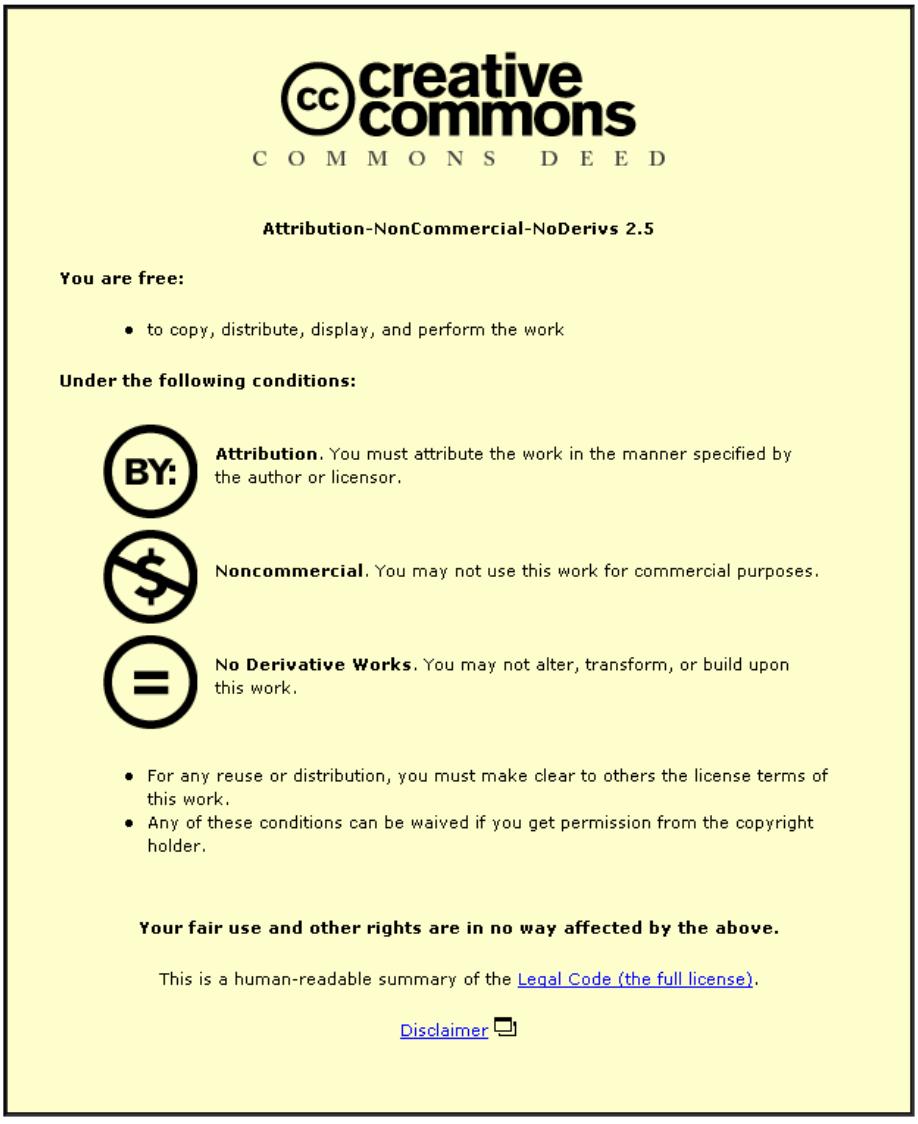

creative commons

COMMONS DEED

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

BY: **Attribution**. You must attribute the work in the manner specified by the author or licensor.

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the Legal Code (the full license).

Disclaimer

On the key equation over a commutative ring. *

Graham H. Norton, Dept. Mathematics, Univ. of Queensland, Brisbane

Ana Sălăgean, Dept. Mathematics, Nottingham Trent Univ., Nottingham, U.K.

Abstract

We define alternant codes over a commutative ring R and a corresponding key equation. We show that when the ring is a domain, e.g. the p-adic integers, the error–locator polynomial is the unique monic minimal polynomial (shortest linear recurrence) of the syndrome sequence and that it can be obtained by Algorithm MR of Norton.

When R is a local ring, we show that the syndrome sequence may have more than one (monic) minimal polynomial, but all the minimal polynomials coincide modulo the maximal ideal of R. We characterise the minimal polynomials when R is a Hensel ring. We also apply these results to decoding alternant codes over a local ring R: it is enough to find any monic minimal polynomial over R and to find its roots in the residue field. This gives a decoding algorithm for alternant codes over a finite chain ring, which generalizes and improves a method of Interlando *et. al.* for BCH and Reed–Solomon codes over a Galois ring.

1 Introduction

Block codes over finite rings were initially studied in [Bla75, Sha79]. A modification of the Berlekamp–Massey algorithm for $\mathbb{Z}/m\mathbb{Z}$ was given in [RS85], where it was claimed [*loc. cit.*, Introduction] (without proof) to decode BCH codes defined over the integers modulo m. This relatively classical topic was recently reinvigorated with the publication of the landmark paper [HKC$^+$94]. An algorithm to decode BCH and Reed–Solomon codes over a Galois ring has been given in [IPE97]. See also [Nor98], which gives an analogue of the Berlekamp–Massey algorithm to find a monic minimal polynomial (shortest linear recurrence) for any finite sequence over a finite chain ring (e.g. a Galois ring) and which has quadratic complexity. Both of the latter algorithms require root–finding in the ring itself.

We define alternant codes and a corresponding key equation over a commutative ring with identity, R. Important examples of alternant codes over a ring are BCH and Reed–Solomon codes over a Galois ring. We concentrate on decoding alternant codes over a domain or a local ring.

When R is a domain, the error–locator polynomial is the unique monic minimal polynomial (shortest linear recurrence) of the syndrome sequence; see Theorem 4.4. We show that it can be easily obtained using Algorithm MR of [Nor95]. Once we have the error–locator polynomial, the error locations and magnitudes can be computed in the same way as over a field. Hence we can decode any alternant code over a domain, e.g. the p-adic integers.

When R is a Hensel ring, i.e. a local ring which admits Hensel lifting (e.g. a finite local ring), we characterize in Theorem 4.8 the set of monic minimal polynomials of a finite syndrome sequence over R (there may be more than one). Our characterization is independent of any particular algorithm, the number of roots of a minimal polynomial and the theory of Linear Systems over

*Research supported in part by the U.K. Engineering and Physical Sciences Research Council under Grant L07680.

a finite ring, *cf.* [IPE97]. We also show (Theorem 4.5) that for any local ring, the minimal polynomials of a syndrome sequence coincide modulo the maximal ideal M of R.

We apply these results to give a new decoding algorithm (Algorithm 5.1) for alternant codes over a local ring R, once a minimal polynomial μ of the syndrome sequence is known. It turns out that for determining the error locations it is enough to find the roots of the image of any such minimal polynomial in the residue field. This is also more efficient than finding roots in R, *cf.* [IPE97, Nor98]. After determining the error locations, the error magnitudes can easily be computed. When R is a finite chain ring i.e. a finite ring in which all the ideals are linearly ordered by inclusion (or equivalently a finite local ring with principal maximal ideal), we invoke Algorithm MP of [Nor98] to find a monic minimal polynomial μ. Our method can be applied to any alternant code over a finite chain ring, in particular to BCH and Reed–Solomon codes over a Galois ring.

It would be interesting to extend Algorithm MP of [Nor98] to a local ring (which is not necessarily finite and where M is not necessarily principal). We would then be able to decode alternant codes over a local ring.

