Fast Decoding of Codes from Algebraic Plane Curves

J. Justesen, K. J. Larsen, H. Elbrønd Jensen, and T. Høholdt

Abstract—**Improvement to an earlier decoding algorithm for codes from algebraic geometry is presented. For codes from an arbitrary regular plane curve we correct up to $d^*/2 - m^2/8 + m/4 - 9/8$ errors, where d^* is the designed distance of the code and m is the degree of the curve. The complexity of finding the error locator is $O(n^{7/3})$, where n is the length of the code. For codes from Hermitian curves the complexity of finding the error values, given the error locator, is $O(n^2)$, and the same complexity can be obtained in the general case if we only correct $d^*/2 - m^2/2$ errors.**

Index Terms—**Decoding, algebraic geometry codes.**

I. Introduction

In [1], we presented an algorithm for the decoding of codes constructed from a nonsingular plane algebraic curve. This algorithm has complexity $O(n^3)$, where n is the length of the code, and corrects $d^*/2 - m^2/4$ errors, where d^* is the designed distance of the code, and m is the degree of the curve involved in the construction.

In [2], this algorithm was treated in the proper algebraic geometry setting by A. N. Skorobogatov and S. G. Vlăduţ, so they could decode codes from arbitrary algebraic curves, i.e., geometric Goppa codes, and in some cases more errors were corrected.

Based on their results, and some deep algebraic geometry, R. Pellikaan [3] proved the existence of a polynomial time algorithm, which corrects $(d^* - 1)/2$ errors for codes from maximal curves and recently S. G. Vlăduţ [4] extended this result to any geometric Goppa code.

In this paper, we return to the codes treated in [1], and improve on the algorithm in several ways. For codes from an arbitrary regular plane curve we correct $d^*/2 - m^2/8 + m/4 - 9/8$ errors. We use a modified version of an algorithm by Sakata [5] to find the error locator in time $O(mt^2)$, where m is the degree of the curve, and t is the number of errors. For good codes one has $m \sim \sqrt{q}$ and $t \le n \sim q\sqrt{q}$, so the complexity is $O(n^{7/3})$, when we consider good codes over increasing fieldsizes.

The error values are then found by a method, which for codes from Hermitian curves has complexity $O(m^2q^2)$, which under the same assumptions as above is $O(n^2)$. We also show how to find the error values with the same complexity in the general case, if we only correct $d^*/2 - m^2/2$ errors.

The paper is organized as follows. Section II reviews the code construction and the overall idea in the decoding method. Section III treats the error locator polynomials and in section IV we present the modified version of Sakata's algorithm, which we use to determine the error locator. Section V presents the method for determining the error values, both in the general case and in the case of Hermitian curve. Finally, section VI contains the conclusion and a discussion.

II. The Codes and Their Decoding

We shall in this section give the construction of the codes and the main ideas of their decoding. Let F_q be a finite field with q elements, and let $C(x, y)$ be a polynomial from $F_q[x, y]$.

The set of points (x, y) where x and y are in the algebraic closure $\bar{F}$ of F_q for which $C(x, y) = 0$ is called an *affine curve*. The points on the curve with both coordinates in F_q are the *rational points*. The curve is *regular* if the projective closure is regular, in particular this implies that $C(x, y)$, is absolutely irreducible. If the curve is regular and $C(x, y)$ has degree m, then the *genus* g of the curve is given by $g = (m - 1)(m - 2)/2$.

In order to describe the code construction and the decoding we need a total ordering of the pairs of natural numbers. We choose the so-called graduated total degree ordering $<_T$ where $(0, 0) <_T (1, 0) <_T (0, 1) <_T (2, 0) \cdots$. Now let $C(x, y) = 0$ be the equation of a regular curve of degree m and let $P_1, P_2, \cdots, P_n$ be the rational points of the curve. Let j be a natural number $m - 2 \le j \le \left\lfloor \frac{n-1}{m} \right\rfloor$ and let $\varphi_0(x, y), \varphi_1(x, y), \cdots, \varphi_s(x, y)$ denote the monomials $x^a y^b$, where $(a, b) \le_T (0, j)$ ordered by $<_T$. The code $C^*(j)$ is then given by its parity check matrix $\underline{\underline{H}}$

$$\underline{\underline{H}} = \begin{bmatrix} \varphi_0(P_1) & \cdots & \varphi_0(P_n) \\ \varphi_1(P_1) & \cdots & \varphi_1(P_n) \\ \vdots & & \vdots \\ \varphi_s(P_1) & \cdots & \varphi_s(P_n) \end{bmatrix}. \tag{2.1}$$

It now follows from [1] that the dimension of $C^*(j)$ is $n - (mj - g + 1)$ and that

$$d_{\min} \ge d^* = mj - 2g + 2.$$

The number d^* is the *designed distance* of the code.

Manuscript received April 17, 1990; revised March 19, 1991. This work was presented in part at the IEEE International Symposium on Information Theory, San Diego, CA, January 14–19, 1990.

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

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

IEEE Log Number 9103459.

Example 1: Let $F_q = \mathrm{GF}(r^2)$, so $q = r^2$ and let

$$C(x, y) = x + x^r - y^{r+1}.$$

This is the affine version of the Hermitian curve as considered by H. Stichtenoth in [6]. It is regular, has degree $m = r + 1$ and therefore genus $g = r(r-1)/2$. It is well known that the affine curve has $n = r^3 = q\sqrt{q}$ rational points. The construction therefore gives a code over $\mathrm{GF}(r^2)$ with parameters $n = r^3$, $k = r^3 - j(r+1) + r(r-1)/2 - 1$, $d \geq (r+1)j - r(r-1) + 2$ for any j, where $r - 1 \leq j \leq \left\lfloor \frac{r^3}{r+1} \right\rfloor$.

In the decoding situation we receive a word $\boldsymbol{r}$ which is the sum of a codeword $\boldsymbol{c}$ and an error vector $\boldsymbol{e}$. We calculate the syndrome $\boldsymbol{s} = \underline{\underline{H}}\boldsymbol{r}^T$. If we number the coordinates of the syndrome vector, like we numbered the rows of $\underline{\underline{H}}$, and the errors occurred in the points with the coordinates (x_i, y_i) $i \in I$, $I \subseteq \{1, 2, \cdots, n\}$, with values e_i, it follows from (2.1) that

$$S_{ab} = \sum_{i \in I} e_i x_i^a y_i^b. \tag{2.2}$$

The decoding problem then is from the syndromes S_{ab}, $a + b \leq j$ to determine the error positions (x_i, y_i) $i \in I$, and the corresponding error values e_i.

