TC10 / **5. Alternant codes**

S. Xambó

Definitions and examples

Let $F = \mathbb{F}_q$ and $\bar{F} = \mathbb{F}_{q^m}$. Let $\alpha_1, \dots, \alpha_n$ and $h_1, \dots, h_n$ be elements of $\bar{F}$ such that $\alpha_i \neq \alpha_j$ when $i \neq j$ and $h_i \neq 0$ for all i. Let us consider the matrix

$$H = V_r(\alpha_1, \dots, \alpha_n) \cdot \operatorname{diag}(h_1, \dots, h_n)$$

$$= \begin{pmatrix} h_1 & \cdots & h_n \\ h_1\alpha_1 & \cdots & h_n\alpha_n \\ \vdots & & \vdots \\ h_1\alpha_1^{r-1} & \cdots & h_n\alpha_n^{r-1} \end{pmatrix} \in M_n^r(\bar{F}).$$

We say that H is the *alternant control matrix* of order r associated to the vectors $\boldsymbol{h} = (h_1, \dots, h_n)$ and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$.

We will also need the vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$, where $\beta_i = 1/\alpha_i$ (defined only if $\alpha_i \neq 0$ for all i).

The F-code defined by H is

$$A_F(\boldsymbol{h}, \boldsymbol{\alpha}, r) = \{x \in F^n \mid xH^T = 0\},$$

and the codes of this kind are called *alternant codes*.

If we define the *H-syndrome* s of a vector $y \in \bar{F}^n$ by $s = yH^T \in \bar{F}^r$, then $A_F(\boldsymbol{h}, \boldsymbol{\alpha}, r)$ is the subspace of F^n whose elements are the vectors with zero H-syndrome.

If $\boldsymbol{h} = \boldsymbol{\alpha}$, we will write $A_F(\boldsymbol{\alpha}, \boldsymbol{r})$ instead of $A_F(\boldsymbol{\alpha}, \boldsymbol{\alpha}, r)$. On the other hand $A(\boldsymbol{h}, \boldsymbol{\alpha}, r)$ or $A(\boldsymbol{\alpha}, r)$ are often used when $F = \mathbb{Z}_2$.

Proposition (Alternant bounds) If $C = A_F(\boldsymbol{h}, \boldsymbol{\alpha}, r)$, then

$$n - r \geq \dim(C) \geq n - rm \text{ i } d_C \geq r + 1.$$

Proof. Let H' be the $rm \times n$ matrix over F obtained after substituting each element of H by the column of its components with respect to a basis of $\bar{F}$ over F. Then C is also the code associated to the control matrix H'. Taking into account that the rank of H' over K is at most rm, it is clear that

$$\dim(C) \geq n - rm.$$

Now the sub-determinant of order r of H corresponding to the columns $i_1, \dots, i_r$ is equal to

$$\begin{vmatrix} h_{i_1} & \dots & h_{i_r} \\ h_{i_1}\alpha_{i_1} & \dots & h_{i_r}\alpha_{i_r} \\ \vdots & & \vdots \\ h_{i_1}\alpha_{i_1}^{r-1} & \dots & h_{i_r}\alpha_{i_r}^{r-1} \end{vmatrix} = h_{i_1} \cdots h_{i_r} \begin{vmatrix} 1 & \dots & 1 \\ \alpha_{i_1} & \dots & \alpha_{i_r} \\ \vdots & & \vdots \\ \alpha_{i_1}^{r-1} & \dots & \alpha_{i_r}^{r-1} \end{vmatrix}$$

$$= h_{i_1} \cdots h_{i_r} D(\alpha_{i_1}, \dots, \alpha_{i_r}) \neq 0$$

where $D(\alpha_{i_1}, \ldots, \alpha_{i_r})$ denotes the determinant of the Vandermonde matrix $V_r(\alpha_{i_1}, \ldots, \alpha_{i_r})$. This means that any r columns of H are linearly independent over $\bar{F}$ and consequently the minimum distance of C is at least $r + 1$. Finally, $\dim(C) \leq n + 1 - d_C$, by the Singleton bound, which together with $d_C \geq r + 1$ gives $\dim(C) \leq n - r$. □

