On minimal realization over a finite chain ring. (Designs, Codes and Cryptography, 16, 161–178,(1999)). *

Graham Norton

Algebraic Coding Research Group
Centre for Communications Research, University of Bristol, England.

July 19, 2001

Abstract

Let R be a finite chain ring, e.g. a Galois ring. We give a compact recursive formula for a minimal realization of a finite R–sequence. In particular, we show how to obtain a monic minimal polynomial and a rational approximation of a finite R–sequence. We also show how to solve the classical key equation of Algebraic Coding Theory over R.

Keywords: finite sequence, finite chain ring, rational approximation, key equation.

1 Introduction

We consider the problem of determining a 'minimal realization' of a finite sequence $s_0, \ldots, s_m$ over a commutative ring R (minimal realization was introduced in [11] and is related to rational approximation). For general commutative rings, division–based techniques (such as the Euclidean algorithm) fail. We show that minimal realization is possible over *finite* chain rings and their finite products. In particular, we can now do rational approximation over such rings. We show how to construct a minimal solution of the key equation $S\sigma \equiv \omega \pmod{g}$, where $g \in R[X]$ is monic (Algorithm 8.3).

Recall that a *chain ring* is a ring in which all its ideals are linearly ordered by inclusion, [7, p. 184]. It is well-known that a finite chain ring R is a local ring and that its maximal ideal M say, is principal. The nilpotency index ν of R and a (fixed) generator of M play an important role in

*Research supported in part by U.K. Science and Engineering Research Grant GR/H15141. Current addresses: Dept. Mathematics, University of Queensland, Brisbane 4072; ghn@maths.uq.edu.au.

our approach. It is also well–known that a Galois ring of characteristic p^e is a finite chain ring; in this case $\nu = e$ and p generates M.

For a characterization of finite chain rings, see Theorem 5.1. Suppose that F_q is a finite field, $f \in F_q[X]$ is reducible and factorises as $f_1^{n_1} \cdots f_k^{n_k}$, where $f_i \in F_q[X]$ is irreducible for $i = 1, \ldots, k$. Then the ring $F_q[X]/(f)$ is a finite product of finite chain rings, as of course is $\mathbb{Z}/(m)$. Trivially, any finite field is a finite chain ring.

Our approach is based on simplifying and generalizing [14] along the lines of [11], [12]. That is, we exploit $R[X^{-1}, X]$, its subrings $R[X^{-1}]$, $R[X]$ and deal with negatively indexed sequences. For $m \leq 0$, $s|m$ denotes the sequence $s_0, \ldots, s_m$ with last term s_m. It turns out that the case $m = 0$ is trivial and that the case $m < 0$ splits naturally into (i) $s|m+1$ 'has constant complexity' and (ii) $s|m+1$ 'does not have constant complexity'. The corresponding (polynomial) constructions are (i) Example 3.2 and Proposition 3.4 and (ii) Proposition 3.6. Theorems 6.2 and 6.4 establish the minimality of certain special constructions for cases (i) and (ii) respectively.

These theorems immediately yield Algorithm 6.6 which computes a minimal polynomial of a finite sequence over a finite chain ring. (An improved version is given in Section 7.3 below.) Algorithm 6.6 differs from [14] in a number of ways: (i) it is valid over more general rings; (ii) it computes a minimal polynomial at each iteration, without using pairs of polynomials; (iii) it is an immediate consequence of Theorem 6.4, not requiring separate 'theorem–proving' and 'computer–implementation' versions; (iv) our theory generalizes results of [11] when R is a finite field.

Section 7 extends Section 6 to minimal realizations ($MR's$) and contains the main result of this paper, Theorem 7.9. This theorem gives a compact formula for an MR of a finite sequence over a finite chain ring and improves Theorems 6.2, 6.4. In Section 7.3, we develop the corresponding compact algorithm to compute a minimal realization, Algorithm MR.

Restricting to first components in Algorithm MR yields an improved minimal polynomial algorithm, which we call Algorithm MP. This algorithm reduces to the monic version of Algorithm MP of [11] when R is a finite field. Proposition 7.13 tabulates the algebraic– and storage–complexity of Algorithms MP and MR. For example, Algorithm MP requires at most $\nu(1-m)^2$ R–multiplications when applied to $s|m$.