The idea is now to treat the two parts of the decoding problem separately, that is, first to determine the error positions and then to determine the error values. The determination of the error positions is based on the observation that if a polynomial

$$\sigma(x, y) = \sum_{l+k \leq h} \sigma_{lk} x^l y^k,$$

has the error positions among its zeros, then

$$\begin{aligned} \sum_{l+k \leq h} \sigma_{lk} S_{a+l, b+k} &= \sum_{l+k \leq h} \sigma_{l,k} \sum_{i \in I} e_i x_i^{a+l} y_i^{b+k} \\ &= \sum_{i \in I} e_i x_i^a y_i^b \sum_{l+k \leq h} \sigma_{lk} x_i^l y_i^k \\ &= 0. \end{aligned} \tag{2.3}$$

This holds for all (a, b) if we use (2.2) as the definition of S_{ab}.

In particular we have from (2.3)

$$\begin{bmatrix} S_{00} & S_{10} & \cdots & S_{0h} \\ S_{10} & S_{20} & \cdots & S_{1h} \\ S_{01} & S_{11} & \cdots & S_{0h+1} \\ \vdots & & & \\ S_{0h'} & S_{1h'} & \cdots & S_{0h+h'} \end{bmatrix} \begin{bmatrix} \sigma_{00} \\ \sigma_{10} \\ \vdots \\ \sigma_{0h} \end{bmatrix} = \underline{0}, \tag{2.4}$$

where $h' = j - h$.

The decoding method of [1] now consists of the following steps.

1) Find a minimal degree solution σ to (2.4) such that $\sigma(x, y)$ does not have $C(x, y)$ as a factor.
2) Find among the points $P_1, \cdots, P_n$ those P_1', $P_2', \cdots, P_{t'}'$ which are zeros of $\sigma(x, y)$.
3) Insert the coordinates of $P_1', \cdots, P_{t'}'$ into (2.2) and solve for e_i's.

Theorem 4 of [1] tells us that this procedure corrects t errors provided there exists a number h such that

$$t + 1 \leq mh - g + 1 \leq d^* - g - t,$$

where

$$m - 2 \leq h \leq j - m + 2.$$

Both Step 1) and Step 2) involve the solution of systems of linear equations, so the proposed algorithm has complexity $O(n^3)$.

The improvements come from using Sakata's algorithm in Step 1), which is described in Sections III and IV, and using a new method to find the error values, which is described in Section V. Moreover, it turns out that this condition is too restrictive, so we are actually able to correct $d^*/2 - m^2/8 + m/4 - 9/8$ errors.

In the following we suppose that the equation of the curve contains the term y^m, which can always be obtained by a suitable choice of coordinates when $m < q$. The case $m \geq q$ are not interesting for regular curves, and is not considered in this paper. Moreover, to avoid treating special cases we suppose that the points $P_1, P_2, \cdots, P_n$ all have both coordinates nonzero.

III. Error Locator Polynomials

For the error positions (x_i, y_i), $i \in I$, and the error values e_i, the syndromes of the code are

$$S_{ab} = \sum_{i \in I} e_i x_i^a y_i^b, \tag{3.1}$$

where $a + b \leq j$. Moreover, we shall refer to S_{ab} defined by (3.1) as a syndrome for any $a, b < q - 1$, and distinguish between known and unknown syndromes. Section V gives a method for calculating the unknown syndromes from the known syndromes and the error locator.

Let us consider the set $\mathcal{L}$ of polynomials

$$\sigma(x, y) = \sum_{l,k} \sigma_{lk} x^l y^k,$$

which define a linear recursion among the syndromes, that is,

$$\sum_{l,k} \sigma_{lk} S_{a+l, b+k} = 0 \tag{3.2}$$

for all a, b where the indexes are calculated modulo $q - 1$. If (3.2) is satisfied, we get by inserting (3.1)

$$\begin{aligned} &\sum_{l,k} \sigma_{lk} \sum_{i \in I} e_i x_i^{a+l} y_i^{b+k} \\ &\quad = \sum_{i \in I} x_i^a y_i^b e_i \sum_{l,k} \sigma_{lk} x_i^l y_i^k = 0, \qquad \text{for all } a, b. \end{aligned}$$

Let $e_i' = e_i \sum_{l,k} \sigma_{lk} x_i^l y_i^k$, then we obtain

$$\sum_{i \in I} e_i' x_i^a y_i^b = 0, \qquad \text{for all } a, b.$$

This means that e_i', $i \in I$ is an error pattern for which all syndromes are zero and, therefore, $e_i' = 0$. Since we may

assume that $e_i \neq 0$, we get

$$\sum_{l,k} \sigma_{lk} x_i^l y_i^k = 0, \tag{3.3}$$

so $\sigma(x, y)$ has the error points as zeros.

On the other hand, if (3.3) holds, then it follows from (2.3) that (3.2) is satisfied, so that elements of $\mathscr{L}$ are exactly those polynomials that have the t error points as zeros. We also note that $\mathscr{L}$ is independent of the error values.

Let us for $f \in \mathrm{GF}(q)^t$, $f = (f_i)$, $i \in I$, consider the array T^f, where

$$T_{ab}^f = \sum_{i \in I} f_i x_i^a y_i^b$$

and let $\mathscr{S}$ be the set of all these arrays.

$\mathscr{S}$ is obviously a linear code, actually it is a two-dimensional cyclic code, and its dimension is t. It is obvious that the dimension cannot exceed t, and on the other hand no combination of error values can give an array of all zeros. By carrying out the same calculations that lead from (3.2) to (3.3) and from (3.3) to (3.2) it can easily be seen that $\mathscr{L}$ is the dual code of $\mathscr{S}$, so in particular the dimension of $\mathscr{L}$, as a vector space over GF(q), is $(q-1)^2 - t$. We will now study $\mathscr{L}$ a little closer. Clearly $C(x, y)$ and all multiples of this polynomial are in $\mathscr{L}$. The error locator should therefore be found among the other codewords of $\mathscr{L}$. We recall from [8] and [10] the properties of a minimal basis for $\mathscr{L}$.

