Computational and Applied Mathematics
Vol. 22, N. 2, pp. 233–247, 2003

Alternant and BCH codes over certain rings

A.A. ANDRADE[1], J.C. INTERLANDO[1] and R. PALAZZO JR.[2]

[1]Department of Mathematics, Ibilce, Unesp, 15054-000 São José do Rio Preto, SP, Brazil
E-mail: andrade@mat.ibilce.unesp.br / E-mail: carmelo@mat.ibilce.unesp.br
[2]Department of Telematics, Feec, Unicamp, P.O. Box 6101, 13081-970 Campinas, SP, Brazil
E-mail: palazzo@dt.fee.unicamp.br

Abstract. Alternant codes over arbitrary finite commutative local rings with identity are constructed in terms of parity-check matrices. The derivation is based on the factorization of $x^s - 1$ over the unit group of an appropriate extension of the finite ring. An efficient decoding procedure which makes use of the modified Berlekamp-Massey algorithm to correct errors and erasures is presented. Furthermore, we address the construction of BCH codes over $\mathbb{Z}_m$ under Lee metric.

Mathematical subject classification: 11T71, 94B05, 94B40.

Key words: codes over rings, alternant codes, BCH codes, Galois extensions of local commutative rings, algebraic decoding, modified Berlekamp-Massey algorithm, errors and erasures decoding, Lee metric.

1 Introduction

Alternant codes form a large and powerful family of codes. They can be obtained by a simple modification of the parity-check matrix of a BCH code. The most famous subclasses of alternant codes are BCH codes and (classical) Goppa codes, the former for their simple and easily instrumented decoding algorithm, and the later for meeting the Gilbert-Varshamov bound. However, most of the work regarding construction and decoding of alternant codes has been done considering codes over finite fields. On the other hand, linear codes over integer rings have recently generated a great deal of interest because of their new role in algebraic coding theory and their successful application in combined coding and modulation. A remarkable paper by Hammons et al. [1] has shown that certain binary

#542/01. Received: 7/III/01.

nonlinear codes with good error correcting capabilities can be viewed through a Gray mapping as linear codes over $\mathbb{Z}_4$. Moreover, Calderbank et al. [2] studied cyclic codes over $\mathbb{Z}_4$. Viewing many BCM (block coded modulation) schemes as group block codes over groups, in [3] it was shown that group block codes over abelian groups can be studied via linear codes over finite rings. Andrade and Palazzo [4] constructed BCH codes over finite commutative rings with identity. Also, Greferath and Vellbinger [5] have investigated codes over integer residue rings under the aspect of decoding. The Lee metric ([6], [7]) was developed as an alternative to the Hamming metric for transmission of nonbinary signals over certain noisy channels. Roth and Siegel [8] have constructed and decoded BCH codes over $GF(p)$ under Lee metric.

In this paper we address the problems of constructing and decoding alternant codes over arbitrary finite commutative local rings with identity and the problems of construction of BCH codes for the Lee metric. The core of the construction technique mimics that of alternant and BCH codes over a finite field, and is based on the factorization of $x^s - 1$ over an appropriate extension ring. The decoder is capable of handling both errors and erasures, which enables the implementation of generalized minimum distance decoding (GMD) to further reduce the probability of decoding error [9].

This paper is organized as follows. In Section 2, we describe a construction of alternant codes over a finite commutative local ring with identity and an efficient decoding procedure is proposed. We show how this decoding procedure can also be used to handle erasures. In Section 3, we describe a construction of BCH codes over $\mathbb{Z}_q$, where q is a prime power, under Lee metric. The question of the existence of a simple decoding algorithm for these codes remains open.

2 Alternant Code

In this section we present a construction technique of alternant codes over finite commutative local rings with identity, in terms of parity-check matrices. First we collect basic definitions and facts from the Galois theory of commutative rings, which are necessary to characterize such matrices. Throughout this section we assume that $\mathcal{A}$ is a finite commutative local ring with identity, maximal ideal $\mathcal{M}$ and residue field $\mathbb{K} = \dfrac{\mathcal{A}}{\mathcal{M}} \cong GF(p^m)$, where m is a positive integer and p is a

prime. Let $f(x)$ be a monic polynomial of degree h in $\mathcal{A}[x]$, such that $\mu(f(x))$ is irreducible in $\mathbb{K}[x]$, where μ is the natural projection. Then by [10, Theorem XIII.7(a)], we have $f(x)$ also irreducible in $\mathcal{A}[x]$. Let $\mathcal{R}$ be the ring $\frac{\mathcal{A}[x]}{\langle f(x)\rangle}$. Then $\mathcal{R}$ is a finite commutative local ring with identity and is called a Galois extension of $\mathcal{A}$ of degree h. Its residue field is $\mathbb{K}_1 = \frac{\mathcal{R}}{\overline{\mathcal{M}_1}} \cong GF(p^{mh})$, where $\overline{\mathcal{M}_1}$ is the maximal ideal. We have that $\mathbb{K}_1^*$ is the multiplicative group of $\mathbb{K}_1$, whose order is $p^{mh} - 1$.