Remark. The last inequality can be established directly. If we set

$$\bar{C} = \{\bar{x} \in \bar{F}^n \mid \bar{x}H^T = 0\},$$

then $\dim_{\bar{F}}(\bar{C}) = n - r$ and $C = \bar{C} \cap F^n \subseteq \bar{C}$. Hence

$$\dim_F(C) \leq \dim_{\bar{F}}(\bar{C}),$$

for linearly independent vectors of C over F are linearly independent over $\bar{F}$, and this shows that $\dim_K(C) \leq n - r$.

Example. Let $\alpha \in \mathbb{F}_8$ and suppose that $\alpha^3 = \alpha + 1$. Consider the matrix

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

and let C be the alternant binary code defined by H. Let us see that $C \sim [7,3,4]$, and hence that $d = 4 > 3 = r + 1$.

To begin with, the minimum distance d of C is ≥ 4, as any three columns of H are linearly independent over $\mathbb{F}_2$. On the other hand, the three first columns and the column of α^5 are linearly dependent, as we have $\alpha^5 = \alpha^2 + \alpha + 1$, and this gives $d = 4$. Finally the dimension of C is 3, because it has a control matrix of rank 4 over $\mathbb{F}_2$:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Examples

Reed–Solomon codes. Given distinct elements $\alpha_1, \dots, \alpha_n \in F$, we know that the Reed–Solomon code

$$C = RS_{\alpha_1,\dots,\alpha_n}(k) \subseteq F^n$$

has a control matrix of the form

$$H = V_{n-k}(\alpha_1, \dots, \alpha_n) \cdot \mathrm{diag}(h_1, \dots, h_n)$$

with $h_i = 1/\prod_{j \neq i}(\alpha_j - \alpha_i)$. Consequently

$$RS_{\alpha_1,\dots,\alpha_n}(k) = A_F(\boldsymbol{h}, \boldsymbol{\alpha}, n-k), \text{ with}$$
$$\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n), \ \boldsymbol{h} = (h_1, \dots, h_n).$$

Notice that in this case $\bar{F} = F$ ($m = 1$), and that the alternant bounds are exact, as we know that the minimum distance of C is $n - k + 1 = r + 1$, where r is the number of rows of H, and $k = n - r$.

Remark (Generalized Reed–Solomon codes). The vector $\boldsymbol{h}$ involved in the definition of the control matrix of the code $RS_{\alpha_1,\ldots,\alpha_n}(k)$ is a function of $\boldsymbol{\alpha}$. If we admit that $\boldsymbol{h}$ can be chosen independently of $\boldsymbol{\alpha}$, but with components in F, the codes obtained, $A_F(\boldsymbol{h}, \boldsymbol{\alpha}, n-k)$, are the *generalized Reed–Solomon* codes, and we write $GRS(\boldsymbol{h}, \boldsymbol{\alpha}, k)$ to denote them.

Note that we have, by definition of alternant codes, the following relation: If $\bar{F}$ is a finite field, r a positive integer and $\boldsymbol{h}, \boldsymbol{\alpha} \in \bar{F}^n$, then the linear code over $\bar{F}$ defined by the alternant control matrix H of order r associated to $\boldsymbol{h}$ and $\boldsymbol{\alpha}$ is the code $GRS(\boldsymbol{h}, \boldsymbol{\alpha}, n-r)$ and

$$A_F(\boldsymbol{h}, \boldsymbol{\alpha}, r) = GRS(\boldsymbol{h}, \boldsymbol{\alpha}, n-r) \cap F^n.$$

