points with x coordinate equal to 1, and hence gives codes $H_c'(j)$ of length $n = r^3 - r$. Since $r+1$ divides $r^3 - r$ (with quotient $r^2 - r$), the necessary condition in Proposition 1 is satisfied.

To prove the statement we have to show that if $F(y, z)$ is any polynomial of degree less than or equal to $r^2 - 2 = s$, then

$$\sum_{(1, y, z) \in C} F(y, z) = 0. \tag{1}$$

We write $F(y, z)$ as

$$F(y, z) = \sum_{j=0}^{s} a_j y^j + \sum_{j=1}^{s} b_j z^j + \sum_{\substack{i>0, j>0 \\ i+j \le s}} a_{ij} y^i z^j. \tag{2}$$

The points $(1, y, z)$ of C have the form $(1, \alpha, 0)$, where α is one of the $r+1$ elements of $\text{GF}(r^2)$, where $\alpha^{r+1} = -1$, or the form $(1, \alpha, \beta)$ where α is an element of $\text{GF}(r^2)$, where $\alpha^{r+1} \ne -1$ and β is any one of the $r+1$ solutions to $z^{r+1} = -1 - \alpha^{r+1}$.

The term $\sum_{j=0}^{s} a_j y^j$ then contributes $na_0 + \sum_{j=0}^{s} a_j \sum_{\alpha \in \text{GF}(r^2)} \alpha^j = 0$ to (1) since r divides n and $j \le r^2 - 2$ [15, p. 321]. The term $\sum_{j=1}^{s} b_j z^j$ contributes (for a fixed α, where $\alpha^{r+1} \ne -1$) $\sum_{j=1}^{s} b_j \sum_{\beta} \beta^j = 0$ to (1) since the inner sum is the sum of the roots of the equation $z^{r+1} = (-1 - \alpha^{r+1})^j$.

The last term

$$\sum_{\substack{i>0, j>0 \\ i+j \le s}} a_{ij} y^i z^j$$

contributes

$$\sum_{\substack{i>0, j>0 \\ i+j \le s}} a_{ij} \sum_{\substack{\alpha \\ \alpha^{r+1} \ne -1}} a^i \sum_{\substack{\beta \\ \beta^{r+1} = -1 - \alpha^{r+1}}} \beta^j$$

which, by the same reasoning, is zero. This result proves the fact corresponding to [9, theorem 2] for the codes we consider.

References

[1] V. D. Goppa, "Codes associated with divisors," *Probl. Peredach, Inform.*, vol. 13, no. 1, pp. 33–39, 1977.
[2] ____, "Codes on algebraic curves," *Soviet Math. Dokl.*, vol. 24, no. 1, pp. 170–172, 1981.
[3] ____, "Algebraico–geometric codes," *Math. USSR Izvestiya*, vol. 21, no. 1, pp. 75–91, 1983.
[4] M. A. Tsfasman, S. G. Vladut, and T. Zink, "Modular curves, Shimura curves and Goppa codes, better than Varshamov–Gilbert bound," *Math. Nachr.*, vol. 104, pp. 13–28, 1982.
[5] G. L. Katsman, M. A. Tsfasman, and S. G. Vladut, "Modular curves and codes with a polynomial construction," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 353–355, 1984.
[6] G. Lachaud, "Les codes geometriques de Goppa," *Seminaire Bourbaki*, no. 641, 1985.
[7] J. P. Hansen, "Codes on the Klein quartic, ideals and decoding," *IEEE Trans. Inform. Theory*, vol. IT-33, no. 6, pp. 919–923, 1987.
[8] J. H. van Lint and T. A. Springer, "Generalized Reed–Solomon codes from algebraic geometry," *IEEE Trans. Inform. Theory*, vol. IT-33, no. 3, pp. 305–310, 1987.

[9] H. J. Tiersma, "Remarks on codes from Hermitian curves," *IEEE Trans. Inform. Theory*, vol. IT-33, no. 4, pp. 605–609, 1987.
[10] A. N. Skorobogatov and S. G. Vladut, "On the decoding of algebraic geometric codes," *Inst. Problems of Information Transmission Moscow*, preprint 1988.
[11] W. Fulton, *Algebraic Curves*. Reading, MA: Benjamin Cummings, 1969.
[12] J. P. Serre, "Sur le nombre des points rationnels d'une courbe algébraique sur un corps fini," *C.R. Acad. Sci.*, vol. 296, pp. 397–402, 1983.
[13] M. Deuring, *Lectures on the Theory of Algebraic Functions of One Variable*. Berlin: Springer Lecture Notes in Mathematics, vol. 314, 1973.
[14] R. D. Hartshorne, *Algebraic Geometry*. New York: Springer-Verlag, Graduate Texts in Mathematics, vol. 52, 1977.
[15] J. W. P. Hirschfeld, *Projective Geometries over Finite Fields*. Oxford, UK: Oxford University, 1979.
[16] W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes*. Cambridge, MA: MIT Press, 1972.