2 Preliminaries

2.1 Subtractive subsets

Let R be a commutative ring with $1 \neq 0$. Let $N(R)$ denote the subset of R consisting of all elements which are *not* zero–divisors. Then $N(R)$ is a multiplicative subset of R which contains the units of R. (If R is a domain which is not a field, then by definition, R has elements which are neither zero–divisors nor units. For example, if R is the domain of p–adic integers then $p \in N(R)$, but p is not a unit of R.)

The following result generalizes the trivial fact that if $a \in N(R)$ and $ax = 0$, then $x = 0$:

Lemma 2.1 *If A is a square matrix over R and $\det(A) \in N(R)$, then the homogeneous linear system $Ax = 0$ has only the trivial solution.*

Proof. For a linear system $Ax = b$ over a ring, we have by Cramér's rule (see [McD84, page 80]) $\det(A)x_i = \det(A_i)$ for $i = 1, \ldots, n$, where A_i is the matrix obtained from A by replacing the i^{th} column by b. In our case $b = 0$, so $\det(A)x_i = 0$ for $1 \leq i \leq n$. Finally, $\det(A) \in N(R)$ so the only solution is $x_i = 0, i = 1, \ldots, n$. □

The following notion is motivated by minimum distance considerations (Theorem 3.3).

Definition 2.2 *We say that $S \subseteq N(R)$ is subtractive in $N(R)$ if for all distinct $a,\ b \in S,\ a - b \in N(R)$.*

We will abbreviate 'subtractive in $N(R)$' to 'subtractive'. Clearly R is a domain iff all subsets of $R \setminus \{0\}$ are subtractive. If $n \geq 2$, then $\{1, 1 + n\} \subseteq \mathbb{Z}_{n^2}$ is not subtractive.

Lemma 2.3 *If $f \in R[X]$ has the distinct roots r_i where $1 \leq i \leq n$ and $\{r_1, \ldots, r_n\}$ is subtractive, then $\prod_{i=1}^{n}(X - r_i)|f$.*

Proof. Since $X - r_1$ is monic, the usual argument over a field shows that $(X - r_1)|f$. Hence $f = (X - r_1)g_1$ for some $g_1 \in R[X]$. Evaluating at r_2, we obtain $(r_2 - r_1)g_1(r_2) = 0$. Since $r_2 - r_1 \in N(R)$, we have $g_1(r_2) = 0$, i.e. $f = (X - r_1)(X - r_2)g_2$ for some $g_2 \in R[X]$. Continuing in this way, we obtain $f = g_n \prod_{i=1}^{n}(X - r_i)$ for some $g_n \in R[X]$. □

Note that this result fails if we drop the condition that $\{r_1, \ldots, r_n\}$ be subtractive. For example, if $n \geq 3$, $f = (X-1)^2 \in \mathbb{Z}_{n^2}[X]$ and $r_i = 1 + (i-1)n$ for $i = 1, \ldots, n$, then $f(r_i)$=0 for $i = 1, \ldots, n$, but $\prod_{i=1}^{n}(X - (i-1)n) \not| f$.

2.2 Local Rings

If R is a local ring, let M be the maximal ideal of R and $K = R/M$ the residue field. For any element $y \in R$ we will denote by $\overline{y}$ its image under the canonical projection from R to K. We extend this projection in the natural way to a projection from $R[X]$ to $K[X]$. For a set $S \subseteq R[X]$ we define $\overline{S} = \{\overline{s} : s \in S\}$. We can regard a field F as a local ring with $M = (0)$ and $F \to K$ the identity map.

Recall that the units of R are exactly the elements of $R \setminus M$ and all zero–divisors of R are contained in M. The following result is known (see for example [McD74, Exercise I.8]). We give a simple proof for completeness.

Proposition 2.4 *Let R be a finite ring. Then every element of $N(R)$ is a unit.*

Proof. Let A be the set of non–units of R. Let $x \in N(R)$ and assume x is not a unit, i.e. $x \in A$. Then $xR \subseteq A$. Since $|A| < |R| < \infty$, we must have $rx = r'x$ for distinct $r,\ r' \in R$ i.e. x is a zero–divisor, which is a contradiction. □

We give now a few simple properties of subtractive sets in local rings.

Lemma 2.5 *Let R be a local ring and $r,\ r' \in R$. Then $\overline{r} \neq \overline{r'}$ iff $r - r'$ is a unit of R.*

Corollary 2.6 *Let R be a local ring and S a finite subset of R.*

(i) If $|S| = |\overline{S}|$, then S is subtractive.

(ii) If all elements of $N(R)$ are units (in particular if R is finite), then S is subtractive iff $|S| = |\overline{S}|$.

(iii) If $S = \{1, \gamma, \ldots, \gamma^{n-1}\}$ and γ is a unit in R such that $\overline{\gamma}$ has order at least n, then $|S| = |\overline{S}|$ and S is subtractive.

Note that for rings with elements that are neither units nor zero–divisors, the property that $|S| = |\overline{S}|$ is strictly stronger than S being subtractive. For example in the ring of p-adic integers the set $\{1, 1+p\}$ is subtractive but $\overline{1} = \overline{1+p} = 1$.

By a *monic* polynomial, we mean a non–zero polynomial with leading coefficient 1.

Definition 2.7 (Hensel ring) *We say that a local ring R is a Hensel ring if R admits Hensel lifting i.e.*

$$\text{if } F \in R[X] \text{ is monic, } \overline{F} = g \cdot h \text{ and } g,\ h \in K[X] \text{ satisfy } (g,h) = 1,$$

then there are monic $G,\ H \in R[X]$ such that

$$F = G \cdot H \text{ and } \overline{G} = g,\ \overline{H} = h.$$

From now on, we will say 'lift' for 'Hensel lift'. The following theorem is an analogue of [McD74, Lemma XV.1] for Hensel rings. The proof is similar.

Theorem 2.8 *If R is a Hensel ring, $F \in R[X]$ is monic and $\alpha \in K$ is a simple root of $\overline{F}$, then there is a unique root $\beta \in R$ of F such that $\overline{\beta} = \alpha$. We call β the lift of α.*

The next result is useful for constructing Galois extensions and cyclic codes. A similar version appears in [Sha79, Theorem 3]. We let ord( ) denote the order function on R and on K.

Theorem 2.9 *Let R be a Hensel ring and let n be a natural number not divisible by the characteristic of K. Assume there is an $\alpha \in K$ such that $\alpha^n = 1$. Then there is a unique β in R such that $\beta^n = 1$ and $\overline{\beta} = \alpha$. Moreover* $\mathrm{ord}(\beta) = \mathrm{ord}(\alpha)$.

PROOF. Since $\mathrm{ord}(\alpha)|n$ and n is not divisible by the characteristic of K, $X^{\mathrm{ord}(\alpha)} - 1$ has only simple roots in K. So by Theorem 2.8, there is a unique β such that $\beta^n = 1$ and $\overline{\beta} = \alpha$.

From $\overline{\beta^{\mathrm{ord}(\beta)}} = 1 = \overline{\beta}^{\mathrm{ord}(\beta)}$ we infer that $\mathrm{ord}(\overline{\beta})|\,\mathrm{ord}(\beta)$. Hence we can write $X^{\mathrm{ord}(\beta)} - 1 = (X^{\mathrm{ord}(\overline{\beta})} - 1)h$ for some $h \in R[X]$. The root $\overline{\beta}$ of $X^{\mathrm{ord}(\overline{\beta})} - 1$ can be lifted to a root γ of $X^{\mathrm{ord}(\overline{\beta})} - 1$ in R and $\mathrm{ord}(\gamma)|\,\mathrm{ord}(\overline{\beta})$. Now both γ and β are roots of $X^{\mathrm{ord}(\beta)} - 1$ and $\overline{\gamma} = \overline{\beta}$. Since $\mathrm{ord}(\beta)|n$, all roots of $X^{\mathrm{ord}(\beta)} - 1$ are simple, so by Theorem 2.8, $\beta = \gamma$ and so $\mathrm{ord}(\beta)|\,\mathrm{ord}(\overline{\beta})$). We conclude that $\mathrm{ord}(\beta) = \mathrm{ord}(\overline{\beta})$. □

It is well–known that if R is complete in its M–adic topology, then R is a Hensel ring. For more details on general Hensel rings, see [Eis95].