BCH codes. Let α be an element of $\bar{F}$, d a positive integer and l an integer. Let n be the order of α. Then we know that a control matrix of $C = BCH_\alpha(d, l)$ (the BCH code over F associated to α and with design *distance* d and *offset* l) is

$$H = \begin{pmatrix} 1 & \alpha^{l} & \alpha^{2l} & \cdots & \alpha^{(n-1)l} \\ 1 & \alpha^{l+1} & \alpha^{2(l+1)} & \cdots & \alpha^{(n-1)(l+1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha^{l+d-2} & \alpha^{2(l+d-2)} & \cdots & \alpha^{(n-1)(l+1)} \end{pmatrix}$$

which is the alternant control matrix of order $d-1$ associated to the vectors

$$\boldsymbol{h} = (1, \alpha^{l}, \alpha^{2l}, \ldots, \alpha^{(n-1)l}) \text{ and } \boldsymbol{\alpha} = (1, \alpha, \alpha^{2}, \ldots, \alpha^{n-1}).$$

The alternant bound on the minimum distance gives that the minimum distance is not less than $(d-1)+1 = d$, which is the BCH bound. The alternant bounds on the dimension coincide with the corresponding bounds for BCH codes.

Classical Goppa codes. Let $g \in \bar{F}[T]$ be a polynomial of degree $r > 0$. Let $\boldsymbol{\alpha} = \alpha_1, \ldots, \alpha_n \in \bar{F}$ be dinstinct elements such that $g(\alpha_i) \neq 0$ for all i. Then the *classical Goppa code* over F associated to g and $\boldsymbol{\alpha}$, which will be denoted $\Gamma(g, \boldsymbol{\alpha})$, can be defined as the code $A_F(\boldsymbol{h}, \boldsymbol{\alpha}, r)$ such that $\boldsymbol{h}$ is the vector $\big((1/g(\alpha_1), \ldots, 1/g(\alpha_n)\big)$. It is thus clear, by the alternant bounds, that the minimum distance of $\Gamma(g, \boldsymbol{\alpha})$ is $\geq r + 1$ and that its dimension k satisfies $n - rm \leq k \leq n - r$.

Proposition (The strict BCH codes are classical Goppa codes). Let ω be a primitive element of $\bar{F} = \mathbb{F}_{q^m}$ and let δ be an integer such that $2 \leq \delta \leq n$. Then $C = BCH_\omega(\delta)$ coincides with $C' = \Gamma\big(X^{\delta-1}, \boldsymbol{\alpha}\big)$, with $\boldsymbol{\alpha} = \big(1, \omega^{-1}, \ldots, \omega^{-(n-1)}\big)$.

Proof. Since the $\boldsymbol{h}$ vector of the control matrix H' of C' is $\big(1, \omega^{\delta-1}, \ldots, \omega^{(\delta-1)(n-1)}\big)$, the i-th row of H' is equal to $\big(1, \omega^{\delta-i}, \ldots, \omega^{(\delta-1)(n-i)}\big)$. Thus we see that H' is the control matrix H that defines C, but with the order of the rows reversed (note that the number of rows of H' is $\deg\big(X^{\delta-1}\big) = \delta - 1$. □

Localization and evaluation of errors. The key equation

Let $C \subseteq F^n$, $F = \mathbb{F}_q$, be the alternant code associated to the alternant matrix H of order r constructed with the vectors $\boldsymbol{h}$ and $\boldsymbol{\alpha}$ (their components are elements of $\bar{F} = \mathbb{F}_{q^m}$). Let $t = \lfloor r/2 \rfloor$, i.e., the greatest integer such that $2t \leq r$. Note that if we set $t' = \lceil r/2 \rceil$, then $t + t' = r$ (later we will use the equivalent equality $r - t = t'$).

Let $x \in C$ (*vector sent*) and $e \in F^n$ (*error vector*). Let $y = x + e$ (*vector received*). The goal of the decoders that we will study is to obtain x from y and H when $s = |e| \leq t$.

Let $M = \{m_1, \ldots, m_s\}$ be the set of error positions, i.e., $m \in M$ if and only if $e_m \neq 0$. Define *the error locators* η_i, $i = 1, \ldots, s$, by the relation $\eta_i = \alpha_{m_i}$. Since the α_j are distinct, the knowledge of the error locators is equivalent to the knowledge of the error positions.