Let $\mathcal{R}^*$ denotes the multiplicative group of units of $\mathcal{R}$. It follows that $\mathcal{R}^*$ is an abelian group, and therefore it can be expressed as a direct product of cyclic groups. We are interested in the maximal cyclic group of $\mathcal{R}^*$, hereafter denoted by $\mathcal{G}_s$, whose elements are the roots of $x^s - 1$ for some positive integer s such that $\gcd(s, p) = 1$. From [10, Theorem XVIII.2], there is only one maximal cyclic subgroup of $\mathcal{R}^*$ having order relatively prime to p. This cyclic group has order $s = p^{mh} - 1$.

Definition 2.1. *Let $\boldsymbol{\eta} = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ be the* locator vector*, consisting of distinct elements of $\mathcal{G}_s$, and let $\boldsymbol{w} = (w_1, w_2, \cdots, w_n)$ be an arbitrary vector consisting of elements of $\mathcal{G}_s$. Define the matrix H by*

$$H = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \\ w_1\alpha_1 & w_2\alpha_2 & \cdots & w_n\alpha_n \\ w_1\alpha_1^2 & w_2\alpha_2^2 & \cdots & w_n\alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ w_1\alpha_1^{r-1} & w_2\alpha_2^{r-1} & \cdots & w_n\alpha_n^{r-1} \end{bmatrix},$$

where r is a positive integer. Then H is the parity-check matrix of a shortened alternant code $\mathcal{C}(n, \boldsymbol{\eta}, \boldsymbol{w})$ *of length $n \leq s$ over $\mathcal{A}$.* □

It is possible to obtain an estimate of the minimum Hamming distance d of $\mathcal{C}(n, \boldsymbol{\eta}, \boldsymbol{w})$ directly from the parity-check matrix. The next theorem provides such an estimate.

Lemma 2.1. *Let α be an element of $\mathcal{G}_s$ of order s. Then the difference $\alpha^{l_1} - \alpha^{l_2}$ is a unit in $\mathcal{R}$ if $0 \leq l_1 < l_2 \leq s-1$.*

Proof. It is sufficient to show that $1 - \alpha^j$, $j = 1, 2, \cdots, s-1$ is a unit. If $1 - \alpha^j \in \overline{\mathcal{M}_1}$ for some $1 \leq j \leq s-1$, it follows that $\alpha^j = 1$, which is a contradiction. □

Theorem 2.1. *$\mathcal{C}(n, \boldsymbol{\eta}, \boldsymbol{w})$ has minimum Hamming distance $d \geq r+1$.*

Proof. Suppose $\boldsymbol{c}$ is a nonzero codeword in $\mathcal{C}(n, \boldsymbol{\eta}, \boldsymbol{w})$, such that the weight $w_H(\boldsymbol{c}) \leq r$. Then, $\boldsymbol{c}H^T = 0$. Deleting $n-r$ columns of the matrix H corresponding to zeros of the codeword, it follows that the new matrix H' is Vandermonde, and therefore its determinant is a unit in $\mathcal{R}$. Thus, the only possibility for $\boldsymbol{c}$ is the all-zero codeword. □

Example 2.1. The polynomial $f(x) = x^3 + x + 1$ is irreducible over $\mathbb{Z}_2$, and over the commutative local ring $\mathcal{A} = \mathbb{Z}_2[i]$, where $i^2 = -1$. Thus $\mathcal{R} = \dfrac{\mathcal{A}[x]}{\langle f(x) \rangle}$ is a Galois extension of $\mathcal{A}$. Let α be a root of $f(x)$. We have that α generates acyclic group $\mathcal{G}_s$ of order $s = 2^3 - 1 = 7$ in $\mathcal{R}^*$. Setting $\boldsymbol{\eta} = (1, \alpha, \cdots, \alpha^6)$ and $\boldsymbol{w} = (1, 1, 1, 1, 1, 1, 1)$, if $r = 2$ then we have an alternant code $\mathcal{C}(7, \boldsymbol{\eta}, \boldsymbol{w})$ over $\mathbb{Z}_2[i]$ with minimum Hamming distance of at least 3. □

2.1 Decoding Procedure

This section is devoted to developing a decoding method for an alternant code as defined in the previous section. Let $\mathcal{C}(n, \boldsymbol{\eta}, \boldsymbol{w})$ be an alternant code with minimum Hamming distance at least $r+1$, i.e., this code can correct up to $t = [(r+1)/2]$ errors, where $[n]$ denotes the largest integer less than or equal to n. Then $t = (r+1)/2$ when r is odd, and $t = r/2$ when r is even. The idea is to extend efficient standard decoding procedures for BCH codes which work well over fields (as described, for example, in [12], [13], [14], and [15]) to finite commutative local rings with identity. Note that these afore mentioned decoding procedures do not work over rings, in general. As an example, the

original Berlekamp-Massey algorithm [12], [16], which is fundamental in the decoding process of a BCH code, cannot be applied directly if the elements of the sequence to be generated do not lie in a field.