A 'Modified Berlekamp–Massey algorithm for shift–register synthesis' over a Galois ring appears in [6]. The authors note 'This procedure leads to a solution; however, not necessarily minimal. So one additional step had to be introduced to check for the minimality of the new solution. In the case where a nonminimal solution is obtained, a search among some candidate polynomials must be carried out', ([*loc. cit.*, Conclusions, p. 1019]). Example 7.12 is Algorithm MP is applied to [6, Example 2], which is a finite sequence over the Galois ring $R = GR(9, 2)$.

The final section discusses the key equation of Algebraic Coding Theory over a finite chain ring.

Possible topics for future work are: (i) decoding Reed–Solomon and BCH codes over a Galois ring;

(ii) studying the uniqueness of a minimal polynomial of $s|m$ over R; (iii) characterizing the set of minimal realizations $s|m$ over R; (iv) applications to algebraic number fields.

The author would like to thank Ana Sălăgean–Mandache for pointing out an error in a version of Theorem 6.2 for more general commutative rings and the referees for their detailed comments, corrections and suggestions. The author gratefully acknowledges financial support from the UK Engineering and Physical Sciences Research Council under grant GR/H15141.

2 Laurent Polynomials and Finite Sequences

We let R denote a commutative ring with $1 \neq 0$. The letters f, g, h denote $f, g, h \in R[X]$ and δf is the degree of f, with $\delta 0 = -\infty$.

As in [12], it is very convenient to work in the ring of Laurent polynomials $R[X^{-1}, X]$, which contains $R[X]$ as subring. Thus $R[X]$ acts on $R[X^{-1}, X]$ in the usual way (by multiplication in $R[X^{-1}, X]$). The letters F, G, H denote elements of $R[X^{-1}, X]$. For $-\infty < i < \infty$, F_i is the i^{th} coefficient of F.

The *support* of $F \neq 0$ is $\mathrm{Supp}(F) = \{i \in \mathbb{Z} : F_i \neq 0\}$, and $\mathrm{Supp}(0) = \emptyset$. Of course, if $\mathrm{Supp}(F) = \emptyset$, then $F = 0$.

We extend δ to $R[X^{-1}, X]$ by $\delta F = \max \mathrm{Supp}(F)$ if $F \neq 0$. If $F \neq 0$, $\lambda F \in R \setminus \{0\}$ is its *leading coefficient*. We have $\delta(FG) \leq \delta F + \delta G$ (with equality if $\lambda F \cdot \lambda G \neq 0$ e.g. if F is monic) and $\delta(F + G) \leq \max\{\delta F, \delta G\}$ (with equality if $\delta F \neq \delta G$).

The letters m, n always denote integers m, $n \leq 0$. We let $s|m$ denote the *finite* sequence with $s_i \in R$ for $m \leq i \leq 0$ and with last term s_m. We write $\Gamma(s|m) = \sum_{i=m}^{0} s_i X^i \in R[X^{-1}]$ for the 'generating polynomial' of $s|m$.

Definition 2.1 *For $l \in \mathbb{Z}$, we define $[F]_l^m \in R[X^{-1}]$ by*

$$[F]_l^m = \begin{cases} \sum_{i=l}^{m} F_i X^i & \text{if } l \leq m \\ 0 & \text{otherwise.} \end{cases}$$

3 Annihilators

3.1 The Annihilator Set

We extend some definitions and elementary results from [11, 12] which only require that R be commutative.

Definition 3.1 *Let $r \in R \setminus \{0\}$. The r–annihilator (or r–characteristic) set of $s|m$ is*

$$\mathrm{Ann}(s|m, r) = \{f \ : \ \lambda f = r, \ [f \cdot \Gamma(s|m)]_{m+\delta f}^{0} = 0\}.$$

If $\delta f \geq 1-m$, then $f \in \mathrm{Ann}(s|m, \lambda f)$ by Definition 2.1. Also, if $r \in R \setminus \{0\}$ and $m < 0$, then $\mathrm{Ann}(s|m,r) \subseteq \mathrm{Ann}(s|m+1,r)$. It is immediate that if

$$f_{0,r} = \begin{cases} r & \text{if } rs_0 = 0 \\ r \cdot X & \text{otherwise,} \end{cases}$$

then $f_{0,r} \in \mathrm{Ann}(s|0,r)$.