Define the *syndrome* vector $S = (S_0, \dots, S_{r-1})$ by the formula $S = yH^T$. Note that $S = eH^T$, as $xH^T = 0$. Consider also the *polynomial syndrome*

$$S(z) = S_0 + S_1 z + \cdots + S_{r-1} z^{r-1}.$$

Since $S = 0$ is equivalent to say that y is a code vector (and therefore that $y = x$), from now on we assume that $S \neq 0$.

Remark. $S_j = \sum_{i=1}^{s} h_{m_i} e_{m_i} \eta_i^j$ $(0 \leq j \leq r-1)$.

Remark. The minimum j such that $S_j \neq 0$ satisfies $j < s$, and hence also $j < t$ (otherwise $h_{m_1} e_{m_1}, \dots, h_{m_s} e_{m_s}$ would satisfy s independent homogeneous linear equations). Since $\gcd(z^r, S(z)) = z^j$, the degree of $\gcd(z^r, S(z))$ is strictly less than s, and hence also strictly less than t. Similarly, $\deg(S(z)) \geq t$, as otherwise we would have $S_t = \cdots = S_{r-1} = 0$, and this also leads to a contradiction, because $r - t = t' \geq t \geq s$.

Remark. For the code $BCH_\omega(\delta, l)$ over $\mathbb{F}_q$, the syndromes $S_0, \dots, S_{\delta-2}$ are the values of the received polynomial (or also of the error

polynomial) on the elements $\omega^l, \dots, \omega^{l+\delta-2}$: $S_j = S(\omega^{l+j})$, $j = 0, \dots, \delta - 2$.

The *error-locator polynomial* $\sigma(z)$ is defined by the formula

$$\sigma(z) = \prod_{i=1}^{s}(1 - \eta_i z).$$

Thus the roots of σ are precisely the reciprocals of the error locators.

We also define the *error-evaluator polynomial* by the formula

$$\epsilon(z) = \sum_{i=1}^{s} h_{m_i} e_{m_i} \prod_{j=1, j\neq i}^{s}(1 - \eta_j z).$$

Proposition (Forney's formula). For $k = 1, \dots, s$, we have

$$e_{m_k} = -\eta_k \epsilon\left(\eta_k^{-1}\right) / h_{m_k} \sigma'\left(\eta_k^{-1}\right),$$

where σ' is the derivative of σ.

Proof. The derivative of σ is given by

$$\sigma'(z) = -\sum_i \eta_i \prod_{j=1, j\neq i}^{s} (1 - \eta_j z)$$

and from this expression we obtain that

$$\sigma'\left(\eta_k^{-1}\right) = -\eta_k \prod_{j\neq k} (1 - \eta_j/\eta_k).$$

On the other hand we have, from the definition of ϵ, that

$$\epsilon\left(\eta_k^{-1}\right) = h_{m_k} e_{m_k} \prod_{j\neq k} (1 - \eta_j/\eta_k).$$

Comparing the last two expressions we obtain the relation

$$\eta_k \epsilon\left(\eta_k^{-1}\right) = -h_{m_k} e_{m_k} \sigma'\left(\eta_k^{-1}\right),$$

which is equivalent to the stated formula.

Theorem (Key equation). The polynomials $\epsilon(z)$ and $\sigma(z)$ satisfy the congruence

$$\epsilon(z) \equiv \sigma(z)S(z) \mod z^r.$$

Proof. By definition of ϵ it is clear that we also have

$$\epsilon(z) = \sigma(z) \sum_{i=1}^{s} \frac{h_{m_i} e_{m_i}}{1-\eta_i z}.$$

But $\sum_{i=1}^{s} \frac{h_{m_i} e_{m_i}}{1-\eta_i z} = \sum_{i=1}^{s} h_{m_i} e_{m_i} \sum_{j\geq 0} (\eta_i z)^j$

$$\equiv \sum_{i=1}^{s} h_{m_i} e_{m_i} \sum_{j=0}^{r-1} (\eta_i z)^j \mod z^r$$

$$= \sum_{j=0}^{r-1} \left(\sum_{i=1}^{s} h_{m_i} e_{m_i} \eta_i^j \right) z^j = \sum_{j=0}^{r-1} S_j z^j = S(z).$$

Remark. The key equation implies that

$$\deg(\gcd(z^r, S(z))) < t$$

as $\gcd(z^r, S(z))$ divides ϵ (this conclusion was obtained before in a different way).