First, we establish some notation. Let $\mathcal{R}$ denotes the ring defined in Section 2 and α be a primitive element of $\mathcal{G}_s$. Let $\boldsymbol{c} = (c_1, c_2, \cdots, c_n)$ be the transmitted codeword and $\boldsymbol{r} = (r_1, r_2, \cdots, r_n)$ be the received vector. The error vector is given by $\boldsymbol{e} = (e_1, e_2, \cdots, e_n) = \boldsymbol{r} - \boldsymbol{c}$. Given a locator vector $\boldsymbol{\eta} = (\alpha_1, \cdots, \alpha_n) = (\alpha^{k_1}, \cdots, \alpha^{k_n})$ in $\mathcal{G}_s^n$, we define the *syndrome values* $s_\ell \in \mathcal{G}_s$, of an error vector $\boldsymbol{e} = (e_1, \cdots, e_n)$, as

$$s_\ell = \sum_{j=1}^{n} e_j w_j \alpha_j^\ell, \quad \ell \geq 0.$$

Suppose that $\nu \leq t$ is the number of errors which occurred at locations $x_1 = \alpha_{i_1}, \cdots, x_\nu = \alpha_{i_\nu}$, with values $y_1 = e_{i_1}, \cdots, y_\nu = e_{i_\nu}$. Since $\boldsymbol{s} = \boldsymbol{r}H^T = \boldsymbol{e}H^T$, where $\boldsymbol{s} = (s_0, \cdots, s_{r-1})$, the first r syndrome values s_ℓ can be calculated from the received vector $\boldsymbol{r}$ as $s_\ell = \sum_{j=1}^{n} e_j w_j \alpha_j^\ell = \sum_{j=1}^{n} r_j w_j \alpha_j^\ell$, $\ell = 0, 1, \cdots, r-1$. The elementary symmetric functions $\sigma_1, \sigma_2, \cdots, \sigma_\nu$ of the error-location numbers $x_1, x_2, \cdots, x_\nu$ are defined as the coefficients of the polynomial

$$\sigma(x) = \prod_{i=1}^{\nu}(x - x_i) = \sum_{i=0}^{\nu} \sigma_i x^{\nu - i},$$

where $\sigma_0 = 1$. Thus, the decoding procedure being proposed consists of four major steps [11]:

Step 1 – Calculation of the syndrome vector from the received vector;

Step 2 – Calculation of the elementary symmetric functions $\sigma_1, \sigma_2, \cdots, \sigma_\nu$ from $\boldsymbol{s}$ using the modified Berlekamp-Massey algorithm [11];

Step 3 – Calculation of the error-location numbers $x_1, x_2, \cdots, x_\nu$ from $\sigma_1, \sigma_2, \cdots, \sigma_\nu$ that are roots of $\sigma(x)$;

Step 4 – Calculation of the error magnitudes $y_1, y_2, \cdots, y_\nu$ from x_i and $\boldsymbol{s}$ by Froney's procedure [13].

Next we analyze each step of the decoding algorithm in some detail. Since calculation of the syndromes is straightforward, we will not make any comments on *Step* 1.

The set of possible error-location numbers is a subset of $\{\alpha^0, \alpha^1, \cdots, \alpha^{s-1}\}$. The elementary symmetric functions $\sigma_1, \sigma_2, \cdots, \sigma_\nu$ (where ν denotes the number of errors introduced by the channel) are defined as the coefficients of the polynomial $(x - x_1)(x - x_2)\cdots(x - x_\nu) = x^\nu + \sigma_1 x^{\nu-1} + \cdots + \sigma_{\nu-1}x + \sigma_\nu$. In *Step* 2, the calculation of the elementary symmetric functions is equivalent to finding a solution $\sigma_1, \sigma_2, \cdots, \sigma_\nu$, with minimum possible ν, to the following set of linear recurrent equations over $\mathcal{R}$

$$\begin{gathered} s_{j+\nu} + s_{j+\nu-1}\sigma_1 + \cdots + s_{j+1}\sigma_{\nu-1} + s_j\sigma_\nu = 0, \\ j = 0, 1, 2, \cdots, (r-1) - \nu, \end{gathered} \tag{1}$$

where $s_0, s_1, \cdots, s_{r-1}$ are the components of the syndrome vector. We make use of the modified Berlekamp-Massey algorithm [11] to find the solutions of Eq. (1), that hold for commutative rings with identity. We call attention to the fact that in rings care must be taken regarding zero divisors, multiple solutions of the system of linear equations, and also with an inversionless implementation of the original Berlekamp-Massey algorithm. The algorithm is iterative, in the sense that the following $n - l_n$ equations (called *power sums*)

$$\begin{cases} s_n\sigma_0^{(n)} + s_{n-1}\sigma_1^{(n)} + \cdots + s_{n-l_n}\sigma_{l_n}^{(n)} = 0 \\ s_{n-1}\sigma_0^{(n)} + s_{n-2}\sigma_1^{(n)} + \cdots + s_{n-l_n-1}\sigma_{l_n}^{(n)} = 0 \\ \quad\vdots \\ s_{l_n+1}\sigma_0^{(n)} + s_{l_n}\sigma_1^{(n)} + \cdots + s_1\sigma_{l_n}^{(n)} = 0 \end{cases}$$