Example 3.2 *Let $m < 0$, $r \cdot s_i = 0$ for $m+1 \leq i \leq 0$ and $r \cdot s_m \neq 0$. Then $r \in \mathrm{Ann}(s|m+1,r) \setminus \mathrm{Ann}(s|m,r)$ and $X^{1-m} \in \mathrm{Ann}(s|m,1)$.*

If $d = \delta f$ and $1 \leq d \leq -m$, then

$$[f \cdot \Gamma(s|m)]_{m+d}^{0} = 0 \Leftrightarrow \lambda f \cdot s_{i-d} = -\sum_{j=0}^{d-1} f_j \cdot s_{i-j} \text{ for } m \leq i-d \leq -d$$

so for $\lambda f = 1$, $f \in \mathrm{Ann}(s|m,1)$ iff f 'generates $s_{-d}, \ldots, s_m$' from $s|(1-d)$.

It is convenient to define $(f \circ s|m)_i = (f \cdot \Gamma(s|m))_i$ for $m + \delta f \leq i \leq 0$. Note that $(f \circ s|m)_i = \sum_{k=0}^{\delta f} f_k \cdot s_{i-k}$ if $m + \delta f \leq i \leq 0$. Also, for $m + \delta f \leq 0$, we define the finite sequence $f \circ s|m$ to be $(f \circ s|m)_i$ for $m + \delta f \leq i \leq 0$. (The sequence $f \circ s|m$ will be used in Lemma 7.3 below.)

When m is understood and $m + \delta f \leq 0$, we write $(f \circ s)_i$ for $(f \circ s|m)_i$. We have

Proposition 3.3 *(i) If $m + \max\{\delta f, \delta g\} \leq i \leq 0$, then $((f+g) \circ s)_i = (f \circ s)_i + (g \circ s)_i$;*

(ii) If $r \in R$, $k \geq 0$ and $m + k + \delta f \leq i \leq 0$, then $((r \cdot X^k f) \circ s)_i = r \cdot (f \circ s)_{i-k}$.

3.2 Two constructions

Suppose now that $r \in R \setminus \{0\}$, $m < 0$, $f \in \mathrm{Ann}(s|m+1,r) \setminus \{0\}$ and $\delta f \leq -m$ (for example we could have $f = f_{0,r}$). Then $f \in \mathrm{Ann}(s|m,r)$ iff the obstruction to extending f to $s|m$

$$\mathcal{O}f \stackrel{\text{def}}{=} (f \circ s)_{m+\delta f}$$

is zero; $\mathcal{O}f$ is called the *discrepancy* of f in the shift–register literature. We conventionally put $\mathcal{O}f = 0$ if $\delta f \geq 1-m$ since such an f is always in $\mathrm{Ann}(s|m,r)$.

Thus in Example 3.2, $f = r \in \mathrm{Ann}(s|m+1,r)$ but $f \notin \mathrm{Ann}(s|m,r)$ since $\mathcal{O}f = r \cdot s_m \neq 0$. Proposition 3.4, which is our second example, is suggested by the equality $\mathcal{O}f - ((\mathcal{O}f/s_0) \circ s)_0 = 0$ when $\mathcal{O}f \neq 0$ and $s_0 | \mathcal{O}f$. The easy verification is omitted.

Proposition 3.4 *Let $m < 0$ and $f \in \mathrm{Ann}(s|m+1, \lambda f)$. If $s_0 \neq 0$, $\mathcal{O}f \neq 0$ and $s_0 | \mathcal{O}f$, then*

$$f_m = X^{-m-\delta f} f - \mathcal{O}f/s_0 \in \mathrm{Ann}(s|m, \lambda f) \setminus \{0\}.$$

The next construction is suggested by the equality $[\mathcal{O}f, \mathcal{O}g] = 0$, and will be used in the non-constant complexity case (i.e. in Theorem 6.4 below):

Definition 3.5 *([11, Definition 3.10]) Let* $m+1 < n \leq 0$, $f \in \mathrm{Ann}(s|m+1, \lambda f)$ *and* $g \in \mathrm{Ann}(s|n, \lambda g)$. *If* $\mathcal{O}f \neq 0$ *and* $\mathcal{O}g \neq 0$, *we define* $\delta = \delta(f,g) = \max\{\delta f, -m+n-1+\delta g\}$ *and*

$$[f,g] = \mathcal{O}g \cdot X^{\delta - \delta f} f - \mathcal{O}f \cdot X^{\delta + m - n + 1 - \delta g} g.$$