A polynomial $f(x, y) = \sum_{i,j} f_{ij} x^i y^j$ has *leading term* $x^a y^b$ if $f_{ab} \neq 0$ and

$$f(x, y) = \sum_{(i,j) \leq_T (a,b)} f_{ij} x^i y^j.$$

A *minimal basis* for $\mathscr{L}$ is a set F of polynomials from $\mathscr{L}$

$$F = \{\sigma^{(1)}, \sigma^{(2)}, \cdots, \sigma^{(l)}\}$$

with leading terms $x^{s_1^{(i)}} y^{s_2^{(i)}}$, such that

$$s_1^{(1)} > s_1^{(2)} > \cdots > s_1^{(l)} = 0 \quad \text{and}$$

$$0 = s_2^{(1)} < s_2^{(2)} < \cdots < s_2^{(l)} \tag{3.4}$$

and, if we define

$$\Delta = \{(h, r) \mid h < s_1^{(i)} \text{ and } r < s_2^{(i+1)}, \quad \text{for some } i,$$

where $1 \leq i \leq l-1\}$, then

no proper polynomial in $\mathscr{L}$ has leading

$$\text{term with exponent in } \Delta. \tag{3.5}$$

Fig. 1 illustrates the concepts.In order to make the paper selfcontained we will prove the following theorem.

Theorem 1: Let $F = \{\sigma^{(1)}, \cdots, \sigma^{(l)}\}$ be a minimal basis for $\mathscr{L}$ and let Δ be the set defined above. Then $|\Delta| = t$ and a polynomial $\sigma'(x, y)$ belongs to $\mathscr{L}$, if and only if

$$\sum_{l,k} \sigma'_{l,k} S_{a+l, b+k} = 0, \qquad \text{for } (a, b) \in \Delta. \tag{3.6}$$

Proof: We will first prove that the $|\Delta|$ arrays $T^{(a,b)}$, $(a, b) \in \Delta$, where $T_{l,k}^{(a,b)} = S_{a+l, b+k}$, are linearly independent elements of $\mathscr{S}$. They are elements of $\mathscr{S}$ since $T^{(a,b)} = T^f$, where $f = (e_i x_i^a y_i^b)$ $i \in I$. To see that they are linearly independent, suppose

$$\sum_{(a,b) \in \Delta} \lambda_{ab} T^{(a,b)} = 0.$$

This means that for all (l, k) we have

$$\sum_{(a,b) \in \Delta} \lambda_{ab} S_{a+l, b+k} = 0$$

and, therefore, that the polynomial

$$\Lambda(x, y) = \sum_{(a,b) \in \Delta} \lambda_{ab} x^a y^b$$

is in $\mathscr{L}$, so by (3.5) we conclude that $\lambda_{ab} = 0$. In particular, we get that $|\Delta| \leq t$. We will next prove that $|\Delta| \geq t$ and therefore we get $|\Delta| = t$ and then the t arrays $T^{(a,b)}$, $(a, b) \in \Delta$, form a basis for $\mathscr{S}$. Therefore, a polynomial $\sigma'(x, y)$ belongs to $\mathscr{L}$, if and only if (3.6) is satisfied. To prove that $|\Delta| \geq t$, we will prove the following lemma.

Lemma 1: For each $(\alpha, \beta) \notin \Delta$, there exists a polynomial $P^{\alpha,\beta}(x, y)$, whose leading term has exponent in Δ, such that $x^\alpha y^\beta + P^{\alpha\beta}(x, y)$ is in $\mathscr{L}$ and has leading term $x^\alpha y^\beta$.

Before proving this, we note that it then follows that these $(q-1)^2 - |\Delta|$ polynomials are linearly independent and therefore $(q-1)^2 - |\Delta| \leq$ dimension of $\mathscr{L} = (q-1)^2 - t$, so $t \leq |\Delta|$. The lemma is now proved by induction. So let (α_0, β_0) be the smallest pair not in Δ, with respect to the total ordering $<_T$. Then $x^{\alpha_0} y^{\beta_0}$ is the leading term of the $\sigma^{(i)}$ whose leading term has the smallest exponent, so $\sigma^{(i)} - x^{\alpha_0} y^{\beta_0}$ has all exponents in Δ, which proves the claim in this case.

Let $(\alpha, \beta) \notin \Delta$ and suppose the claim is true for all smaller exponents. In particular, it is true for either $(\alpha - 1, \beta)$ or $(\alpha, \beta - 1)$, since $(\alpha, \beta) \notin (0, 0)$. Suppose that it is true for $(\alpha - 1, \beta)$. Then $f(x, y) = x^{\alpha-1} y^\beta - P^{\alpha-1,\beta}(x, y) \in \mathscr{L}$

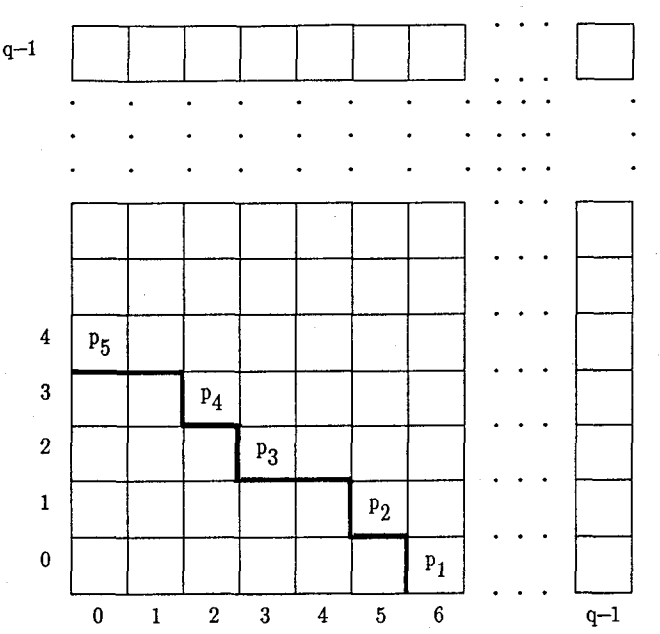

Fig. 1. Leading terms of polynomials in the minimal basis, $t = 16$, p_i is the leading term of $\sigma^{(i)}$, e.g., $p_3 = (3, 2)$.

and $P^{(\alpha-1,\beta)}(x, y)$ has leading term with exponents in Δ. Therefore, $xf(x, y) \in \mathcal{L}$ and $xf(x, y) = x^{\alpha}y^{\beta} + x \cdot P^{\alpha-1,\beta}(x, y)$. Here, either $x \cdot P^{\alpha-1,\beta}(x, y)$ has leading term with exponent in Δ, or the leading term has an exponent $<_T (\alpha, \beta)$. In the latter case, we subtract from $xf(x, y)$, for all terms with exponents not in Δ, the polynomials of the form $x^{\alpha_1}y^{\beta_1} + P^{\alpha_1,\beta_1}(x, y)$, which exist by the induction hypothesis, and this gives us the desired result. We emphasize that it follows from the theorem, that if we want to determine if a polynomial σ' belongs to $\mathcal{L}$ then it suffices to check the t conditions corresponding to (3.6).