are satisfied with l_n as small as possible and $\sigma^{(0)} = 1$. The polynomial $\sigma^{(n)}(x) = \sigma_0^{(n)} + \sigma_1^{(n)}x + \cdots + \sigma_{l_n}^{(n)}x^n$ represents the solution at the n-th stage. The n-th *discrepancy* will be denoted by d_n and defined by $d_n = s_n\sigma_0^{(n)} + s_{n-1}\sigma_1^{(n)} + \cdots + s_{n-l_n}\sigma_{l_n}^{(n)}$. The modified Berlekamp-Massey algorithm for commutative rings with identity is formulated as follows: The inputs to the algorithm are the syndromes $s_0, s_1, \cdots, s_{r-1}$ which belong to $\mathcal{R}$. The output of the algorithm is a set of values σ_i, $1 \leq i \leq \nu$, such that the equations in Eq. (1) hold with minimum ν. In order to start the algorithm, set the initial conditions: $\sigma^{(-1)}(x) = 1, l_{-1} = 0, d_{-1} = 1, \sigma^{(0)}(x) = 1, l_0 = 0$, and $d_0 = s_0$ [15]. Thus, we have the following steps:

1) $n \leftarrow 0$.

2) If $d_n = 0$, then $\sigma^{(n+1)}(x) \leftarrow \sigma^{(n)}(x)$, and $l_{n+1} \leftarrow l_n$, and go to 5).

3) If $d_n \neq 0$, then find an $m \leq n-1$ such that $d_n - yd_m = 0$ has a solution in y and $m - l_m$ has the largest value. Then, $\sigma^{(n+1)}(x) \leftarrow \sigma^{(n)}(x) - yx^{n-m}\sigma^{(m)}(x)$, and $l_{n+1} \leftarrow \max\{l_n, l_m + n - m\}$.

4) If $l_{n+1} = \max\{l_n, n+1-l_n\}$ then go to 5), else search for a solution $D^{(n+1)}(x)$ with minimum degree l in the range $\max\{l_n, n+1-l_n\} \leq l < l_{n+1}$ such that $\sigma^{(m)}(x)$ defined by $D^{(n+1)}(x) - \sigma^{(n)}(x) = x^{n-m}\sigma^{(m)}(x)$ is a solution for the first m power sums, $d_m = -d_n$, and with $\sigma_0^{(m)}$ a zero divisor in $\mathcal{R}$. If such a solution is found, $\sigma^{(n+1)}(x) \leftarrow D^{(n+1)}(x)$, and $l_{n+1} \leftarrow l$.

5) If $n < r-1$, then $d_{n+1} \leftarrow s_{n+1} + s_n\sigma_1^{(n+1)} + \cdots + s_{n+1-l_{n+1}}\sigma_{l_{n+1}}^{(n+1)}$.

6) $n \leftarrow n+1$; if $n < r$ go to 2), else stop.

The coefficients $\sigma_1^{(r)}, \sigma_2^{(r)}, \cdots, \sigma_\nu^{(r)}$ satisfy the equations in Eq. (1). The basic difference between the modified Berlekamp-Massey algorithm and the original one lies in the fact that the modified algorithm allows updating a minimal polynomial solution $\sigma^{(n)}(x)$ (at the n-th step) from a previous solution $\sigma^{(m)}(x)$, whose discrepancy can even be a noninvertible element in the commutative ring under consideration. This process does not necessarily lead to a minimal solution $\sigma^{(n+1)}(x)$ (at the $(n+1)$-th stage). So, Step 4, calculated at Step 3, is checked to be a minimal solution. This search consists of finding a polynomial $\sigma^{(m)}(x)$, satisfying certain conditions, and being a solution for the first m power sums. Since the number of polynomials $\sigma^{(m)}(x)$ to be checked is not too large, Step 4 does not essentially increase the complexity.

In Step 3, the calculation of error location numbers over rings requires one more step than over fields, because in $\mathcal{R}$ the solution of Eq. (1) is generally not unique and the reciprocal of the polynomial $\sigma^{(r)}(z)$ (output by the modified Berlekamp-Massey algorithm), namely $\rho(z)$, may not be the right error locator polynomial $(z - x_1)(z - x_2)\cdots(z - x_\nu)$ where $x_i = \alpha^{k_j}$ (j is an integer in the range $1 \leq j \leq n$ such that k_j indicates the position of the error in the codeword) are the correct error-location numbers, ν is the number of errors, and α is the generator of $\mathcal{G}_s$. Thus, the procedure for the calculation of the correct error-location numbers [11] is given by

- Compute the roots of $\rho(z)$ (the reciprocal of $\sigma^{(r)}(z)$), say, $z_1, z_2, \cdots, z_\nu$,
- Among the $x_i = \alpha^{k_j},\ j = 1, 2 \cdots, n$, select those x_i's such that $x_i - z_i$ are zero divisors in $\mathcal{R}$. The selected x_i's will be the correct error-location numbers and $k_j,\ j = 1, 2, \cdots, n$, indicates the position of the error in the codeword.