A finite local ring is a Hensel ring (see [McD74, Theorem XII.4]). Finite local rings are completely classified (see the structure theorems in [McD74, Chapter XVII]). Recall that chain ring is a ring in which all its ideals are linearly ordered by inclusion and that R is a finite chain ring iff it is a finite local ring with M principal.

If p is a prime and $a, l \in \mathbb{N}$ are strictly positive, the Galois ring $R = GR(p^a, l)$ is the quotient ring $\mathbb{Z}_{p^a}[y]/(f)$, where f is a monic irreducible polynomial of degree l such that $\overline{f}$ is irreducible in $\mathbb{Z}_p[y]$. It is a finite local ring with $M = (p)$ and $K = GF(p^l)$. The integers modulo a power of a prime and their Galois extensions are important special cases of finite chain rings and we refer the reader to [McD74, Ch. XVI] for the general theory.

An important example of an infinite Hensel ring is the ring of p–adic integers, denoted $\mathbb{Z}_{p^\infty}$. For details on the construction and properties of p–adic numbers we refer the reader to [Coh89, Ch. 8]. The ring of p–adic integers is a unique factorization domain with $M = (p)$. We can construct Galois extensions of $\mathbb{Z}_{p^\infty}$ in a way similar to the Galois extensions of $\mathbb{Z}_{p^a}$ above (see [CS95]). Namely, we put $GR(p^\infty, l) = \mathbb{Z}_{p^\infty}[y]/(f)$, where f is a monic irreducible factor of degree l of $X^{p^l-1} - 1$. (Such an f can be obtained by lifting the factorisation of $X^{p^l-1} - 1$ from $\mathbb{Z}_p[y]$). Such a Galois ring is again a Hensel ring with $M = (p)$. The canonical projections from $\mathbb{Z}_{p^\infty}$ to each $\mathbb{Z}_{p^a}$, $a \geq 1$ can be extended in a natural way to projections from $GR(p^\infty, l)$ to $GR(p^a, l)$.

2.3 Laurent series

We will use the following fact without further mention.

Proposition 2.10 *If $f, g \in R[X]$, f is monic and* $\deg(g) \leq \deg(f)$, *then $g/f \in R[[X^{-1}]]$.*

PROOF. If $d = \deg(f)$ then $f = X^d(1 + f_{d-1}X^{-1} + \ldots + f_0X^{-d}) = X^d(1 - h)$ say, and $(1 - h)^{-1} = 1 + h + h^2 + \ldots \in R[[X^{-1}]]$. Thus $1/f \in X^{-d}R[[X^{-1}]]$, and if $\deg(g) \leq d$, $g/f \in X^{\deg(g)-d}R[[X^{-1}]] \subseteq R[[X^{-1}]]$. □

As usual, $R((X^{-1}))$ denotes the ring of Laurent series in X^{-1}. For $i \in \mathbb{Z}$ and $F \in R((X^{-1}))$, F_i is the coefficient of X^i in F and we extend the degree function on $R[X]$ by $\delta(F) = \max\{i : F_i \neq 0\}$ for $F \neq 0$, with the convention $\delta(0) = -\infty$. Thus $F = \sum_{i \leq \delta(F)} F_i X^i$ and $\delta(FG) \leq \delta(F) + \delta(G)$. The ring of Laurent polynomials $R[X^{-1}, X]$ is a subring of $R((X^{-1}))$.

As in [Nor95], we prefer to study finite and linear recurring sequences by exploiting $R((X^{-1}))$ as an $R[X]$–module (i.e. to let $R[X]$ act on $R((X^{-1}))$ by multiplication in $R((X^{-1}))$), rather than using $R[[X]]$ and reciprocals of polynomials. [Thus we avoid defining linear recurrences using reciprocals. (Writing f^* for the reciprocal of $f \in R[X]$, it is easy to find polynomials f, g with $(f+g)^* \neq f^* + g^*$.) We also avoid an additional order function on $R[[X]]$ and 'linear feedback shift–registers (LFSR's)'.]

We note that the theory and resulting Algorithm MP of [Nor98] is both simpler and more general than the LFSR approach of [RS85]. Also, the algorithm reduces to the monic version of Algorithm MP of [Nor95] when R is a field.

Thus we index R–sequences negatively: the letter s denotes the infinite sequence $s_0, s_{-1}, s_{-2}, \ldots$ of elements of R, so that the generating function of s is $\Gamma(s) = \sum_{i \leq 0} s_i X^i$. We will also use the fact that $R((X^{-1})) = R[[X^{-1}]] \oplus XR[X]$.

We say that $f \in R[X]$ annihilates s if $(f \cdot \Gamma(s))_i = 0$ for all $i \leq 0$, and write $Ann(s)$ for the set of all polynomials f which annihilate s. Clearly $f \in Ann(s)$ iff $f \cdot \Gamma(s) \in XR[X]$ and s is a *linear recurring sequence* iff $Ann(s) \neq (0)$. By a *minimal polynomial of* s, we mean a monic annihilating polynomial of s which has minimal degree amongst all monic annihilating polynomials of s, and we write $Min(s)$ for the set of minimal polynomials of s. (In fact $Ann(s)$ is an ideal, and when R is a field, it is generated by a minimal polynomial of s.) The degree of a polynomial in $Min(s)$ is called the complexity of s. For any polynomial f we define

$$\beta(f, s) = \sum_{i=1}^{\delta(f)} (f\Gamma(s))_i X^i.$$

We now extend these definitions to cover the case of a finite sequence as in [Nor98]. For $m \in \mathbb{Z}$, $m \leq 0$, we denote by $s|m$ the finite sequence $s_0, s_{-1}, s_{-2}, \ldots, s_m$ of elements in R. The generating function of $s|m$ is $\Gamma(s|m) = \sum_{m \leq i \leq 0} s_i X^i$. A polynomial $f \in R[X]$ annihilates $s|m$ if $(f\Gamma(s|m))_i = 0$ for all i such that $m + \delta(f) \leq i \leq 0$. We denote by $Ann(s|m)$ the set of polynomials that annihilate $s|m$. Note that any polynomial of degree at least $1 - m$ is vacuously in $Ann(s|m)$. Also, we have $Ann(s) \subseteq Ann(s|m-1) \subseteq Ann(s|m)$.

By a *minimal polynomial of* $s|m$, we mean a monic annihilating polynomial of $s|m$ which has minimal degree amongst all monic annihilating polynomials of $s|m$. The set of all monic polynomials will be denoted by $Min(s|m)$. Note that any $s|m$ has a monic annihilating polynomial e.g. X^{1-m} and hence has a (monic) minimal polynomial. The unique degree of the polynomials in $Min(s|m)$ is called the complexity of $s|m$.

In the case of a finite sequence, we define

$$\beta(f, s|m) = \sum_{i=1}^{\delta(f)} (f\Gamma(s|m))_i X^i.$$

From the definition of an annihilating polynomial, we see that $f \in Ann(s|m)$ iff $f\Gamma(s|m) = \beta(f, s|m) + F$ for some $F \in R[X^{-1}]$ with $\delta(F) < m + \delta(f)$.