In the following, we will by an *error locator polynomial* mean a polynomial from $\mathcal{L}$ that do not have $C(x, y)$ as a factor. The polynomial $\sigma^{l}(x, y)$ has leading term with exponent $(0, s_2^{(l)})$, and here $s_2^{(l)} \leq m$. If $s_2^{(l)} = m$ we will use as $\sigma^{l}(x, y)$ the polynomial $C(x, y)$, since we have assumed that $C(x, y)$ contains the term y^m. The other polynomials in the minimal basis can therefore not have $C(x, y)$ as a factor.

As mentioned before, we distinguish between known and unknown syndromes. By the degree of a syndrome $S_{l,k}$, we mean the number $l + k$. If we take a certain polynomial $\sigma^{(s)}$ from the minimal basis, only syndromes up to degree j_s are involved in (3.6), where

$$j_s = \max_i \{\deg \sigma^{(i)}\} - 1 + \deg \sigma^{(s)}. \tag{3.7}$$

We say that $\sigma^{(s)}$ can be determined from the known syndromes if $j \geq j_s$.

In the following, we shall discuss two problems.

1) Up to which degree shall the syndromes be known in order that all error locator polynomials can be determined from the known syndromes?
2) Up to which degree shall the syndromes be known in order that the error locator polynomials of smallest degree can be determined from the known syndromes?

To carry out this discussion we shall first establish some bounds on the degrees of the error locator polynomials. It follows from the theorem of Bezout that

$$\deg \sigma^{(i)} \geq t/m. \tag{3.8}$$

Let $D(k)$ denote the number of polynomials from $\mathcal{L}$ of degree less than or equal to k, which are linearly independent modulo $C(x, y)$, that is linearly independent when considered as vectors in $F_q[x, y]/C(x, y)$ over F_q. Now a special case of the Riemann-Roch theorem [7] gives the following theorem.

Theorem 2:

$$D(k) \geq mk - g + 1 - t, \tag{3.9}$$

and (3.9) holds with equality provided

$$mk - t > 2(g - 1). \tag{3.10}$$

We define the numbers $a(k)$, $k \geq 1$, as

$$a(k) = D(k) - D(k - 1). \tag{3.11}$$

We always have $D(k) \leq D(k - 1) + m$, since the number of polynomials linearly independent modulo $C(x, y)$, with no further restrictions, has this property [1, Theorem 1]. Therefore,

$$a(k) \leq m. \tag{3.12}$$

We can now prove Lemma 2.

Lemma 2: If $a(k) = m$, then $a(l) = m$ for $l \geq k$.

Proof: Since the polynomials are reduced modulo $C(x, y)$, which contains the term y^m, the m polynomials of degree k have leading terms among x^k, $x^{k-1}y$, $\cdots, x^{k+1-m}y^{m-1}$. From these we get by multiplication with x, m polynomials of degree $k + 1$ and these are also linearly independent modulo $C(x, y)$. So $a(k + 1) = m$ and by repeating the argument the lemma follows. □

From Lemma 2 we get the following.

Lemma 3: If $a(k) = m$, then

$$\max_s \{\deg \sigma^{(s)}\} \leq k. \tag{3.13}$$

Proof: It follows from (3.12) and Lemma 2 that $D(k + l) = D(k) + lm$, and from the proof of Lemma 2 follows then that there exists an ideal basis for $\mathcal{L}$ consisting of polynomials of degrees less than or equal to k. From this a minimal basis can be obtained [8], in which all polynomials have degrees less than or equal to k. □

Now let k_o be the smallest number, such that (3.10) is satisfied, that is the smallest number such that $k_o > t/m + m - 3$. It follows from Theorem 1 that $a(k_o + 1) = m$, and therefore from Lemma 3, we get

$$\max_s \{\deg \sigma^{(s)}\} \leq t/m + m - 1. \tag{3.14}$$

We can now answer the first question as follows.

Theorem 3: All error locators can be determined from the known syndromes if

$$j \geq 2(t/m + m - 1) - 1,$$

or, equivalently,

$$t \leq d^*/2 - m^2/2.$$

Proof: From (3.14) and (3.7), we have $j_s \leq 2(t/m + m - 1) - 1$ for each s and hence, the theorem. □

In order to answer Question 2 we have to use more algebraic geometry. The main idea in the following argument is implicit in the proof of Theorem 8 in [2]. So let us consider the projective plane over GF(q) and let C' be the projective closure of C. Let H be the intersection divisor of the curve C', with the line with equation $z = 0$ with respect to the homogenous coordinates $(x : y : z)$. With the usual notation in algebraic geometry we then have $l(kH - \sum P_i) = D(k)$, where $D(k)$ is the dimension introduced earlier, and $P_1, \cdots, P_t$ are the error points. Suppose now that $D(p - 1) = 0$ and $D(p) > 0$, and that $a(p + s) < m$. Then the divisor $(p + s)H - \sum P_i$ is special and hence equivalent to $K - J$, where K is a canonical divisor, and J is effective. We have

$$\deg J = -(p + s)m + t + 2g - 2. \tag{3.15}$$

From the equivalence, it follows that $(p-1)H - \sum P_i$ is equivalent to $K - J - (s+1)H$. Since $l((p-1)H - \sum P_i) = 0$, we, therefore, have

$$l\big(K - (s+1)H\big) \leq \deg J. \tag{3.16}$$

From the Riemann–Roch Theorem, we have

$$l\big(K - (s+1)H\big) = l\big((s+1)H\big) - (s+1)m + g - 1,$$

and combining this with (3.15) and (3.16) we get

$$l\big((s+1)H\big) + (p-1)m + 1 \leq t + g. \tag{3.17}$$

Now let s_o be the smallest number such that (3.17) is *not* satisfied. Then $a(p+s) = m$ and it follows from Lemma 3 that $\max_i(\text{degree } \sigma^{(i)}) \leq p + s_o$. An upper bound on the minimal j_s from (3.7) can, therefore, be found as the maximal value of $2p+s$, where p and s are connected by the equation $l((s+1)H) + (p-1)m + 1 = t + g$.