In *Step* 4, the calculation of the error magnitudes is based on Forney's procedure [13]. The error magnitudes $y_1, y_2, \cdots, y_\nu$ are given by

$$y_j = \frac{\sum_{\ell=0}^{\nu-1} \sigma_{j\ell} s_{\nu-1-\ell}}{E_j \sum_{l=0}^{\nu-1} \sigma_{j\ell} x_j^{\nu-1-\ell}}, \qquad j = 1, 2, \cdots, \nu, \tag{2}$$

where the coefficients $\sigma_{j\ell}$ are recursively defined by $\sigma_{j,i} = \sigma_i + x_j \sigma_{j,i-1},\ i = 0, 1, \cdots, \nu - 1$, starting with $\sigma_{j,0} = \sigma_0 = 1$. The $E_j = w_{i_j},\ j = 1, 2, \cdots, \nu$ are the corresponding location of errors in the vector $\boldsymbol{w}$. Again making use of Lemma 2.1, it can be shown that the denominator in Eq. (2) is always a unit in $\mathcal{R}$.

Example 2.2. Let $\mathcal{G}_7$ be the cyclic group as in Example 2.1. Considering $\boldsymbol{\eta} = (\alpha^5, \alpha, \alpha^4, \alpha^2) = (\alpha^{k_1}, \cdots, \alpha^{k_4})$, $\boldsymbol{w} = (\alpha^4, \alpha, \alpha^4, \alpha)$ and $r = 2$, we have an alternant code over $\mathbb{Z}_2[i]$ of length 4 and minimum Hamming distance at least 3. Let H be the parity-check matrix. Assume that the all-zero codeword $\boldsymbol{c} = (0, 0, 0, 0)$ is transmitted, and the vector $\boldsymbol{r} = (0, 0, i, 0)$ is received. Then the syndrome vector is $\boldsymbol{s} = \boldsymbol{r}H^T = (i\alpha^4, i\alpha)$. By the modified Berlekamp-Massey algorithm we obtain $\sigma^{(2)}(z) = 1 + \alpha^4 z$. The root of $\rho(z) = z + \alpha^4$ (the reciprocal of $\sigma^{(2)}(z)$) is $z_1 = \alpha^4$. Among the elements $\alpha^0, \cdots, \alpha^6$, we have that $x_1 = \alpha^4$ satisfies $x_1 - z_1 = 0$ (zero divisor in $\mathcal{R}$). Therefore, x_1 is the correct error-location number since $k_3 = 4$ indicates that one error has occurred in the third coordinate of the codeword. The correct elementary symmetric function $\sigma_1 = \alpha^4$ is obtained from $x - x_1 = x - \sigma_1 = x - \alpha^4$. Finally, applying Forney's method to $\boldsymbol{s}$ and σ_1, gives $y_1 = i$. Therefore, the error pattern is $\boldsymbol{e} = (0, 0, i, 0)$. □

2.2 *Error-and-Erasure Decoding*

In this subsection we briefly discuss how the decoding procedure for alternant codes can be used to correct errors and erasures. The development is based in [15, pp. 305-307]. We know that if the minimum distance d of a code $\mathcal{C}$ satisfies $d \geq 2t + e + 1$, then e erasures and up to t errors can be corrected by $\mathcal{C}$. Suppose that $v \leq t$ errors occur in positions $x_1 = \alpha_{k_1}, x_2 = \alpha_{k_2}, \cdots, x_v = \alpha_{k_v}$, with respective nonzero magnitudes $y_1, y_2, \cdots, y_v$. Suppose further that e erasures occur in positions $u_1 = \alpha_{\ell_1}, u_2 = \alpha_{\ell_2}, \cdots, u_e = \alpha_{\ell_e}$, with respective magnitudes $v_1, v_2, \cdots, v_e$. Note that whereas e and the u_i are known to the decoder, the v_i are not. The syndrome of a received vector $\boldsymbol{r}$ is given by

$$s_j = \sum_{i=1}^{v} y_i' x_i^j + \sum_{p=1}^{e} v_p' u_p^j ; \qquad 0 \leq j \leq r-1, \tag{3}$$

where $y_i' = y_i w_{k_i}$, and $v_p' = v_p w_{\ell_p}$. Defining the elementary symmetric functions τ_k of the known erasure locations by $\prod_{i=1}^{e}(u - u_i) = \sum_{k=0}^{e}(-1)^k \tau_k u^{e-k}$, and the modified syndromes t_j by $t_j = \sum_{k=0}^{e}(-1)^k \tau_k s_{j-k}$; $j = e, e+1, \ldots, r-1$, it can be shown that

$$\begin{aligned} t_j &= \sum_{i=1}^{v} f_i x_i^j ; \qquad j = e, e+1, \ldots, r-1, \\ \text{where } f_i &= y_i x_i^{-e} \sum_{k=0}^{e} (-1)^k \tau_k x_i^{e-k}. \end{aligned} \tag{4}$$

Since $x_i \neq 0$, and $x_i \neq u_i$, it follows that $f_i \neq 0$. The equations in Eq. (4) can be efficiently solved for the x_i using the modified Berlekamp-Massey algorithm. We call attention to the fact that the first value assumed by the exponent j is e, instead of 0, as before. Now, with the x_i known, all we need to complete the decoding process is to solve the equations in Eq. (3) in order to find y_i and v_i. To this end, Forney's procedure can be applied again as in *Step 4* of the decoding procedure for the alternant codes.