Example 2.11 *As an exercise for the reader, we observe that if $m = -1$, s_0 is not a unit in R and s_{-1} is not divisible by s_0, then $s_0 X - s_{-1} \in Ann(s|m)$, but the complexity of $s|m$ is 2.*

We conclude by collecting here a few technical results needed below. The following lemma is straightforward.

Lemma 2.12
(i) If $f \in Ann(s)$, then $f\Gamma(s) = \beta(s| - \delta(f))$
(ii) For any $m \leq -\delta(f)$ we have $\beta(f, s|m) = \beta(f, s| - \delta(f))$.

The next lemma is [Nor95, Corollary 3.25] which holds, with the same proof, for arbitrary R. We include the proof for the convenience of the reader.

Lemma 2.13 *Let $f, g \in Ann(s|m)$. If $\delta(f) + \delta(g) \leq 1 - m$ then $f\beta(g, s|m) = g\beta(f, s|m)$.*

PROOF. We have $f\Gamma(s|m) = \beta(f, s|m) + F$ for some F with $\delta(F) \leq m + \delta(f) - 1$, and $g\Gamma(s|m) = \beta(g, s|m) + G$ with $\delta(G) \leq m + \delta(g) - 1$. Then $fg\Gamma(s|m) = g\beta(f, s|m) + gF = f\beta(g, s|m) + fG$, hence $f\beta(g, s|m) - g\beta(f, s|m) = gF - fG$. For the degrees we have $\delta(f\beta(g, s|m) - g\beta(f, s|m)) = \delta(gF - fG) \leq m + \delta(f) + \delta(g) - 1 \leq 0$. But $f\beta(g, s|m) - g\beta(f, s|m) \in XR[X]$, so we can only have $f\beta(g, s|m) - g\beta(f, s|m) = 0$. □

3 Codes and decoding

3.1 Alternant codes

Definition 3.1 (Alternant codes) *Let $d \geq 2$, $N = N(R)$ and T be a subring of R. Suppose that $\alpha = (\alpha_1 \ldots, \alpha_n)$ and $y = (y_1, \ldots, y_n)$ are such that $\{\alpha_1 \ldots, \alpha_n\}$ is subtractive and $y_i \in N$ for $1 \leq i \leq n$. If*

$$H = \begin{bmatrix} y_1 & y_2 & \ldots & y_n \\ y_1\alpha_1 & y_2\alpha_2 & \ldots & y_n\alpha_n \\ y_1\alpha_1^2 & y_2\alpha_2^2 & \ldots & y_n\alpha_n^2 \\ \vdots & \vdots & & \vdots \\ y_1\alpha_1^{d-2} & y_2\alpha_2^{d-2} & \ldots & y_n\alpha_n^{d-2} \end{bmatrix} \tag{1}$$

then the alternant code of length n and alphabet T defined by H is the T–module

$$\mathcal{A}(\alpha, y, d) = \{c \in T^n : \ Hc^{\mathrm{tr}} = 0\}.$$

As usual, H is called the parity check matrix.

Example 3.2 *When $R = GF(q^a)$ and $T = GF(q)$, we obtain the usual notion of an alternant code, see e.g. [MS77, Chapter 12]. Indeed, the natural projection induces a T–homomorphism of codes*

$$\mathcal{A}(\alpha, y, d) \to \mathcal{A}(\overline{\alpha}, \overline{y}, d),$$

where $\overline{\alpha}, \overline{y}$ have the obvious meaning.

Theorem 3.3 *The minimum Hamming distance of $\mathcal{A}(\alpha, y, d)$ is at least d.*

PROOF. We use the classical argument. If there were a codeword of weight $d - 1$ or less, having the positions of the non–zero entries among $i_1, \ldots, i_{d-1}$ say, then the homogeneous system $Ax = 0$, where A consists of columns $i_1, \ldots, i_{d-1}$ of H, would have a non–trivial solution. But we will show this system can only have the trivial solution.

The matrix A is vanderMonde with determinant $\det(A) = y_{i_1} y_{i_2} \ldots y_{i_{d-1}} \prod_{1 \leq j < k \leq d-1} (\alpha_{i_k} - \alpha_{i_j}) \in N$ since all $y_{i_j} \in N$ and $\{\alpha_1, \ldots, \alpha_n\}$ is subtractive. The result now follows from Lemma 2.1. □

BCH and Reed–Solomon codes are particular cases of alternant codes and we can also specialize R to Galois and p–adic rings:

Example 3.4 (BCH codes over a Galois ring) *To define a BCH code of length n over a Galois ring $T = GR(p^a, l_1)$, we take an extension ring $R = GR(p^a, l)$ such that $l_1 | l$ and $n | p^l - 1$.*

Then there exists γ a primitive n^{th} root of unity in R (see [McD74, Theorem XVI.9]). Actually, by Theorem 2.9 γ is the lift of a primitive n^{th} root of unity in $GF(p^a)$, hence the order of $\overline{\gamma}$ is n. In the definition of alternant codes we put $y_i = (\gamma^b)^{i-1}$ for some $b \geq 0$, and $\alpha_i = \gamma^{i-1}$ for $1 \leq i \leq n$. Since $\overline{\gamma}$ has order n, by Corollary 2.6, $\overline{\alpha}_1, \ldots, \overline{\alpha}_n$ are distinct and $\{\alpha_1, \ldots, \alpha_n\}$ is subtractive. We then obtain the classical form of the parity check matrix for a BCH code of length n and designed distance d:

$$H = \begin{bmatrix} 1 & \gamma^b & (\gamma^b)^2 & \ldots & (\gamma^b)^{n-1} \\ 1 & \gamma^{b+1} & (\gamma^{b+1})^2 & \ldots & (\gamma^{b+1})^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \gamma^{b+d-2} & (\gamma^{b+d-2})^2 & \ldots & (\gamma^{b+d-2})^{n-1} \end{bmatrix} \tag{2}$$

The code is cyclic and the generator polynomial is the product of the distinct minimal polynomials of $\gamma^b, \gamma^{b+1}, \ldots, \gamma^{b+d-2}$ over $T[X]$. This construction coincides with that of [Sha79].

Example 3.5 (Reed–Solomon codes over a Galois ring) *These are defined like BCH codes, except that now the alphabet T of the code is R, which is taken to contain the n^{th} roots of unity. In [Bla75], Reed–Solomon codes over $R = \mathbb{Z}_{p^a}$ of length $n|p-1$ are defined in a slightly different way, namely $\gamma \in \mathbb{Z}_{p^a}$ is taken to be an element of $\mathbb{Z}_p$ of order n in $\mathbb{Z}_p$. The parity check matrix is as in (2). The order of γ in $\mathbb{Z}_{p^a}$ is not necessarily n, but a multiple of n, and therefore these codes are not necessarily cyclic. (See e.g. [IPE97, Example 1, p. 1018].) The order of $\overline{\gamma}$ is still n, so again, by Corollary 2.6, $\overline{\alpha}_1, \ldots, \overline{\alpha}_n$ are distinct and $\{\alpha_1, \ldots, \alpha_n\} = \{\gamma^0, \ldots, \gamma^{n-1}\}$ is subtractive. Hence these Reed–Solomon codes are alternant codes as well. Moreover, they are a particular case of the generalised Reed–Solomon codes we will introduce below in Example 3.7.*