Now $l((s+1)H) = (s+1)m - g + 1$ if $s+1 \geq m$ and $l((s+1)H) = \frac{1}{2}(s+2)(s+3)$ if $s+1 \leq m-1$. Carrying out the calculations it turns out that the maximal value is obtained in the second case and more precisely for $s = m/2 - 5/2$. The maximal value of $2p+s$ is therefore $2t/m + 5m/4 + 1/4m - 7/2$. We formulate the result in the next theorem.

Theorem 4: The error locator of lowest degree can be determined from the known syndromes if

$$j > 2t/m + 5m/4 + 1/4m - 7/2, \tag{3.18}$$

or, equivalently,

$$t < d^*/2 - m^2/8 + m/4 - 1/8. \tag{3.19}$$

At this point we will explain the consequences of Theorem 4.

In the decoding situation the only thing that we know are the syndromes S_{ab}, $a + b \leq j$, but we do not know the set Δ. So we consider all equations of the form

$$\sum_{l,k} \sigma'_{l,k} S_{a+l,\, b+k} = 0, \tag{3.20}$$

which only involve the known syndromes.

Let $\hat{\sigma}(x, y)$ be the solution to (3.20), which has leading term with smallest exponent, and does not have $C(x, y)$ as a factor. Furthermore, let $\sigma^{(s)}(x, y) \neq C(x, y)$ be a polynomial with lowest degree in the minimal basis for $\mathcal{L}$. In particular $\sigma^{(s)}(x, y)$ satisfy (3.20) so the degree of $\hat{\sigma}(x, y)$ is smaller than or equal to the degree of $\sigma^{(s)}(x, y)$. If the condition (3.18) is satisfied then Theorem 4 tells us that the equations for $\sigma^{(s)}(x, y)$ involved in (3.6) are a subset of the equations involved in (3.10). Consequently, this is also true for the polynomial $\hat{\sigma}(x, y)$ and therefore $\hat{\sigma}(x, y)$ is an error locator. We remark that the procedure previously described is basically the same as the modified algorithm of [2]. In the next section we discuss how to find $\hat{\sigma}(x, y)$.

IV. Calculation of Error Locators

The error locator polynomials defined in Section III may be obtained as solutions to the system of linear equations (3.20). However, the complexity of this approach is high and the algorithm is not practical for some of the most interesting geometric codes (e.g., Hermitian codes with large q and moderate rate).

We shall assume that the reader is familiar with the algorithms for correcting errors in BCH codes, as the present decoding problem may be interpreted as a generalization to two dimensions of the decoding of BCH codes. In one dimension the Berlekamp–Massey algorithm [9] and many later improvements allow decoding of t errors with complexity $O(t^2)$ or less. Sakata [5] has generalized the Berlekamp–Massey algorithm to two dimensions, and we shall present a modified version of this algorithm, which computes the error locators using at most $(6 + A)m^2j^3$ GF(q) additions and multiplications, where A is the number of terms in $C(x, y)$.

Sakata's algorithm was developed for calculating the recursions consistent with a given two dimensional array. The input in our situation is the array of syndromes S_{ab} $a + b \leq j$, $b < 2m$. A step in the algorithm consists of reading the next element, with respect to the total ordering $<_T$, and then finding a minimal set of recursions for the array of elements read so far. A minimal set is expressed as a set of polynomials $F = \{f^{(1)}, f^{(2)}, \cdots, f^{(j)}\}$ for which the leading terms satisfy (3.4), and a condition like (3.5) holds when $\mathcal{L}$ is substituted by the set of valid recursions for the array at this step. At each step the current set of F polynomials are tested on the new array, and if some of the $f^{(i)}$'s are not consistent, they are updated.

In the one-dimensional Berlekamp–Massey algorithm, the recursion $f(z)$ is updated by means of a polynomial $g(z)$ that has failed at an earlier point. When for some input $f(z)$ is not satisfied, a multiple of $g(z)$ is added.

$$g(z) \leftarrow f(z)$$
$$f(z) \leftarrow z^s f(z) + a g(z)$$

In Sakata's algorithm the set F is updated by a set of polynomials $G = \{g^{(1)}, g^{(2)}, \cdots\}$ that have failed at earlier points in the algorithm. The polynomials in G satisfies a condition like (3.4) and the set is updated during the algorithm. For details of the algorithm the reader is referred to [5].

In our situation we are not interested in polynomials which have $C(x, y)$ as a factor, and, therefore, we reduce the polynomials in F modulo $C(x, y)$, so that the reduced polynomial has y degree less than m, when this is possible. Since $C(x, y)$ contains the term y^m, this reduction has the following consequences.

1) There are at most m polynomials in the sets F and G, since a condition like (3.4) is satisfied.
2) When the input is S_{ab} it follows from [5, Section 5] that no polynomial in F (and hence, in G) have leading terms greater than (a, b) in the total ordering, so at this step the polynomials in F and G have at most mj terms.

To get the complexity of the algorithm we count the number of GF(q) multiplications and additions. After each new input element we have to 1) check whether the polyno-

mials in F still are valid, 2) update F and G and 3) reduce the elements in the new F modulo $C(x, y)$.

Since the polynomials in F have at most mj terms the cost of checking one polynomial in a point is at most mj multiplications and additions. Since there are at most m polynomials in F and at most $2mj$ elements in the input array the total
cost of 1) is at most $m \cdot 2mj \cdot mj = 2m^3j^2$ additions and multiplications.