Example 2.3. Let $\mathcal{R} = \mathbb{Z}_4[x]/(x^3 + x + 1)$. The element $\alpha = x^2 = (0, 0, 1)$ generates $\mathcal{G}_7$. Considering $\boldsymbol{\eta} = (\alpha^2, \alpha, \alpha^3, \alpha^5, \alpha^4, \alpha^6) = (\alpha^{k_1}, \cdots, \alpha^{k_6})$, $\boldsymbol{w} = (1, 1, 1, 1, 1, 1)$ and $r = 4$, we have an alternant code over $\mathbb{Z}_4$ of length 6 and

minimum Hamming distance at least 5. This code can correct 1 error and 2 erasures. Assume that the all-zero codeword $\boldsymbol{c} = (0, 0, 0, 0, 0, 0)$ is transmitted, and the vector $\boldsymbol{r} = (2, 0, ?, 0, 0, ?)$ is received, where ''?'' denotes an erasure. Note that the erasures in $\mathbf{r}$ can take values on $\mathbb{Z}_4$. For example, we can ''guess'' that the erasure in the third coordinate is a 3, and the erasure in the sixth coordinate is a 2. Thus the received vector is $\boldsymbol{r} = (2, 0, 3, 0, 0, 2)$. It follows that the components of the syndrome vector are $s_0 = (3, 0, 0), s_1 = (1, 2, 3), s_2 = (1, 1, 3)$, and $s_3 = (2, 1, 3)$. From the equation $(u - \alpha^3)(u - \alpha^6) = \tau_0 u^2 + \tau_1 u + \tau_2$ we obtain the elementary symmetric functions τ_k of the known erasures locations, that is, $\tau_0 = (1, 0, 0)$, $\tau_1 = (2, 3, 2)$, and $\tau_2 = (0, 3, 3)$. The modified syndromes are therefore $t_2 = (0, 2, 0)$ and $t_3 = (0, 0, 2)$. Applying the modified Berlekamp-Massey algorithm to the sequence $\{t_2, t_3\}$, we obtain $\sigma(z) = 1 + (0, 1, 0)z$. The root of $\rho(z) = z + (0, 1, 0)$ is $z_1 = (0, 3, 0)$. Among the elements $\alpha^0, \cdots, \alpha^6$, we have that $\alpha^4 = (2, 1, 0)$ satisfies $\alpha^4 - z_1 = 0$ (zero divisor in $\mathcal{R}$). Therefore, $x_1 = \alpha^{k_1} = \alpha^2$ is the correct error-location number, since $e = 2$. It indicates that one error has occurred in the first position of the codeword. Forney's procedure applied to $\tau_1 = \alpha^2$ gives $y_1 = 2$, $v_1 = 3$, and $v_2 = 2$. Therefore, the error pattern is $\boldsymbol{e} = \boldsymbol{r}$. □

3 BCH code

In this section we present a construction technique of BCH codes over commutative ring of integers modulo q, where q is a prime power, in terms of parity-check matrices under Lee metric based in the work by Roth and Siegel [8]. First we collect basic definitions and facts from Lee metric over $\mathbb{Z}_m$, where m is a positive integer.

Definition 3.1. *Let $\mathbb{Z}_m$ denotes the commutative ring of integers modulo m, where m is a positive integer.*

- *The Lee value of an element $\alpha \in \mathbb{Z}_m$ is defined by*

$$| \alpha | = \begin{cases} \alpha, & \text{for} \quad 0 \leq \alpha \leq \left[\frac{m}{2}\right], \\ m - \alpha, & \text{for} \quad \left[\frac{m}{2}\right] < \alpha \leq m - 1. \end{cases},$$

where $\left[\frac{m}{2}\right]$ *is the greatest integer smaller than or equal to* $\frac{m}{2}$.

- *The Lee distance between two vectors* $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ *and* $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ *over* $\mathbb{Z}_m$ *is defined by*

$$d_L(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{n} d_L(a_i, b_i),$$

where $d_L(a_i, b_i) = min\{a_i - b_i, b_i - a_i\}(mod\ m),\ i = 1, 2, \ldots, n.$

- *The Lee weight of a vector* $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ *over* $\mathbb{Z}_m$ *is defined by*

$$w_L(\mathbf{a}) = d_L(\mathbf{a}, \mathbf{0}) = \sum_{i=1}^{n} \mid a_i \mid .$$

- *The minimum Lee distance,* $d_L(X)$*, of a subset* X *of* $\mathbb{Z}_m^n$ *is the minimum Lee distance between any pair of distinct vectors in* X. □

Remark 3.1.