Solution of the key equation. The key equation shows that there exists a unique polynomial $\tau(z)$ such that

$$\epsilon(z) = \tau(z)z^r + \sigma(z)S(z).$$

This equation is equivalent to the key equation, and one of the crucial steps in the decoding of alternant codes is to find a solution (σ and ϵ) in terms of z^r and $S(z)$.

The method that we will present is a modification of Euclid's algorithm for finding the gcd of two polynomials.

The algorithm Sugiyama(z^r,S,t). It is a variation of Euclid's algorithm. The input is a pair of polynomials $r_0 = z^r$ and $r_1 = S(z)$ (recall that we have assumed that $S \neq 0$), and an integer t. The description of the algorithm is as follows:

1. Let r_i, $i = 0, \dots, j$, be the polynomials that are calculated with Euclid's algorithm applied to r_0 and r_1, but with j equal to the first index such that $\deg(r_j) < t$. For $i = 2, \dots, j$, let q_i be the quotient of the Euclidean division of r_{i-2} by r_{i-1}, so that
$$r_i = r_{i-2} - q_i r_{i-1}.$$
2. Note that since $\deg(r_1) = \deg(S) \geq t$ (as seen before), we have $j \geq 2$. Note also that the integer j exists, for the degree of the gcd d of r_0 and r_1 is less than t, and we know that the full Euclid algorithm returns d.
3. Define $v_0, v_1, \dots, v_j$ so that $v_0 = 0$, $v_1 = 1$ and $v_i = v_{i-2} - q_i v_{i-1}$ for $i = 2, \dots, j$.
4. Return the pair $\{v_j, r_j\}$.

In order to establish that the Sugiyama algorithm produces the wanted solution of the key equation, it is convenient to calculate, together with $v_0, v_1, \ldots, v_j$, the sequence $u_0, u_1, \ldots, u_j$ such that $u_0 = 1$, $u_1 = 0$, and $u_i = u_{i-2} - q_i u_{i-1}$ for $i = 2, \ldots, j$.

r_0	r_1	r_2	r_3	$\cdots$	r_{j-2}	r_{j-1}	r_j
		q_2	q_3	$\cdots$	q_{j-2}	q_{j-1}	q_j
$v_0 = 0$	$v_1 = 1$	v_2	v_3	$\cdots$	v_{j-2}	v_{j-1}	v_j
$u_0 = 1$	$u_1 = 0$	u_2	u_3	$\cdots$	u_{j-2}	u_{j-1}	u_j

- We have that $u_i r_0 + v_i r_1 = r_i$ for all $i = 0, \ldots, j$.

Remark (Extended Euclid Algoritm). If we modify Sugiyama's algoritm so that j is the greatest integer such that $r_j \neq 0$, then $d = r_j$ is equal to $\gcd(r_0, r_1)$ and the identity above shows that $a_0 = u_j$ and $a_1 = v_j$ yield a solution of *Bezout's identity*:

$$a_0 r_0 + a_1 r_1 = d.$$

Let us continue with the notations introduced in the description of the Sugiyama algorithm. Recall that $t' = \lceil r/2 \rceil$ and $r - t = t'$.

Lemma. Let $\bar{\epsilon} = r_j$, $\bar{\tau} = u_j$, $\bar{\sigma} = v_j$. Then

$$\bar{\epsilon}(z) = \bar{\tau}(z)z^r + \bar{\sigma}(z)S(z), \text{ with } \deg(\bar{\sigma}) \leq t', \ \deg(\bar{\epsilon}) < t.$$

Proof: We have $u_i r_0 + v_i r_1 = r_i$, $i = 0, \dots, j$. For $i = j$ it coincides with the equality of the statement. Now we will use induction on i to show that $\deg(v_i) = r - \deg(r_{i-1})$ for $i = 1, \dots, j$ (and thereby that $\deg(v_i)$ is strictly increasing with i). Since the relation is clearly true for $i = 1$, we can assume that $i > 1$. Then $v_i = v_{i-2} - q_i v_{i-1}$ and

$$\begin{aligned} \deg(v_i) &= \deg(q_i) + \deg(v_{i-1}) \\ &= \deg(r_{i-2}) - \deg(r_{i-1}) + r - \deg(r_{i-2}) \\ &= r - \deg(r_{i-1}) \end{aligned}$$

(in the second step we have used the definition of q_i and the induction hypothesis).