Example 3.6 (BCH codes over the p–adic integers) *Cyclic codes over the p–adic integers were introduced in [CS95]. To construct a BCH code of length n and designed distance d over $T = \mathbb{Z}_{p^\infty}$, we consider a Galois ring $R = GR(\mathbb{Z}_{p^\infty}, l)$ where $n|p^l - 1$. The n simple roots of $X^n - 1$ over $GF(p^l)$ lift to n simple roots of $X^n - 1$ in $GR(\mathbb{Z}_{p^\infty}, l)$. By Theorem 2.9, the lift of a primitive root will be primitive. We take γ a primitive n^{th} root of unity in $GR(\mathbb{Z}_{p^\infty}, l)$ and construct the parity check matrix of the code as in (2). Again, $\overline{\gamma}$ has order n, so by Corollary 2.6, $\overline{\alpha}_1, \ldots, \overline{\alpha}_n$ are distinct and $\{\alpha_1, \ldots, \alpha_n\}$ is subtractive. The code is cyclic and the generator polynomial is the product of the distinct minimal polynomials of $\gamma^b, \gamma^{b+1}, \ldots, \gamma^{b+d-2}$ over $T[X]$. Note that by projecting this code to $\mathbb{Z}_{p^a}$ we obtain a BCH code as in Example 3.4.*

Example 3.7 (Generalised Reed–Solomon codes over finite rings) *When $T = R$ we will call the alternant code a generalised Reed–Solomon code. We will assume that R is a finite ring. As in [MS77, Ch. 10,§8.], we denote this code by $GRS_k(\alpha, y)$, with $k = n - d + 1$. We will show that $GRS_k(\alpha, y)$ is a free R–module of rank k, its minimum distance is d and the dual code is $GRS_{n-k}(\alpha, y')$, for suitably chosen y'.*

By elementary row operations we get an equivalent form of H having the first $d-1$ columns in triangular form:

$$\begin{bmatrix} y_1 & y_2 & y_3 & \ldots & y_k & \ldots & y_n \\ 0 & y_2(\alpha_2 - \alpha_1) & y_3(\alpha_3 - \alpha_1) & \ldots & y_k(\alpha_k - \alpha_1) & \ldots & y_n(\alpha_n - \alpha_1) \\ 0 & 0 & y_3 \prod_{i=1}^{2}(\alpha_3 - \alpha_i) & \ldots & y_k \prod_{i=1}^{2}(\alpha_k - \alpha_i) & \ldots & y_n \prod_{i=1}^{2}(\alpha_n - \alpha_i) \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \\ 0 & 0 & 0 & \ldots & y_k \prod_{i=1}^{k-1}(\alpha_k - \alpha_i) & \ldots & y_n \prod_{i=1}^{k-1}(\alpha_n - \alpha_i) \end{bmatrix}$$

So far we did not need any divisions, we only subtracted from some rows of H other rows multiplied by suitable constants. Since R is finite, by Proposition 2.4 the set $N(R)$ coincides with the units of R, hence all y_i and all differences $\alpha_i - \alpha_j$ are units. We can therefore apply further row operations, including division by units, and bring H to a standard form $H = (I_{d-1}|A)$ for some matrix A. As

usual, a generator matrix can be obtained as $G = (-A^{\mathrm{tr}}|I_k)$. The rows of the generator matrix are linearly independent hence they are a basis for the module. By Theorem 3.3 we know the minimum distance is at least d. Any row of G is a codeword of weight d, hence the minimum distance of the code is exactly d.

The proof that the dual of $GRS_k(\alpha, y)$ is $GRS_{n-k}(\alpha, y')$, for suitably chosen y', is similar to the one for fields ([MS77, Theorem 4, Ch. 8]). For the two codes to be orthogonal, it suffices that y' satisfies the system of n equations in $n-1$ unknowns $\sum_{i=1}^{n} y_i y'_i \alpha_i^j = 0, j = 0, \ldots, n-1$. The determinant of any $n-1$ columns of the matrix of the system is vanderMonde and it is a unit. Putting for example $y_n = 1$ and solving the system by Cramér's rule, we get a solution y with all the y_i's units.

3.2 A key equation

For decoding alternant codes we follow the main steps of algebraic decoding of alternant codes over finite fields. Suppose that the codeword c is received as $r = c + e$. We have to find the error vector e given the syndrome vector $Hr^{\mathrm{tr}} = He^{\mathrm{tr}}$.

We will assume henceforth that $d = 2t + 1 \geq 3$ and that the number of errors is $w = wt_H(e) \leq t$. Let $i_1, i_2, \ldots, i_w$ be the positions of the errors. As usual, $\alpha_{i_1}, \ldots, \alpha_{i_w}$ are called the error locations and $e_{i_1}, \ldots, e_{i_w}$ the error magnitudes. We will treat the syndrome vector as a finite sequence.

Definition 3.8 (Syndrome sequence) *The syndrome sequence of the error e is the finite sequence $s|m$, with $m = 1 - 2t$, defined by:*

$$s_i = \sum_{k=1}^{n} e_k y_k \alpha_k^{-i} = \sum_{j=1}^{w} e_{i_j} y_{i_j} \alpha_{i_j}^{-i}.$$

Definition 3.9 (Error polynomials) *We define the error–locator and –evaluator polynomials by*

$$\sigma_e = \prod_{j=1}^{w} (X - \alpha_{i_j}) \text{ and } \omega_e = \sum_{j=1}^{w} e_{i_j} y_{i_j} \prod_{\substack{k=1,\ldots,w \\ k \neq j}} (X - \alpha_{i_k}).$$

If we know the coefficients of σ_e and ω_e then we can find the error locations by finding the roots of σ_e in R. From the definition of the error–locator and error–evaluator polynomial we get, for each unknown e_{i_j} a linear equation in R:

$$e_{i_j} \sigma'_e(\alpha_{i_j}) y_{i_j} = \omega_e(\alpha_{i_j})$$

Let us put $a_j = \sigma'_e(\alpha_{i_j}) y_{i_j}$ and $b_j = \omega_e(\alpha_{i_j})$, so that the equation is $a_j e_{i_j} = b_j$. We know that a solution of this equation does exist (we assumed the number of errors is at most t), hence $a_j | b_j$. Moreover, we can easily check that $a_j \in N(R)$, hence the equation cannot have more than one solution. This means that the quotient b_j/a_j is well defined and that we can compute the error magnitudes as:

$$e_{i_j} = \omega_e(\alpha_{i_j}) / (\sigma'_e(\alpha_{i_j}) y_{i_j}). \tag{3}$$

Remark 3.10 *In the classical literature σ_e^* and $X^{\delta\sigma_e - 1 - \delta\omega_e} \omega_e^*$ are called the error–locator and the error–evaluator polynomial respectively.*

We first set up an equation in $R[[X^{-1}]]$:

Definition 3.11 (Key equation) *Let $G = \sum_{m \le i \le 0} G_i X^i \in R[X^{-1}]$. We say that $(f,h) \in R[X] \times XR[X]$ is a solution of the key equation defined by G and X^{m-1} if f is monic, $\delta(h) \le \delta(f) \le -m$ and*

$$G \equiv h/f \bmod X^{m-1}. \tag{4}$$

A solution (f,h) is called minimal if $\delta(f)$ is minimal.

Remark 3.12 *Equation (4) is an analogue of the usual key equation in $R[[X]]$, see e.g. [MS77, Eqn. (68), p. 366]. A simpler, equivalent criterion for (f,h) to be a solution of Equation (4) is that $\delta(fG - h) < m + \delta(f)$ in $R((X^{-1}))$ (i.e. that (f,h) 'realize' $G_0, \ldots, G_m$ in the terminology of [Nor95]), but we have given an equation in $R[[X^{-1}]]$ to reflect the usual one.*

Lemma 3.13 *Suppose given an alternant code with parity check matrix H and designed distance $2t+1$. If $w = wt_H(e) \le t$ and $s|1-2t$ is the syndrome sequence, then $(\sigma_e, X\omega_e)$ solves the key equation defined by $\Gamma(s|1-2t)$ and X^{-2t}.*

The proof that $(\sigma_e, X\omega_e)$ solves our key equation (defined by $\Gamma(s|1-2t)$ and X^{-2t}) is similar to [MS77, Ch. 12,§9], and is omitted. However, the minimality of the solution $(\sigma_e, X\omega_e)$ is not obvious and will be proved in Section 4. We will also show that, unlike in the case of a field, the key equation does not necessarily have a unique solution. Nevertheless, it will turn out that *any* minimal solution can be used for determining the error polynomials and hence the error.