It follows directly from [5, Section 5], that the updating of the sets F and G costs at most $4m^3j^2$ additions and multiplications, and that a new F element has the form

$$f'(x, y) = \alpha x^{i_1} y^{j_1} f(x, y) + \beta x^{i_2} y^{j_2} g(x, y), \quad (4.1)$$

where $j_1 < m$ and $j_2 < m$, $\alpha, \beta \in \mathrm{GF}(q)$ and $f(x, y)$ is from the old F and $g(x, y)$ is from the old G. The degree of $f'(x, y)$ is at most $2m - 1$ in the variable y and the total degree is at most j.

The reduction of $f'(x, y)$ modulo $C(x, y)$ proceeds through decreasing powers of y. First the terms $y^{2m-1} \cdot x^i$ are eliminated by addition of a suitable multiple of $C(x, y)$, then the terms $y^{2m-2} \cdot x^i$, etc. Each term requires A additions and multiplications, where A is the number of terms in $C(x, y)$, and since there are at most mj terms, the total number of operations is at most $A \cdot mj$. The reduction of at most m polynomials at most mj times therefore costs at most Am^3j^2 operations.

The total cost of using the modified version of Sakata's algorithm is therefore bounded by $(6 + A)m^3j^2$ additions and multiplications in $\mathrm{GF}(q)$. If the number t of errors satisfies the condition from Theorem 3, that is $t \le d^*/2 - m^2/2$, then the set of syndromes S_{ab}, $a + b \le j$, $b < 2m$ is so large that the polynomials in F all satisfies (3.6) and hence, it follows from Theorem 3 that the set F is a minimal basis for $\mathscr{L}$, that is they are error locators. If the number t of errors satisfy the condition from Theorem 4, that is, $t \le d^*/2 - m^2/8$, then the set of input syndromes is so large that a polynomial in $F \setminus C(x, y)$ with the lowest degree satisfy (3.6) and, therefore, by Theorem 4, these polynomials are error locators.

We note that Sakata [10] has generalized his algorithm to higher dimensions. In principle this algorithm could be applied to a wider class of codes.

V. Determination of the Error Values

The fast determination of the error values presupposes knowledge of all syndromes $S_{a,b}$ $a \le q - 1$, $b \le q - 1$. We will first show how the error values are found from these (this is an extension of the familiar transformation method) and then describe the method to obtain the unknown syndromes from the curve, the locator polynomial and the known syndromes.

Let us suppose we have S_{ab}, $a \le q - 1$, $b \le q - 1$, and the possible error points (x_i, y_i), $i = 1, 2, \cdots, s$, where $s \geqq t$. For a fixed c, where $0 < c \le q - 1$, we define

$$\tilde{S}_a(c) = \sum_{b=1}^{q-1} S_{ab}(\alpha^c)^b, \quad (5.1)$$

where α is a primitive element of the field $\mathrm{GF}(q)$.

Substitution of $S_{ab} = \sum_{i=1}^{s} e_i x_i^a y_i^b$ into (5.1) gives

$$\tilde{S}_a(c) = \sum_{b=1}^{q-1} \sum_{i=1}^{s} e_i x_i^a (y_i \alpha^c)^b = \sum_{i=1}^{s} e_i x_i^a \sum_{b=1}^{q-1} (y_i \alpha^c)^b.$$

So we have

$$\tilde{S}_a(c) = - \sum_{\text{those } i\text{'s where } y_i = \alpha^{-c}} e_i x_i^a. \quad (5.2)$$

If we now define for a fixed d, where $0 < d \le q - 1$,

$$E_{c,d} = \sum_{a=1}^{q-1} \tilde{S}_a(c)(\alpha^d)^a, \quad (5.3)$$

we have

$$E_{c,d} = \sum_{a=1}^{q-1} \sum_{b=1}^{q-1} S_{ab}(\alpha^c)^b (\alpha^d)^a = e_i,$$

where

$$(x_i, y_i) = (\alpha^{-d}, \alpha^{-c}). \quad (5.4)$$

Now the error values can be determined directly from (5.4). The calculation of $E_{c,d}$ costs at most q^2 additions and multiplications using a procedure like Horners method, so the cost of finding the error values using (5.4) is at most sq^2. Alternatively, we calculate the $\tilde{S}_a(c)$'s using (5.1). This calculation needs only to be done when α^{-c} is the second coordinate of an error point, but there may be $q - 1$ of those. However, if we use (5.2) to find the error values, we use at most m values of a, so using a fast transform on (5.1), the calculation of the $\tilde{S}_a(c)$'s costs at most $C_1 \cdot mq \log q$ operations. To find the error values from (5.2), we can either use Forney's algorithm or a simple matrix inversion. This can be done at a cost of $C_2 m^2 q$ operations, so the cost of finding the error values using (5.1) and (5.2) is bounded by $(C_1 \cdot mq \log q + C_2 m^2 q)$ operations.

We will now turn our attention to the determination of the syndromes S_{ab}, $a \le q - 1$, $b \le q - 1$, from the known syndromes S_{ab}, $a + b \le j$. The basic observation is (again) that since from (2.2) we have

$$S_{ab} = \sum_{i \in I} e_i x_i^a y_i^b,$$

any polynomial $f(x, y)$, which have the points (x_i, y_i) $i \in I$ among its zeros, gives a recursion among the syndromes. In particular, the polynomials $C(x, y)$ and $\sigma(x, y)$ give such recursions, as well as do all polynomials in the ideal in $F[x, y]$ generated by these two.

The idea is now to use these two recursions to generate the remaining syndromes from the given ones. To this end, let

$$C(x, y) = \sum_{l+k \le m} c_{lk} x^l y^k$$

and

$$\sigma(x, y) = \sum_{l+k \le h} \sigma_{lk} x^l y^k.$$

We then have the following recursions:

$$\sum_{l+k\le m} c_{lk}S_{a+l,\,b+k} = 0,$$

$$\sum_{l+k\le h} \sigma_{lk}S_{a+l,\,b+k} = 0,$$

which we write as

$$c_{mo}S_{a+m,\,b} + c_{m-1,1}S_{a+m-1,\,b+1} + \cdots + c_{om}S_{a,\,b+m} = -\sum_{l+k\le m-1} c_{lk}S_{a+l,\,b+k} \quad (5.5)$$

and

$$\sigma_{ho}S_{a+h,\,b} + \sigma_{h-1,1}S_{a+h-1,\,b+1} + \cdots + \sigma_{oh}S_{a,\,b+h} = -\sum_{l+k\le h-1} \sigma_{lk}S_{a+l,\,b+k}. \quad (5.6)$$