In particular

$$\deg(\bar{\sigma}) = \deg(v_j) = r - \deg(r_{j-1}) \leq r - t = t',$$

and this establishes the first inequality in the statement. The second inequality is a direct consequence of the definition of j and $\bar{\epsilon}$.

Remark. With the same notations as in the lemma, it is straightforward that

$$u_i v_{i-1} - v_i u_{i-1} = (-1)^i, \ i = 1, \ldots, j.$$

This implies that $\gcd(u_i, v_i) = 1$. So, in particular, $\gcd(\bar{\tau}, \bar{\sigma}) = 1$.

Theorem. With notations as in the lemma, there exists $\rho \in \mathbb{F}_{q^m}^*$ such that

$$\sigma = \rho\bar{\sigma} \text{ i } \epsilon = \rho\bar{\epsilon}.$$

Proof: Multiplying the key equation $\epsilon(z) = \tau(z)z^r + \sigma(z)S(z)$ by $\bar{\sigma}$, the equation $\bar{\epsilon}(z) = \bar{\tau}(z)z^r + \bar{\sigma}(z)S(z)$ in the lemma by σ, and subtracting the results, we obtain the identity

$$\bar{\sigma}\epsilon - \sigma\bar{\epsilon} = (\bar{\sigma}\tau - \sigma\bar{\tau})z^r.$$

The degree of the polynomial in the left hand side is $< r$, as

$$\deg(\bar{\sigma}\epsilon) = \deg(\bar{\sigma}) + \deg(\epsilon) \leq t' + t - 1 = r - 1,$$

$$\deg(\sigma\bar{\epsilon}) = \deg(\sigma) + \deg(\bar{\epsilon}) \leq t + t - 1 \leq r - 1.$$

Since the polynomial on the right hand side contains the factor z^r, we infer that $\bar{\sigma}\epsilon = \sigma\bar{\epsilon}$, $\bar{\sigma}\tau = \sigma\bar{\tau}$. Hence $\sigma|\bar{\sigma}\epsilon$ and $\bar{\sigma}|\bar{\tau}\sigma$. As $\gcd(\sigma, \epsilon) \equiv 1$, because σ and ϵ have no common roots, and also $\gcd(\bar{\tau}, \bar{\sigma}) = 1$, we obtain $\sigma|\bar{\sigma}$, $\bar{\sigma}|\sigma$ and $\sigma|\bar{\sigma}\epsilon$, $\bar{\sigma}|\sigma$. Therefore there exists $\rho \in \mathbb{F}_{q^m}^*$ such that $\sigma = \rho\bar{\sigma}$ and $\epsilon = \rho\bar{\epsilon}$, as claimed.

Remark. The theorem shows that $\bar{\sigma}$ and σ have the same roots, and so we can use $\bar{\sigma}$ instead of σ for finding the error locations. Moreover, the Forney's formula shows that we can use $\bar{\sigma}$ and $\bar{\epsilon}$ instead of σ and ϵ in order to find the error values:

$$\frac{\eta_k \bar{\epsilon}(\eta_k^{-1})}{h_{m_k} \bar{\sigma}'(\eta_k^{-1})} = \frac{\eta_k \epsilon(\eta_k^{-1})}{h_{m_k} \sigma'(\eta_k^{-1})} .$$

The Berlekamp–Massey–Sugiyama algorithm

Let H be the control matrix of the alternant code $C = C_H \subseteq F^n$ of order r associated to the vectors $\boldsymbol{h}, \boldsymbol{\alpha} \in \bar{F}^n$, and let

$$\boldsymbol{\beta} = (\beta_1, \dots, \beta_n), \beta_i = 1/\alpha_i, \text{ i } t = \lfloor r/2 \rfloor.$$

Let $y \in F^n$ be the received vector.