The connection between the key equation and minimal polynomials becomes clear from the following lemma:

Lemma 3.14 *The pair $(f,h) \in R[X] \times XR[X]$ is a minimal solution of the key equation defined by $\Gamma(s|m)$ and $X^{m-1} \Leftrightarrow \delta(f) \le -m,\ f \in Min(s|m)$ and $h = \beta(f, s|m)$.*

Proof. $\Rightarrow$: This is an easy consequence of the definitions. $\Leftarrow$: This is similar to the proof of [Nor95, Proposition 2.6] and is omitted. □

Hence we can concentrate on finding minimal polynomials.

4 Characterization of minimal polynomials

The purpose of this section is to characterize the (monic) minimal polynomials of the finite (and infinite) sequences which can be written as finite sums of geometric sequences. In particular, the syndrome sequence of an alternant code is of this type.

Let $w \ge 1$ and $a_1, \ldots, a_w,\ \gamma_1, \ldots, \gamma_w$ non–zero elements of R, with $\gamma_1, \ldots, \gamma_w$ all distinct. For the moment we will not impose other restrictions on the γ_i's. We define the sequence s by:

$$s_i = \sum_{j=1}^{w} a_j \gamma_j^{-i} \text{ for all } i \le 0.$$

Let $m = 1 - 2w$, so $s|m$ is the finite sequence consisting of the first $2w$ terms of s.

Throughout this section, we set $\sigma = \prod_{j=1}^{w}(X - \gamma_j)$.

The following is a more general form of Lemma 3.13. The proof is similar.

Lemma 4.1 *The polynomial σ satisfies:*
(i) $\sigma \in Ann(s)$

(ii) $\sigma\Gamma(s) = X\omega$ *where* $\omega = \sum_{j=1}^{w} a_j \prod_{\substack{k=1,\dots,w \\ k\neq j}} (X - \gamma_k)$

(iii) $\sigma'(\gamma_j) = \prod_{\substack{k=1,\dots,w \\ k\neq j}} (\gamma_j - \gamma_k)$ *for* $j = 1, \dots, w$

(iv) $\omega(\gamma_j) = a_j\sigma'(\gamma_j)$ *for* $j = 1, \dots, w$

(v) $\beta(\sigma, s| - w) = X\omega$.

Theorem 4.2 *If* $z_j \in R$ *are such that* $z_j a_j = 0$ *for* $j = 1, \dots, w$, *then* $f = \prod_{j=1}^{w}(X - \gamma_j - z_j) \in Ann(s)$.

Proof. Let u be the sequence given by $u_i = \sum_{j=1}^{w} a_j(\gamma_j + z_j)^{-i}$ for $i \leq 0$. By Lemma 4.1, $f \in Ann(u)$. If $z_j a_j = 0$ for $1 \leq j \leq w$, then $s_i = \sum_{j=1}^{w} a_j\gamma_j^{-i} = u_i$ for $i \leq 0$ and so $f \in Ann(u) = Ann(s)$. □

Note that with the assumptions so far, σ is not necessarily a minimal polynomial for s. For example, in $R = \mathbb{Z}_{p^2}$, let $w = 2$, $a_1 = a_2 = 1$, $\gamma_1 = 1 + p$, $\gamma_2 = 1 + 3p$. Then $s_i = (1 + p)^{-i} + (1 + 3p)^{-i} = 2 - 4ip = 2(1 + 2p)^{-i}$ and $\sigma = (X - 1 - p)(X - 1 - 3p)$ annihilates s, but so does $X - 1 - 2p$.

Lemma 4.3 *Let* $f \in Ann(s|m)$ *with* $\delta(f) \leq w$. *Then* $f(\gamma_j)a_j\gamma_j\sigma'(\gamma_j) = 0$ *for* $j = 1, \dots, w$.

Proof. By Lemma 2.13 we have $f\beta(\sigma, s|m) = \sigma\beta(f, s|m)$ since $\delta(f) + \delta(\sigma) \leq 2w = 1 - m$. By Lemmas 2.12 and 4.1, $\beta(\sigma, s|m) = \beta(\sigma, s| - w) = X\omega$. Hence $fX\omega = \sigma\beta(f, s|m)$. Evaluating this identity at γ_j we get $f(\gamma_j)\gamma_j\omega(\gamma_j) = 0$ since γ_j is a root of σ. Using Lemma 4.1*(iv)* we get the identity we are looking for. □

For the remainder of this section, we will assume that $\{\gamma_1, \dots, \gamma_w\} \subseteq N(R)$ *is subtractive.*

Theorem 4.4 *The following assertions are equivalent:*
(i) $\{f \in Ann(s|m) \setminus \{0\} : f \text{ of minimal degree}\} = \{r\sigma : r \in R \setminus \{0\}\}$
(ii) $Min(s|m) = \{\sigma\}$
(iii) $\{a_1, \dots, a_w\} \subseteq N(R)$.

Proof. (i) ⇒ (ii) When there are monic polynomials among the elements of $Ann(s|m)$ of minimal degree (as is the case in (i)), these polynomials are exactly the elements of $Min(s|m)$.

(ii) ⇒ (iii) Suppose without loss of generality that a_1 is a zero–divisor, with $z_1 \neq 0$ such that $a_1 z_1 = 0$. By Theorem 4.2, $f = (X - \gamma_1 - z_1)\prod_{j=2}^{w}(X - \gamma_j) = \sigma - z_1\prod_{j=2}^{w}(X - \gamma_j) \in Ann(s|m)$. The polynomial f is monic, of the same degree as σ but distinct from σ, so that (ii) fails.

(iii) ⇒ (i) Let $f \in Ann(s|m)$ be non–zero and of minimal degree. Then $\delta(f) \leq \delta(\sigma) = w$. By Lemma 4.3, $f(\gamma_j)a_j\gamma_j\sigma'(\gamma_j) = 0$ for $1 \leq j \leq w$. All of the factors in this expression except $f(\gamma_j)$ are in $N(R)$, so the relation simplifies to $f(\gamma_j) = 0$. Since $\{\gamma_1, \dots, \gamma_w\}$ is subtractive by Lemma 2.3, there is some $g \in R[X]$ such that $f = g\prod_{j=1}^{w}(X - \gamma_j) = g\sigma$, and g must be a constant since $\deg(f) \leq w$. □

In particular, the assertions in the theorem above always hold when R is a domain.

We note that Theorem 4.4 does not apply to the sequence of Example 2.11 since it is not a sum of geometric sequences.