Now, suppose we know S_{ab}, $a + b \le j + \alpha - 1$, where $\alpha \ge 1$, then by putting $b = 0, 1, \cdots, j + \alpha - m$, and $a = j + \alpha - b - m$ in (5.5) and (5.6) we get the following system of linear equations

$$\begin{matrix} j+\alpha+1-m \\ \text{rows} \\ \\ j+\alpha+1-h \\ \text{rows} \end{matrix} \begin{bmatrix} c_{mo} & c_{m-11} & \cdots & c_{om} & & \\ & c_{mo} & \cdots & & c_{om} & \underline{0} \\ \underline{0} & & \ddots & & & \ddots \\ & & & c_{mo} & & c_{om} \\ \sigma_{ho} & \sigma_{h-11} & \cdots & \sigma_{oh} & & \\ & \sigma_{ho} & \cdots & & \sigma_{oh} & \underline{0} \\ \underline{0} & & \ddots & & & \\ & & & \sigma_{ho} & & \sigma_{oh} \end{bmatrix} \begin{bmatrix} S_{j+\alpha\,o} \\ S_{j+\alpha-11} \\ \\ S_{oj+\alpha} \end{bmatrix} = v, \quad (5.7)$$

where v only depends on $S_{a,b}$, where $a + b \le j + \alpha - 1$ and the coefficients c_{lk} with $l + k \le m - 1$ and σ_{lk} with $l + k \le h - 1$.

To ensure that the system (5.7) has at most one solution a sufficient condition is that the rank of the coefficient matrix is $j + \alpha + 1$, and a necessary condition for this is that $2(j + \alpha + 1) - (m + h) \ge j + \alpha + 1$, that is, $j + \alpha + 1 \ge m + h$, so if we let $\alpha = 1$ we must have $j \ge m + h - 2$.

The condition $j \ge m + h - 2$ is always satisfied. In the case corresponding to Theorem 3, we have $j \ge 2(t/m + m - 1) - 1$ and (3.14) gives an upper bound on h. In the case corresponding to Theorem 4, we have $t \le mj/2 - 5m^2/8 + 3m/2$ and, since h is upper bounded by the p appearing in the proof of the theorem, $h \le j/2 - m/4 + 2$. Hence, $m + h - 2 \le j/2 + 3m/4$, so if $j \ge 3m/2$ we have $j \ge m + h - 2$. If $j < 3m/2$, we have $t \le 1/8m^2 + 3m/2$, and therefore, there are error locators of degree less than m. The degree h is then upper bounded by the smallest number such that $(h + 2)(h + 1)/2 \ge mj/2 - 5m^2/8 + 3m/2$. By putting $j = (1 + x)m$, $0 \le x < 1/2$ and carrying out the calculation, it follows that also in this case we have $j \ge m + h - 2$.

When $j \ge m + h - 2$, then it is well known, [7, p. 29], that the coefficient matrix has rank $j + \alpha + 1$, if and only if the two polynomials

$$\tilde{c}(x, y) = c_{mo}x^m + c_{m-1,1}x^{m-1}y + \cdots + c_{om}y^m$$

and

$$\tilde{\sigma}(x, y) = \sigma_{ho}x^h + \sigma_{h-1,1}x^{h-1}y + \cdots + \sigma_{oh}y^h$$

do not have a common nonconstant factor.

This is in particular the case when $\tilde{\sigma}(x, y) = \sigma_{ho}x^h$, $\sigma_{ho} \ne 0$, and $c_{om} \ne 0$, and then the solution of (5.7) is easily obtained. From the rows of the lower part of the matrix, we calculate $S_{j+\alpha,\,o}, S_{j+\alpha-1,\,1}, \cdots, S_{h,\,j+\alpha-h}$ and then, from rows of the upper part of the matrix, we get $S_{h-1,\,j+\alpha-h-1}, \cdots, S_{o,j+h}$.

We formulate the above result as follows.

Theorem 5: Suppose

$$C(x, y) = \sum_{l+k\le m} c_{lk}x^l y^k,$$

where

$$c_{om} \ne 0 \quad \text{and} \quad \sigma(x, y) = \sum_{l+k\le h-1} \sigma_{lk}x^l y^k + \sigma_{ho}x^h,$$

where $\sigma_{ho} \ne 0$, both gives recursions among the S_{ab}'s. Then all S_{ab}'s can be determined from S_{ab}, $a + b \le j$, by the method just described.

The cost of finding the S_{ab}'s of a given degree is at most hm multiplications and additions for at most $(j + \alpha - h)$ of these, and for the remaining ones it is at most m^2. The total cost is therefore bounded by m^2q^2. If we combine this with the remarks following (5.2), we have the following theorem.

Theorem 6: If the curve has an equation of the form

$$C(x, y) = \sum_{l+k\le m} c_{lk}x^l y^k,$$

where $c_{om} \ne 0$, and we have an error locator of the form

$$\sigma(x, y) = \sum_{l+k\le h-1} \sigma_{lk}x^l y^k + \sigma_{ho}x^h, \qquad \sigma_{ho} \ne 0,$$

then the error values of the code $C^*(j)$ can be found using at most Am^2q^2 additions and multiplications in GF(q), where A is a constant (independent of $C(x, y)$, j and q).

We will now treat the codes from Hermitian curves. It turns out, that due to the special equation, it is possible to determine the remaining syndromes from the given ones, without any conditions on the error locator polynomials. We will prove the following theorem.

Theorem 7: Let $\sigma(x, y)$ be an error locator of degree h for a code $C^*(j)$ from the Hermitian curve

$$x + x^r - y^{r+1} = 0, \qquad q = r^2.$$

The syndromes S_{ab}, where $0 \le a < q-1$, $0 \le b < q-1$ can be determined from S_{ab}, where $a + b \le j$, using at most Aq^3 additions and multiplications in GF(q), where A is a constant independent of q and j.