Algorithm BMS

Input: y

1. Calculate $s = yH^T$, $s = (s_0, \dots, s_{r-1})$, and form the polynomial $S = s(z)$ in the variable z, $S = s_0 + s_1 z + \cdots + s_{r-1} z^{r-1}$ (we say that S is the *polynomic syndrome*).
2. Let $\{\sigma, \epsilon\}$ be the pair returned by Sugiyama(z^r, S, t).
3. Form the list $M = \{m_1, \dots, m_s\}$ of the indeces $m \in \{1, \dots, n\}$ such that $\sigma(\beta_m) = 0$ (we call them *error position*). If $s < \deg(\sigma)$, return *Error*.
4. Let x be the result of substituting y_m by $y_m + e_m$, for each $m \in M$, where

$$e_m = \frac{\alpha_m \cdot \epsilon(\beta_m)}{h_m \cdot \sigma'(\beta_m)}$$

(σ' is the derivative of σ). If $e_m \notin F$ for some m, return *Error*. Otherwise return x if x is a code vector, or *Error* if not.

Theorem. The algorithm BMS corrects $t = \lfloor r/2 \rfloor$ errors.

The algorithm PGZ

We shall present another decoding algorithm for alternant codes. Define the *error locating polynomial* by

$$\sigma(z) = \prod_{i=1}^{s}(z - \eta_i),$$

so that now the roots of σ are the error locators η_i. From now on we will assume that $s \leq t$.

Determination of the number of errors. For each integer ℓ such that $s \leq \ell \leq t$, define

$$A_\ell = \begin{pmatrix} S_0 & S_1 & \cdots & S_{\ell-1} \\ S_1 & S_2 & \cdots & S_\ell \\ \vdots & \vdots & & \vdots \\ S_{\ell-1} & S_\ell & \cdots & S_{2\ell-2} \end{pmatrix},$$

which is called the *Hankel matrix* associated to the vector

$$(S_0, S_1, \ldots, S_{2\ell-2}).$$

Note that this vector exists, as $2\ell - 2 \leq r - 2$.

Lemma. We have that $\det(A_\ell) = 0$ for $s < l \leq t$ and $\det(A_s) \neq 0$. In other words, s is the greatest integer (among those satisfying $s \leq t$) such that $\det(A_s) \neq 0$.

Proof: Let $M' = \{m'_1, \dots, m'_\ell\} \subseteq \{0, \dots, n-1\}$ be any subset such that $M \subseteq M'$. For $i = 1, \dots, \ell$, let $\eta_i = \alpha_{m'_i}$. As we already know,

$$S_j = \sum_{k=1}^{s} h_{m_k} e_{m_k} \alpha_{m_k}^j = \sum_{k=1}^{\ell} h_{m'_k} e_{m'_k} \eta_k^j$$

for $j = 0, \dots, r-1$.

Let $D = \operatorname{diag}\left(h_{m'_1} e_{m'_1}, \dots, h_{m'_\ell} e_{m'_\ell}\right)$, so that $\det(D) \neq 0$ if $\ell = s$ and $\det(D) = 0$ if $\ell > s$. Set $W = V_\ell(\eta_1, \dots, \eta_\ell)$ (the Vandermonde matrix of ℓ rows associated to the elements η_i. Note that in particular we have $\det(W) \neq 0$.

We also have that $WDW^T = A_\ell$, since the i-th row of W is $(\eta_1^i, \dots, \eta_\ell^i)$, the j-th column of DW^T is $\left(h_{m'_1} e_{m'_1} \eta_1^j, \dots, h_{m'_\ell} e_{m'_\ell} \eta_\ell^j\right)$, and their product

$$\sum_{k=1}^{\ell} h_{m'_k} e_{m'_k} \eta_k^{i+j} = S_{i+j} \quad (i, j = 0, \dots, \ell-1).$$

Thus $\det(A_\ell) = \det(D) \cdot \det(W)^2$, which is 0 if $\ell > s$ (in this case $\det(D) = 0$), and $\neq 0$ if $\ell = s$ (in this case $\det(D) \neq 0$ and $\det(W) \neq 0$).