- The elements $0, 1, 2, \ldots, \left[\frac{m}{2}\right]$ of $\mathbb{Z}_m$ are defined as the *positive elements*. The rest of the elements are the *negative* ones [8].

- The *minimum Lee distance* of a code is defined as the minimum Lee distance between all pairs of codewords. For linear codes, the difference of any two codewords is also a codeword. Thus, the minimum Lee distance of a linear code is equal to the minimum Lee weight of its nonzero codewords.

- The minimum Lee distance of a code is greater than or equal to the minimum Hamming distance of the same code, and smaller than or equal to the Lee distance between the two codewords which define the minimum Hamming distance. Thus

$$d_H \leq d_L \leq \left[\frac{m}{2}\right] d_H.$$

- The Lee distance defines a metric over $\mathbb{Z}_m$.

- For $m = 2$ and 3, Lee and Hamming distance coincide. For $m > 3$, the Lee distance between two n-tuples is greater than or equal to the Hamming distance between them. □

Let $\mathbb{Z}_q[x]$ denotes the ring of polynomials in the variable x over $\mathbb{Z}_q$, where q is a prime power p. Let $p(x)$ be a monic polynomial of degree h, irreducible over $\mathbb{Z}_p$. We have that $f(x)$ is also irreducible over $\mathbb{Z}_q$. Let $\mathcal{R} = \frac{\mathbb{Z}_m[x]}{\langle f(x)\rangle}$ denotes the set of residue classes of polynomials in x over $\mathbb{Z}_q$, modulo the polynomial $f(x)$. This ring is a local commutative with identity and is called a Galois extension of $\mathbb{Z}_q$ of degree h. Let $\mathcal{G}_s$, where $s = p^h - 1$, be the maximal cyclic subgroup of $\mathcal{R}^*$ such that $gcd(s, p) = 1$.

Definition 3.2. [8] *Let $\eta = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ be the locator vector consisting of distinct elements of $\mathcal{G}_s$. Now define matrix H as*

$$H = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{r-1} & \alpha_2^{r-1} & \cdots & \alpha_n^{r-1} \end{bmatrix},$$

where r is a positive integer. Then H is the parity-check matrix of a shortened BCH code $\mathcal{C}(n, \eta)$ of length $n \leq s$ over $\mathbb{Z}_q$. □

Thus, a word $\mathbf{c} = (c_1, c_2, \cdots, c_n) \in \mathbb{Z}_q^n$ is in $\mathcal{C}(n, \eta)$ if and only if it satisfies the following r parity-check equations over $\mathcal{R}$:

$$\sum_{j=1}^{n} c_j \alpha_j^l = 0, \qquad l = 0, 1, \ldots, r-1.$$

The codes $\mathcal{C}(n, \eta)$ for which $n = p^h - 1$ will be called *primitive*. In this case, η is unique, up to permutation of coordinates.

Given a transmitted word $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in \mathcal{C}(n, \eta)$ and a received word $\mathbf{b} \in \mathbb{Z}_q^n$, the error vector is defined by $\mathbf{e} = \mathbf{b} - \mathbf{c}$. The number of *Lee errors* is given by $w_L(\mathbf{e})$, that is, the number of Lee errors is the smallest number of additions of ± 1 to the coordenates of the transmitted codeword $\mathbf{c}$ which yields

the received word $\mathbf{b}$. Since the Lee weight satisfies the triangle inequality, using a code of minimum Lee distance d_L allows one to correct any pattern of up to $(d_L - 1)/2$ Lee errors.

Given a locator vector $\eta = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ of a code $\mathcal{C}(n, \eta)$ and a word $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{Z}_q^n$, we define the *locator polynomial* associated with $\mathbf{b}$ as the polynomial

$$\sigma(x) = \prod_{j=1}^{n} (1 - \alpha_j x)^{w_L(c_j)}.$$

We define the syndrome values s_l of an error vector $\mathbf{e} = (e_1, e_2, \ldots, e_n)$ by

$$s_l = \sum_{j=1}^{n} e_j \alpha_j^l, \qquad l \geq 0.$$

The formal syndrome series $S(x)$ is defined by

$$s(x) = \sum_{j=1}^{\infty} s_l x^l.$$

Given a codeword $\mathbf{c} \in \mathcal{C}(n, \eta)$, following the approach in Roth and Siegel [8], we define the word $\mathbf{c}^+ = (c_1^+, c_2^+, \ldots, c_n^+)$ by

$$c_j^+ = \begin{cases} c_j & \text{if } c_j \in \left\{1, 2, \ldots, \left[\frac{q}{2}\right]\right\} \\ 0 & \text{otherwise,} \end{cases}$$

and let $\mathbf{c}^- = \mathbf{c}^+ - \mathbf{c}$. That is, $\mathbf{c}^+$ is equal to $\mathbf{c}$ at the late positive entries, and is zero otherwise, whereas the entries of $\mathbf{c}^-$ take the Lee values of the negative entries of $\mathbf{c}$, leaving the other locations zero.