Proof: We will describe a method that also proves the theorem. We first use the curve to obtain the following recursion:

$$S_{a,b+r+1} = S_{a+1,b} + S_{a+r,b}. \tag{5.8}$$

So from S_{ab}, where $a + b \le j + \alpha - 1$, we find

$$S_{o,j+\alpha}, S_{1,j+\alpha-1}, \cdots, S_{j+\alpha-(1+r),r+1}$$

by putting $a = 0, 1, \cdots, j + \alpha - (1 + r)$, and $b = j + \alpha - (r + 1) - a$. To find more syndromes we use the polynomial $\sigma(x, y)$.

Let

$$\sigma(x,y) = \sum_{l+k\le h-1} \sigma_{lk} x^l y^k + \alpha_0 x^s y^{h-s} + \alpha_1 x^{s-1} y^{h-s+1} + \cdots + \alpha_s y^h,$$

where $\alpha_0 \neq 0$. The corresponding recursion is

$$\sum_{l+k\le h-1} \sigma_{l,k} S_{a+l,b+k} + \alpha_0 S_{a+s,b+h-s} + \cdots + \alpha_s S_{a,b+h} = 0,$$

from which we determine

$$S_{j+\alpha-r,r}, \cdots, S_{j+\alpha-h+s,h-s}$$

by putting $a = j + \alpha - r - s, \cdots, j + \alpha - h$ and $b = j + \alpha - h - a$, so if $s = h$ we have got all the syndromes, but if $s < h$ we must do a little more. We can suppose that $h - s < r + 1$, since higher powers in y can be removed by using the equation of the curve.

Let us next consider the polynomial

$$\begin{aligned}\sigma_1(x,y) &= -(y^{r+1} - x^r - x)(\alpha_o x^s + \cdots + \alpha_s y^s) \\ &\quad + y^{r+1-h+s}\sigma(x,y) \\ &= (x^r + x)(\alpha_o x^s + \cdots + \alpha_s y^s) \\ &\quad + y^{r+1-h+s} \sum_{l+k\le h-1} \sigma_{lk} x^l y^k.\end{aligned}$$

This polynomial also gives a recursion, namely,

$$\begin{aligned}&\alpha_o S_{a+r+s,b} + \cdots + \alpha_s S_{a+r,b+s} \\ &\quad + \alpha_o S_{a+s+1,b} + \cdots + \alpha_s S_{a+1,b+s} \\ &\quad + \sum_{l+k\le h-1} \sigma_{lk} S_{a+l,b+r+1-h+s+k} = 0.\end{aligned}$$

From this we determine

$$S_{j+\alpha-h+s+1,h-s-1}, \cdots, S_{j+\alpha,o}$$

by putting $a = j + \alpha - h + 1, \cdots, j + \alpha - r - s$ and $b = j + \alpha - r - s - a$.

What remains is to calculate the number of GF(q) additions and multiplications used in the method described above. Let α be fixed. The cost of finding the first $j + \alpha - (r + 1) + 1$ syndromes is $2(j + \alpha - r)$ additions. The cost of finding the remaining $r + 1$ syndromes is at most $(r + 1)[(r + 1)h]$ additions and multiplications, but since we can suppose that $h < 2q$, since we are only interested in GF(q) points, the total cost is upper bounded by $B \cdot q^3$, where B is a constant independent of q and j.

We have seen that the number of GF(q) additions and multiplications needed to find the error locator is bounded by $(6 + A)m^3 j^2$, where A is the number of terms in $C(x, y)$. We have also seen that we need at most $B \cdot m^2 q^2$ additions and multiplications to find the error values.

Since we are looking only at plane curves the number n of points is bounded by q^2, so the only way to get longer codes is to increase the field size. Good codes are obtained when the curve has many rational points, that is, $n \sim m^2\sqrt{q}$, which follows from the Weil bound. Moreover, since a curve of degree m has at most $m\,q$ rational points, we see that good codes are obtained if $m \sim \sqrt{q}$ and $n \sim q\sqrt{q}$. If we, therefore, consider a family of curves for increasing q, which satisfies the above conditions, and for which the number of terms do not depend on q, we see that the complexity of the decoding algorithm is $O(n^{7/3})$, where $n \sim q^{3/2}$, $m \sim q^{1/2}$, since $mj \le n$.

VI. Conclusion and Discussion

We have used a modified version of Sakata's generalization of the Berlekamp–Massey algorithm to find error locator polynomials for codes from regular plane curves. This is one with complexity $O(n^{7/3})$, and we correct $d^*/2 - m^2/8 + m/4 - 9/8$ errors. Moreover, we can find the error values with the same complexity in most cases. Examples show that it is not always possible to decode $(d^* - 1)/2$ errors, but the examples seem to be rare and to have a common feature, so in most cases the algorithm does decode $(d^* - 1)/2$ errors and the information in the examples where this is not the case, may be useful in an improvement of the algorithm. We feel confident that the algorithm can be generalized to higher dimensions, but we still lack good explicit constructions of curves in higher dimensional spaces.

References

[1] J. Justesen, K. J. Larsen, A. Havemose, H. E. Jensen, and T. Høholdt, "Construction and decoding of a class of algebraic geometry codes," *IEEE Trans. Inform. Theory*, vol. 35, pp. 811–821, July 1989.

[2] A. N. Skorobogatov and S. G. Vlăduţ, "On the decoding of algebraic-geometric codes," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1051-1061, Sept. 1990.

[3] R. Pellikaan, "On a decoding algorithm for codes on maximal curves," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1228-1232, Nov. 1989.

[4] S. G. Vlăduţ, "On the decoding of algebraic-geometric codes over $\mathbb{F}_q$ for $q \geq 16$," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1461-1463, Nov. 1990.

[5] S. Sakata, "Finding a minimal set of linear recurring relations capable of generating a given finite two-dimensional array," *J. Symbolic Computat.*, vol. 5, pp. 321-337, 1988.

[6] H. Stichtenoth, "A note on Hermitian codes over GF(q^2)," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1345-1348, Sept. 1988.

[7] R. J. Walker, *Algebraic Curves*. New York: Dover, 1962.

[8] S. Sakata, "On determining the independent point set for doubly periodic arrays and encoding two-dimensional cyclic codes and their duals," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 556-565, Sept. 1981.

[9] J. L. Massey, "Shift-register synthesis and BCH decoding," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 122-127, Jan. 1969.

[10] S. Sakata, "Extension of the Berlekamp-Massey algorithm to N dimensions," *Inform. Computat.*, vol. 84, no. 2, pp. 207-239, Feb. 1990.