How to find the error locator polynomial. Once the number s of errors is known (recall that we assume that it $\leq t$), we can find the coefficients of the error locating polynomial as follows. Note that

$$\sigma(z) = \prod_{i=1}^{s}(z - \eta_i) = z^s + a_1 z^{s-1} + \cdots + a_s \,,$$

where $a_j = (-1)^j \sigma_j(\eta_1, \ldots, \eta_s)$, with $\sigma_j = \sigma_j(\eta_1, \ldots, \eta_s)$ the j-th elementary symmetric polynomial in the η_j $(0 \leq j \leq s)$.

Remark. The polynomial $1 + a_1 z + \cdots + a_s z^s$ is the error locator $\prod_{i=1}^{s}(1 - \eta_i z)$ considered in the study of the BMS algorithm.

Proposition. If $\boldsymbol{a} = (a_s, \ldots, a_1)$ is the coefficient vector of σ and $\boldsymbol{b} = (S_s, \ldots, S_{2s-1})$, then $\boldsymbol{a} A_s = -\boldsymbol{b}$. Since $\det(A_s) \neq 0$, this relation determines $\boldsymbol{a}$ uniquely: $a = -b(A_s)^{-1}$.

Proof: Substituting z by the η_i in the identity

$$\prod_{i=1}^{s}(z - \eta_i) = z^s + a_1 z^{s-1} + \cdots + a_s$$

we obtain the relations

$$\eta_i^s + a_1\eta_i^{s-1} + \cdots + a_s = 0,\ i = 1, \ldots, s\ .$$

Multiplying by $h_{m_i} e_{m_i} \eta_i^j$ and adding with respect to i, we obtain the relations

$$S_{j+s} + a_1 S_{j+s-1} + \cdots + a_s S_j = 0,\ j = 0, \ldots, s-1\ ,$$

and it is immediate to check that these relations are equivalent to the claimed matrix relation.

The algorithm PGZ**.** The lemma and the proposition allow us to formulate an algorithm, essentially due to Peterson, Gorenstein and Zierler, to descode alternant codes. With conventions similar to those used in the algorithm BMS, including those relative to the meaning of *Error*, the PGZ algorithm goes as follows:

1) Calculate the syndrome vector,

$$S = (S_0, \dots, S_{r-1}) = yH^T.$$

If $S = 0$, return y.

2) Thus we assume that $S \neq 0$. Beginning with $s = t$, and while $\det(A_s) = 0$, set $s = s - 1$. The value of s at the end of this loop is the number of errors ($s > 0$, as otherwise S would be 0).

3) Calculate $(a_s, \dots, a_1) = -(S_s, \dots, S_{2s-1})(A_s)^{-1}$, and form the polynomial

$$\sigma(z) = z^s + a_1 z^{s-1} + \cdots + a_s.$$

4) Find the elements α_j that aret he roots of σ. If the number of these roots is $< s$, return *Error*. Otherwise, let $\eta_1, \dots, \eta_s$ be the error locators corresponding to the roots ans set $M = \{m_1, \dots, m_s\}$, where $\eta_i = \alpha_{m_i}$.

5) Determine the error evaluator $\epsilon(z)$ by reducing mod z^r the product

$$(1 + a_1 z + \cdots + a_s z^s) S(z) \, .$$

6) Calculate the errors e_{m_i} by means of Forney's formula with the error locator $1 + a_1 z + \cdots + a_s z^s$ and the error evaluator $\epsilon(z)$. If any of these values does not lie in K, return *Error*. Otherwise, return $y - e$ or *Error* according to whether $y - e$ is or is not a code vector.

Proposition. PGZ correct up to t errors.

Remark. The determination of the errors $e_{m_1}, \dots, e_{m_s}$ can also be done by solving the system of linear equations

$$h_{m_1} e_{m_1} \eta_1^j + \cdots + h_{m_s} e_{m_s} \eta_s^j = S_j .$$