The next result is essentially [11, Proposition 3.11], and so we omit the straightforward proof:

Proposition 3.6 *Let* $m+1 < n \leq 0$, $f \in \mathrm{Ann}(s|m+1, \lambda f)$, $g \in \mathrm{Ann}(s|n, \lambda g)$ *and* $\mathcal{O}f \neq 0$, $\mathcal{O}g \neq 0$. *If* $\lambda f \cdot \mathcal{O}g \neq 0$ *then* $\lambda[f,g] = \lambda f \cdot \mathcal{O}g$, $\delta[f,g] = \delta \leq -m$ *and* $[f,g] \in \mathrm{Ann}(s|m, \lambda f \cdot \mathcal{O}g)$. *If in addition* $\mathcal{O}g | \mathcal{O}f$, *then* $\lambda([f,g]/\mathcal{O}g) = \lambda f$ *and* $[f,g]/\mathcal{O}g \in \mathrm{Ann}(s|m, \lambda f)$.

4 Minimality

We write $\mathrm{Min}(s|m, r)$ for the polynomials in $\mathrm{Ann}(s|m, r) \setminus \{0\}$ of minimal degree, with typical element $\mu_{m,r} \in \mathrm{Min}(s|m, r)$. We can take $\mu_{0,r} = f_{0,r}$ as defined in Section 3.1. Our goal is to determine a $\mu_{m,r}$, given $m < 0$, an $s|m$ and an $r \in R \setminus \{0\}$.

We call the unique minimal degree of a $\mu_{m,r}$ the *r–complexity* of $s|m$, written $\kappa_{m,r}$ when $s|m$ is understood. Clearly $\kappa_{m,r} \leq 1 - m$ for any $r \in R$ and since $\mathrm{Ann}(s|m-1) \subseteq \mathrm{Ann}(s|m)$, $\kappa_{m,r} \leq \kappa_{m-1,r}$. If R is a domain, then $\kappa_{m,r}$ is independent of r. In Example 3.2, $\kappa_{m+1,r} = 0$, and we will see below that if $ur \cdot s_m \neq 0$, then $\kappa_{m,u} = 1 - m$.

We give a key extension of the Minimality Lemma ([11, Lemma 4.2]) to commutative rings (Lemma 4.3 below). Recall from [11] that the *border polynomial* of f and $s|m$ is

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

where $f_j = 0$ for $j > \delta f$. For example, $\beta(X^{1-m}, s|m) = \sum_{i=1}^{1-m} s_{i-1+m} X^i$. Also, $\delta\beta(f, s|m) \leq \delta f$, X divides $\beta(f, s|m)$ and $\beta(r, s|m) = 0$ if $r \in R$. It is clear that $\beta(\ , s|m)$ is R–linear i.e. that $\beta(uf + vg, s|m) = u\beta(f, s|m) + v\beta(g, s|m)$ for any $u,\ v \in R$.

The foregoing definitions easily yield:

Proposition 4.1 $f\Gamma(s|m) = F + [f\Gamma(s|m)]_{m+\delta f}^{0} + \beta(f, s|m)$ *for a unique* $F \in R[X^{-1}]$ *satisfying* $\delta F < m + \delta f$.

As in [12], Proposition 4.1 implies that if $f \in \mathrm{Ann}(s|m, 1)$, then $\delta(\Gamma(s|m) - \beta(f, s|m)/f) < m$ i.e. $\beta(f, s|m)/f$ is an order m *rational approximation* of $\Gamma(s|m)$.

Corollary 4.2 *If* $m < 0$, $f \in \mathrm{Ann}(s|m+1, \lambda f)$, *then*

$$f\Gamma(s|m) = F + \mathcal{O}f \cdot X^{m+\delta f} + \beta(f, s|m)$$

for a unique $F \in R[X^{-1}]$ *satisfying* $\delta F < m + \delta f$.

The next lemma is pivotal for constructing minimal polynomials and indicates a need to keep track of leading coefficients when there are zero–divisors:

Lemma 4.3 *(cf. [14], [12, Lemma 4.2])*

Let $m < 0$ *and* $f \in \mathrm{Ann}(s|m+1, \lambda f) \setminus \{0\}$, $g \in \mathrm{Ann}(s|m+1, \lambda g) \setminus \{0\}$. *Then*

(i) if $f \in \mathrm{Ann}(s|m, \lambda f)$ *and* $\lambda f \cdot \mathcal{O}g \neq 0$, *then* $\delta f \geq 1 - m - \delta g$;

(ii) if $\delta f + \delta g \leq -m$, *then* $\lambda f \cdot \mathcal{O}g = \lambda g \ \cdot \mathcal{O}f$.