Theorem 4.5 *Let* R *be a local ring,* $\overline{\gamma_j}$ *be distinct for* $1 \leq j \leq w$ *and let* $\mu \in Min(s|m)$. *Then*

$$\overline{\mu} = \overline{\sigma} = \prod_{j=1}^{w}(X - \overline{\gamma}_j).$$

PROOF. From Lemma 4.3 we know $\mu(\gamma_j)a_j\gamma_j \prod_{\substack{k=1,\ldots,w \\ k\neq j}} (\gamma_j-\gamma_k) = 0$. Since $\{\gamma_1,\ldots,\gamma_w\}$ is subtractive, all γ_j's and all differences $\gamma_j-\gamma_k$ are in $N(R)$. Hence the previous relation simplifies to $a_j\mu(\gamma_j) = 0$. Since $a_j \neq 0$, this implies $\mu(\gamma_j)$ is a zero–divisor and so $\overline{\mu}(\overline{\gamma}_j) = \overline{\mu(\gamma_j)} = 0$ for $j = 1,\ldots,w$. Since the $\overline{\gamma_j}$ are distinct, $\overline{\sigma}|\overline{\mu}$. Now μ is monic and of minimal degree, $\delta(\overline{\mu}) = \delta(\mu) \leq \delta(\sigma) = w$, so $\overline{\mu} = \overline{\sigma}$. □

Corollary 4.6 *If R is a domain or R is a local ring and the $\overline{\gamma_j}$ are distinct for $1 \leq j \leq w$, then (i) the complexity of s and of $s|m$ is w and (ii) $\sigma \in Min(s) \subseteq Min(s|m)$.*

The next Corollary follows from Theorem 4.2 and Corollary 4.6 *(i)*.

Corollary 4.7 *If R is a local ring, $\overline{\gamma_j}$ are distinct for $1 \leq j \leq w$, and $z_i \in R$ are such that $a_i z_i = 0$ for $1 \leq j \leq w$ then $\prod_{j=1}^{w}(X-\gamma_j-z_j) \in Min(s)$.*

The following theorem characterizes the polynomials in $Min(s|m)$ and $Min(s)$ when R is a Hensel ring, and in particular when R is a finite local ring.

Theorem 4.8 *If R is a Hensel ring and the $\overline{\gamma_j}$ are distinct for $1 \leq j \leq w$, then $\mu \in Min(s|m)$ iff $\mu = \prod_{j=1}^{w}(X-\gamma_j-z_j)$ for some z_j such that $z_j a_j = 0$ for $j = 1,\ldots,w$.*

PROOF. The sufficiency of the condition is Corollary 4.7. We now prove that the condition is necessary. Let $\mu \in Min(s|m)$. From Theorem 4.5 we know $\overline{\mu} = \prod_{j=1}^{w}(X-\overline{\gamma}_j)$. The factors of $\overline{\mu}$ are pairwise coprime because the $\overline{\gamma}_j$ are distinct. We can therefore lift $\overline{\mu}$ and factor μ as $\prod_{j=1}^{w}(X-\gamma_j-z_j)$ with $z_j \in M$. We still have to prove that $z_j a_j = 0$. From Lemma 4.3 we know that $\mu(\gamma_j)a_j\gamma_j \prod_{\substack{k=1,\ldots,w \\ k\neq j}} (\gamma_j-\gamma_k) = 0$. Evaluating $\mu(\gamma_j)$, we get:

$$-z_j a_j \gamma_j \prod_{\substack{k=1,\ldots,w \\ k\neq j}} (\gamma_j-\gamma_k) \prod_{\substack{l=1,\ldots,w \\ l\neq j}} (\gamma_j-\gamma_l-z_l) = 0$$

Since all $\overline{\gamma}_i$ are distinct, by Lemma 2.5 all differences $\gamma_j-\gamma_k$ are units. Also all $\gamma_j-\gamma_l-z_l$ are units and all γ_j are in $N(R)$, hence the relation above simplifies to $z_j a_j = 0$. □

Corollary 4.9 *If R is a domain or R is a Hensel ring and the $\overline{\gamma_j}$ are distinct for $1 \leq j \leq w$, then $Min(s|m) = Min(s)$.*

This result says intuitively that to find a minimal polynomial for the infinite sequence $s_i = \sum_{j=1}^{w} a_j\gamma_j^{-i}$ it suffices to know the first $2w$ consecutive terms of the sequence and to compute their minimal polynomial. This result is well–known over fields, but its extension to rings is not trivial.

Remarks 4.10
(i) A result similar to Theorem 4.5 is proven for Galois rings in [IPE97, Proposition 3 and the discussion following it]. Our result is more general since it is valid for arbitrary local rings, does not rely on the fact that the minimal polynomial has at least as many roots as the number of errors and that it is obtained by a particular algorithm. Nor do we use the theory of Linear Systems over a finite ring, [McD84].
(ii) Theorem 4.4 gives a different proof of the necessary and sufficient condition for the unicity of σ given in [IPE97, Appendix]. Our result holds in a more general context, as noted in the previous remark.
(iii) Lemma 4.3 can also be proved by Linear Algebra arguments, using [McD84, Theorem I.29].

5 Decoding algorithms

We continue our discussion on decoding alternant codes begun in Section 3, applying the new results proven in Section 4. We will consider two types of rings: domains and local rings.

The sequence of syndromes is of the same form as the sequence considered in Section 4, putting $\gamma_j = \alpha_{i_j}$ and $a_j = y_{i_j} e_{i_j}$ for $1 \leq j \leq w$. From Lemma 3.13, Lemma 3.14 and Corollary 4.6 we know that the error locator polynomial is a minimal polynomial for the syndrome sequence.

First we consider the case when R is a domain. For $\{\alpha_1, \ldots, \alpha_n\}$ to be subtractive, it is enough that the α_i's are distinct, since there are no non–trivial zero–divisors in R. We know from Theorem 4.4 that the error locator polynomial is the only monic minimal polynomial for the syndrome sequence. Algorithm MR of [Nor95] computes for *any sequence $s|m$ over a domain*, an annihilating polynomial f (not necessarily monic) of minimal degree and the corresponding $\beta(f, s|m)$. But again, from Theorem 4.4, we know that for a syndrome sequence, such a polynomial f must be the error locator multiplied by some non–zero constant. Hence, after applying algorithm MR to the sequence of syndromes we divide the output polynomials f and $\beta(f, s|m)$ by the leading coefficient of f obtaining thus σ_e and $X\omega_e$. The whole algorithm has quadratic complexity. We proceed as in the classical (field) case: we compute the roots of σ_e, which are of the form $\alpha_{i_1}, \ldots, \alpha_{i_w}$, giving the error locations. Then we compute the error magnitudes as $e_{i_j} = \omega_e(\alpha_{i_j})/(\sigma'_e(\alpha_{i_j}) y_{i_j})$.

Next we will consider the case of an alternant code $\mathcal{A}(\alpha, y, d)$ over a local ring R. We make the additional assumption that $\overline{\alpha}_1, \ldots, \overline{\alpha}_n$ are distinct. The definition of alternant codes only requires $\{\alpha_1, \ldots, \alpha_n\}$ to be subtractive, which is in general a weaker condition than $\overline{\alpha}_1, \ldots, \overline{\alpha}_n$ being distinct. For finite rings, however, the two conditions are equivalent (see Corollary 2.6). Note that all the codes in Examples 3.4, 3.5, 3.6 as well as the codes in Example 3.7 for R local satisfy the property that $\overline{\alpha}_1, \ldots, \overline{\alpha}_n$ are distinct.

We will also assume that for the ring R there is an algorithm of quadratic complexity ($O(m^2)$) which computes a minimal polynomial for any sequence $s|m$ of syndromes. We know that such algorithms exist for any sequence when R is a field (the well–known Berlekamp–Massey algorithm), $R = \mathbb{Z}_{p^a}$ (see [RS85]), and more generally, any finite chain ring, in particular Galois rings (see [Nor98, Algorithm MP]). We do not know whether such an algorithm exists for other local rings, for example for Hensel rings which are either infinite or have a non–principal maximal ideal.

We remark that the algorithm of [IPE97] is valid for a Galois ring, but may involve searching, so we do not know if it has quadratic complexity. See [IPE97, Conclusions, p. 1019].