In the next Proposition, a lowerbound for the minimum Lee distance is obtained when $w_L(\mathbf{c}^+) \neq w_L(\mathbf{c}^-)$.

Proposition 3.1. [8] *If* $\mathbf{c} \in \mathcal{C}(n, \eta)$ *and* $w_L(\mathbf{c}^+) \neq w_L(\mathbf{c}^-)$ *then* $w_L(\mathbf{c}) \geq q$.

Proof. Since $\mathbf{c}H^T = 0$, we have that $\mathbf{c}^+ H^T = \mathbf{c}^- H^T$. The first equation in this last equality reads $w_L(\mathbf{c}^+) = w_L(\mathbf{c}^-)(mod\ q)$, that is, $w_L(\mathbf{c}^+) = w_L(\mathbf{c}^-) \pm lq$, for some integer l. Therefore, $w_L(\mathbf{c}) = w_L(\mathbf{c}^+) + w_L(\mathbf{c}^-) \geq q$. □

Example 3.1. The polynomial $f(x) = x^3 + x + 1$ is irreducible over $\mathbb{Z}_4$. Thus the finite commutative ring $\mathcal{R} = \frac{\mathbb{Z}[x]}{\langle f(x)\rangle}$ is a Galois extension of $\mathbb{Z}_4$. Let α be a root of $f(x)$. We have that $\beta = \alpha^8$ generates a cyclic group $\mathcal{G}_s$ of order $s = 2^3 - 1 = 7$ in $\mathcal{R}^*$. Letting $\eta = (1, \beta, \beta^2, \beta^3, \beta^4, \beta^5, \beta^6)$ and $r = 2$, we have an BCH code $\mathcal{C}(7, \eta)$ over $\mathbb{Z}_4$. Let

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \beta & \beta^2 & \beta^3 & \beta^4 & \beta^5 & \beta^6 \end{bmatrix}$$

be the parity-check matrix. We have that $\mathbf{c} = (3102101) \in \mathcal{C}(7, \eta)$ and $w_L(\mathbf{c}^+) = 5$, $w_L(\mathbf{c}^-) = 1$ and $w_L(\mathbf{c}) = 6 > 4$. □

4 Acknowledgments

The authors would like to thank the referees for theirs helpful suggestions and comments which improved the presentation of this paper.

REFERENCES

[1] A.R. Hammons Jr., P.V. Kumar, A.R. Calderbank, N.J.A. Sloane and P. Solé, *The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals and related codes*, IEEE Trans. Inform. Theory, IT-40 (1994), pp. 301–319.

[2] A.R. Calderbank, G. McGuire, P.V. Kumar and T. Helleseth, *Cyclic codes over $\mathbb{Z}_4$, locator polynomials, and Newton' identities*, IEEE Trans. on Inform. Theory, **42** (1996), pp. 217–226.

[3] E. Biglieri and M. Elia, *On the construction of group block codes*, Annales des Telecommunications, Tome 50, No. **9-10** (1995), pp. 817–823.

[4] A.A. Andrade and R. Palazzo Jr., *Construction and decoding of BCH codes over finite commutative rings*, Linear Algebra and Its Applications, **286** (1999) pp. 69–85.

[5] M. Greferath and U. Vellbinger, *Efficient decoding of $\mathbb{Z}_{p^k}$-linear codes*, IEEE Trans. Inform. Theory, **44** (1998), pp. 1288–1291.

[6] C.Y. Lee, *Some properties of nonbinary error-correcting codes*, IRE Trans. Inform. Theory, vol. **4**, no. 4 (1958), pp. 77–82.

[7] W. Ulrich, *Non-binary error correction codes*, Bell Sys. Tech. J., vol. **36**, no. 6 (1957), pp. 1341–1387.

[8] R.M. Roth and P.H. Siegel, *Lee-metric BCH codes and their application to constrained and partial-reponse channels*, IEEE Trans. Inform. Theory, vol. **40**, no. 4 (1994), pp. 1083–1096.

[9] G.D. Forney Jr., *Generalized minimum distance decoding*, IEEE Trans. Inform. Theory, IT-12 (1966) pp. 125–131.

[10] B.R. McDonald, *Finite Rings with Identity*, Marcel Dekker, Inc., New York (1974).

[11] J.C. Interlando, R. Palazzo Jr. and M. Elia, *On the decoding of Reed-Solomon and BCH codes over integer residue rings*, IEEE Trans. Inform. Theory, IT-43 (1997), pp. 1013–1021.

[12] E.R. Berlekamp, *Algebraic Coding Theory*, McGraw-Hill, New York, (1968).

[13] G.D. Forney Jr., *On decoding BCH codes*, IEEE Trans. Inform. Theory, IT-11 (1965), pp. 549–557.

[14] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error Correcting Codes*, North-Holland, Amsterdam (1977).

[15] W.W. Peterson and E.J. Weldon Jr., *Error Correcting Codes*, MIT Press, Cambridge, Mass., (1972).

[16] J.L. Massey, *Shift-register synthesis and BCH decoding*, IEEE Trans. Inform. Theory, IT-15 (1969), pp. 122–127.