Proof. Let $h = g\beta(f, s|m) - f\beta(g, s|m) \in XR[X]$, so that $\mathrm{Supp}(h) \subseteq [1, \delta h]$ and let $d = m + \delta f + \delta g$. Expanding h via Corollary 4.2 yields $h_d = \lambda f \cdot \mathcal{O}g - \lambda g \cdot \mathcal{O}f$. Parts (i) and (ii) are now simple consequences of $h_d \neq 0 \Rightarrow d \geq 1$. □

Remark 4.4 *In the preceding proof,* $h_d \neq 0 \Rightarrow \delta h = d$, *but we will not need this fact.*

5 Finite chain rings

The following result is well–known (see e.g. [1, 9]):

Theorem 5.1 *Let* R *be a finite commutative local ring with maximal ideal* M *and residue field* K. *The following are equivalent:*

(i) R *is a finite chain ring;*

(ii) R *is the homomorphic image of* $A[X]$ *given in [9, p. 342], where* A *is a Galois ring;*

(iii) M *is principal;*

(iv) $\dim_K M/M^2 \leq 1$.

Proof. The equivalence of (i) and (ii) is the characterization theorem for finite chain rings [9, Theorem XVII.5, p. 342]; [1, Proposition 8.8, p. 91] gives the equivalence of (i), (iii) and (iv), since any finite ring is Artinian. □

Convention: For the remainder of this paper, R denotes a *finite* commutative chain ring with principal maximal ideal M and nilpotency index ν. We fix a generator γ of M.

For example, a finite commutative local ring with characteristic p^e is a Galois ring if and only if its maximal ideal is pR [9, Exercise XVI.9(a), p. 332]. Thus if R is a Galois ring with characteristic p^e, we can take $\gamma = p$ and have $\nu = e$.

The following form of unique factorization in R plays an important role in our approach:

Proposition 5.2 *Any element $r \in R \setminus \{0\}$ can be written as $r = u\gamma^t$ where u is a unit of R, t is unique and $0 \leq t \leq \nu - 1$.*

PROOF. Let $r \in R \setminus \{0\}$. If r is a unit, we take $t = 0$. If r is not a unit, $r \in M$ by [1, Corollary 1.4]. In fact, since R is finite, $r \in M^t \setminus M^{t+1}$ for some t, $1 \leq t \leq \nu - 1$, i.e. $r = u\gamma^t$ for some $u \in R \setminus M$. This implies that $(u) = R$, so u is a unit of R.

To show that t is unique, suppose that $up^s = vp^t$ for units u, v and $0 \leq s, t \leq \nu - 1$. We can assume that both s and t are strictly positive. If $s > t$ and $uu' = 1$, then $u\gamma^t(\gamma^{s-t} - u'v) = 0$, where $\gamma^{s-t} - u'v$ is a unit — otherwise $u'v \in M$, which is impossible. This implies that $\gamma^t = 0$ where $t < \nu$, which is a contradiction. Similarly $s < t$ is impossible and we conclude that $s = t$. □

If $r \in R \setminus \{0\}$ and $r = u\gamma^t$, we write $\log r$ for the uniquely defined power t of γ, where $0 \leq \log r \leq \nu - 1$. Since R is finite, any element of $R \setminus \{0\}$ is either a unit or a zero–divisor (see e.g. [9, Exercise I.8, p. 5]). It is easy to check that if a, b are zero–divisors of $R \setminus \{0\}$, then $a|b$ iff $\log a \leq \log b$.

Finite chain rings also appear as quotients of $F_q[X]$ as follows:

Proposition 5.3 *Let F_q be a finite field, $f \in F_q[X]$ irreducible and $e \geq 2$. Then $R = F_q[X]/(f^e)$ is a finite chain ring with proper ideals $f^i R, 1 \leq i \leq e - 1$, and unique maximal ideal fR.*

PROOF. Ideals of R are ideals of $F_q[X]$ which contain (f^e). Thus the first statement follows from the fact that $F_q[X]$ is a principal ideal domain and hence is a factorial ring. The second statement is an easy application of the double quotient isomorphism theorem. □

6 Minimal Polynomials

In what follows, we set $\ell = \gamma^{\nu-1}$. Thus if R is a Galois ring of characteristic p^e, then $\ell = p^{e-1}$. In particular, if R is a finite field of characteristic p then $\ell = 1$.