From Corollary 4.7 we know that when some of the error magnitudes are zero–divisors the minimal polynomial is not unique. None of the algorithms above is guaranteed to find the error locator polynomial, but from Theorem 4.5, any minimal polynomial μ satisfies $\overline{\mu} = \overline{\sigma_e} = \prod_{j=1}^{w}(X - \overline{\alpha}_{i_j})$. Hence we have an even simpler method of finding the error locations. Namely, once we have some minimal polynomial μ (given by its coefficients), all we have to do is find the roots of $\overline{\mu}$ in K. From Theorem 4.5, they will be of the form $\overline{\alpha}_{i_1}, \ldots, \overline{\alpha}_{i_w}$, and $i_1, \ldots, i_w$ give the positions of the errors. We can then compute σ_e as $\sigma_e = \prod_{j=1}^{w}(X - \alpha_{i_j})$. To find the roots of $\overline{\mu}$ over the field K we can either use factorization algorithms for our particular K, or test the value of $\overline{\mu}$ at $\overline{\alpha}_1, \ldots, \overline{\alpha}_n$.

We summarise our proposed decoding algorithm for an alternant code with parity check matrix H and designed distance $2t + 1$ over a local ring R.

Algorithm 5.1 (Decoding an alternant code over a local ring)
Input: $r = (r_1, \ldots, r_n)$, the received vector, containing at most t errors.
Output: $c = (c_1, \ldots, c_n)$, the nearest codeword.

0. Let $m = 1 - 2t$.
1. Compute the syndrome sequence $s_i = (Hr^{\mathrm{tr}})_{1-i},\ m \leq i \leq 0$. If $s|m = (0, \ldots, 0)$, return r.
2. Compute a minimal polynomial μ for the sequence $s|m$.
3. Compute the roots $\overline{\alpha}_{i_1}, \ldots, \overline{\alpha}_{i_w}$ of $\overline{\mu}$ in K. Then the errors occurred at positions $i_1, \ldots, i_w$.

4. Compute $\sigma_e = \prod_{j=1}^{w}(X - \alpha_{i_j})$.
5. Compute σ'_e *and* $\omega_e = \beta(\sigma_e, s|m)/X$.
6. Set $e = (0, \ldots, 0)$ *and for* $j = 1, \ldots, w$, *put* $e_{i_j} = \omega_e(\alpha_{i_j})/(\sigma'_e(\alpha_{i_j})y_{i_j})$. *Return* $r - e$.

An example of our decoding procedure follows:

Example 5.2 *(cf. [IPE97, Example 2, p. 1019]) Let* $R = GR(9, 2) = \mathbb{Z}_9[y]/(y^2 + y + 2)$ *and* $\gamma = 2 - y$. *The successive powers of* γ *are* $2 + 4y,\ 3 + y,\ -1,\ y + 7,\ 5y + 7,\ 8y + 6$ *and* 1 *i.e.* γ *is a primitive* 8^{th} *root of* 1. *Let* C *be the BCH code of length* 8, *alphabet* $\mathbb{Z}_9$ *and designed distance* 5 *defined using* γ *as in Example 3.4.*

For the error $e = (0, 3, 0, 0, 0, 0, 6, 0)$, *we have the syndrome sequence* $3, 3y, 3, 3$. *Algorithm MP computes the minimal polynomial* $\mu = X^2 - yX - y$, *[Nor98, Example 7.12]. Thus* $\overline{\mu} = X^2 - yX - y = \mu$, *which has roots* $\overline{\gamma}$ *and* $\overline{\gamma}^6$ *in* $K = GF(3)[y]/(y^2 + y + 2) = GF(9)$.

We compute $\sigma_e = (X - \gamma)(X - \gamma^6) = X^2 + 5yX + 8y + 6$ *and* $\sigma'_e = 2X + 5y$. *Multiplication gives* $\omega_e = 3X$. *The error magnitudes are:*

$$e_2 = 3\gamma/(\gamma(2\gamma + 5y)) = 3/(1 + 3(y + 1)) = 3(1 - 3(y + 1)) = 3$$

and

$$e_7 = 3\gamma^6/(\gamma^6(2\gamma^6 + 5y)) = -3/(1 - 3(2y + 2)) = -3(1 + 3(2y + 2)) = -3 = 6.$$

Thus our algorithm returns the 0 *codeword, as expected.*

Remark 5.3 *Another method of computing the error locations once we have a minimal polynomial of the syndromes is proposed in [IPE97] for BCH and Reed–Solomon codes over a Galois ring. It involves computing the roots of a minimal polynomial in* R *and then determining which powers of* γ *differ from these roots by a zero–divisor (*cf. *Theorem 4.8).*

Our computation of the roots of $\overline{\mu}$ *in* K *is more efficient than the computation of roots in* R. *If we do it by searching, our approach requires the values of* $\overline{\mu}$ *at* n *points (viz.* $\alpha_1, \ldots, \alpha_n$*), whereas the approach of [IPE97] requires the values of a minimal polynomial at* $np^{l(a-1)}$ *points in the worst case (all the lifts of* $\alpha_1, \ldots, \alpha_n$*). Finding the roots of a polynomial* $f \in R[X]$ *by lifting requires the computation of the roots of* $\overline{f}$ *in* K *in the first place.*

The algorithms we described can decode, in particular, the examples of alternant codes we presented in Section 3. Namely, for decoding BCH codes over the p-adic integers (Example 3.6) we use the algorithm for domains described at the beginning of this section. For decoding BCH and Reed–Solomon codes over a Galois ring (Examples 3.4 and 3.5) and for generalised Reed–Solomon codes over a finite chain ring (Example 3.7), we use Algorithm 5.1.

Acknowledgement. The authors gratefully acknowledge financial support from the U.K. Engineering and Physical Sciences Research Council (EPSRC). The second author was supported by EPSRC Grant L07680.

References

[Bla75] I. A. Blake. Codes over integer residue rings. *Inform. Contr.*, 29:295–300, 1975.

[Coh89] P. M. Cohn. *Algebra*, volume 1. John Wiley & Sons, 1989.

[CS95] A. R. Calderbank and N. J. A. Sloane. Modular and p-adic codes. *Designs, Codes and Cryptography*, 6:21–35, 1995.

[Eis95] D. Eisenbud. *Commutative Algebra with a View toward Algebraic Geometry*, volume 150 of *GTM*. Springer Verlag, 1995.

[HKC+94] 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*, 40:301–319, 1994.

[IPE97] J. C. Interlando, R. Palazzo, and M. Elia. On the decoding of Reed-Solomon and BCH codes over integer residue rings. *IEEE Trans. Inform. Theory*, 43(3):1013–1021, 1997.

[McD74] B. R. McDonald. *Finite Rings with Identity*. Marcel Dekker, New York, 1974.

[McD84] B. R. McDonald. *Linear Algebra over Commutative Rings*. Marcel Dekker, New York, 1984.

[MS77] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-correcting Codes*. North Holland, Amsterdam, 1977.

[Nor95] G. H. Norton. On the minimal realizations of a finite sequence. *J. Symbolic Computation*, 20:93–115, 1995.

[Nor98] G. H. Norton. On minimal realization over a finite chain ring. *Designs, Codes and Cryptography*, 1998. Accepted for publication.

[RS85] J. A. Reeds and N. J. A. Sloane. Shift-register synthesis (modulo m). *SIAM J. Computing*, 14:505–513, 1985.

[Sha79] P. Shankar. On BCH codes over arbitrary integer rings. *IEEE Trans. Inform. Theory*, 25(4):480–483, 1979.

E-mail: Graham.Norton@bristol.ac.uk, Ana.Salagean@ntu.ac.uk