Recall that $a, b \in R \setminus \{0\}$ are associates if there is a unit $u \in R \setminus \{0\}$ such that $a = ub$. This defines an equivalence relation: $[r]$ denotes the class of associates of $r \in R \setminus \{0\}$ and $[R \setminus \{0\}]$ denotes the set of associate classes of $R \setminus \{0\}$. We will use $[\gamma^i]$, $0 \leq i \leq \nu - 1$ as the equivalence classes in $[R \setminus \{0\}]$.

Since $r = ut$ for some unit $u \in R$ for each $t \in [r]$, $\mathrm{Min}(s|m, r) = u \cdot \mathrm{Min}(s|m, t)$ where $u \cdot$

$\mathrm{Min}(s|m,t) = \{u \cdot f \ : f \in \mathrm{Min}(s|m,t)\}$. For this reason, it suffices to determine $\mathrm{Min}(s|m,[r])$ for each $r \in R \setminus \{0\}$ and thus *from now on, we will assume that r denotes an associate class.*

We need one more item of notation before stating the main theorems:

Definition 6.1 *For $r \in [R \setminus \{0\}]$, we define $r^* = [\gamma^{\nu-1-\log r}]$.*

For $r \in [R \setminus \{0\}]$, r^* is well–defined by Proposition 5.2, and it is easy to see that $rr^* = [\ell]$. If R is a Galois ring of characteristic p^2 (e.g. $R = \mathbb{Z}_4$) the map $r \to r^*$ simply interchanges $[1]$ and $[p]$ and if R is a finite field, $r^* = [1]$ for all $r \in [R \setminus \{0\}]$.

6.1 The Minimality Theorems

From now on, we will abbreviate $\mathcal{O}\mu_{i,r}$ to $\mathcal{O}_{i,r}$ for $i \leq 0$ and $r \in [R \setminus \{0\}]$. First we treat the constant complexity case:

Theorem 6.2 *(cf. [14].) Let $m < 0$, $r \in [R \setminus \{0\}]$, $\mu = \mu_{m+1,r}$, $\mathcal{O}_{m+1,r} \neq 0$ and $c = (\mathcal{O}_{m+1,r})^*$. If $\kappa_{m+1,c} = \kappa_{0,c}$ and*

$$\mu_{m,r} = \begin{cases} rX^{1-m} & \text{if } \kappa_{0,c} = 0 \\ X^{-m-\delta\mu}\mu - \mathcal{O}_{m+1,r}/s_0 & \text{otherwise,} \end{cases}$$

then

(i) $\mu_{m,r}$ is well–defined and $\mu_{m,r} \in \mathrm{Ann}(s|m,r)$;

(ii) If $\delta\mu_{m,r} > \delta\mu_{m+1,r}$ then (a) $\delta\mu_{m,r} = 1 - m - \delta\mu_{m+1,c}$ and (b) $[\mathcal{O}_{m+1,c}] = r^*$;

(iii) $\mu_{m,r} \in \mathrm{Min}(s|m,r)$.

Proof. Suppose first that $\kappa_{0,c} = 0$. Certainly $\delta\mu_{m,r} = 1 - m$, so that $\mu_{m,r} \in \mathrm{Ann}(s|m,r)$.

Part (ii)(a) is clear. We have $\delta\mu_{m+1,r} \leq -m$ and so $\delta\mu_{m+1,r} + \delta\mu_{m+1,c} = \delta\mu_{m+1,r} \leq -m$. Lemma 4.3 (ii) implies that $r \cdot \mathcal{O}_{m+1,c} = c \cdot \mathcal{O}_{m+1,r} = \ell$. This implies (ii)(b) and since $\ell \neq 0$, $\kappa_{m,r} \geq 1 - m$ by Lemma 4.3(i). Thus $\mu_{m,r} \in \mathrm{Min}(s|m,r)$.

Suppose now that $\kappa_{0,c} = 1$. We have $t = c \cdot s_0 \neq 0$ — otherwise $\delta\mu_{0,c} = 0$ — and so $s_0\ell = t \cdot \mathcal{O}_{m+1,r}$. If t is a unit, $s_0 | \mathcal{O}_{m+1,r}$. If not, $\log t$ is defined and $\log s_0 + \nu - 1 = \log \mathcal{O}_{m+1,r} + \log t$. Hence $\log s_0 \leq \log \mathcal{O}_{m+1,r}$ which implies that $s_0 | \mathcal{O}_{m+1,r}$. Thus $\mu_{m,r}$ is well–defined, and $\mu_{m,r} \in \mathrm{Ann}(s|m)$ by Proposition 3.4.

It remains to prove (ii) and show that $\delta\mu_{m,r} = -m$ is minimal. If $\delta\mu_{m+1,r} = -m$ then $\delta\mu_{m,r}$ is already minimal. Suppose now that $\delta\mu_{m+1,r} < -m$. Then $\delta\mu_{m,r} = -m = 1 - m - \delta\mu_{m+1,c}$ which proves (ii)(a). Also, $\delta\mu_{m+1,r} + \delta\mu_{m+1,c} \leq -m - 1 + 1 = -m$ and so by Lemma 4.3(ii)

$$r \cdot \mathcal{O}_{m+1,c} = c \cdot \mathcal{O}_{m+1,r} = \ell$$

which proves (ii)(b). To prove (iii), we have $r \cdot \mathcal{O}_{m+1,c} = \ell \neq 0$ and so by Lemma 4.3(i) applied to $g = \mu_{m+1,c}$, we deduce that $\kappa_{m,r} \geq 1 - m - \delta\mu_{m+1,c} = -m$ and so $\mu_{m,r} \in \mathrm{Min}(m, r)$. □

Remark 6.3 *We may also prove part (ii)(b) for the case $\kappa_{0,c} = 0$ directly. Using the same notation as Theorem 6.2, $\mu \in \mathrm{Ann}(s|m+1, r)$ implies that*

$$\delta(\mu\Gamma(s|m) - \beta(\mu, s) - \mathcal{O}_{m+1,r} X^{m+\delta\mu}) < m + \delta\mu$$

by Corollary 4.2. Now $c\, s_j = 0$ for $m+1 \leq j \leq 0$, so multiplying by c and equating coefficients of $X^{m+\delta\mu}$ gives $r \cdot \mathcal{O}_{m+1,c} = rcs_m = c \cdot \mathcal{O}_{m+1,r} = \ell$. (cf. [14, Case IIa, p.511].)

Now we treat the remaining (non–constant complexity) case. In the following theorem, we think of the index $\alpha_{m+1,c}$ of $s|m$ as the 'antecedent' of $\kappa_{m+1,c}$.

Theorem 6.4 *(cf. [14].) Suppose that $m + 1 < 0$, $r \in [R \setminus \{0\}]$, $\mu_{n,a} \in \mathrm{Min}(s|n, a)$ for all $a \in [R \setminus \{0\}]$ and $m + 1 \leq n \leq 0$. Let $\mathcal{O}_{m+1,r} \neq 0$ and $c = (\mathcal{O}_{m+1,r})^*$. If $\kappa_{m+1,c} > \kappa_{0,c}$, define*

$$\alpha = \alpha_{m+1,c} = \min_{m+1<n\leq 0}\{n \ : \ \delta\mu_{n,c} < \delta\mu_{m+1,c}\}$$

$$d = (\mathcal{O}_{\alpha,c})^*$$

$$\mu_{m,r} = [\mu_{m+1,r},\ \mu_{\alpha,d}]/\mathcal{O}_{\alpha,d}.$$

Then

(i) $\mu_{m,r}$ is well–defined and $\mu_{m,r} \in \mathrm{Ann}(s|m, r)$;

(ii) If $\delta\mu_{m,r} > \delta\mu_{m+1,r}$ then (a) $\delta\mu_{m,r} = 1 - m - \delta\mu_{m+1,c}$ and (b) $[\mathcal{O}_{m+1,c}] = r^*$;

(iii) $\mu_{m,r} \in \mathrm{Min}(s|m, r)$.

Proof. Suppose inductively that the result is true for $s|n$, where $m + 1 \leq n < 0$. That is, for all $a \in [R \setminus \{0\}]$, $\mu_{n,a} \in \mathrm{Min}(s|n, a)$ satisfies: if $\mathcal{O}_{n+1,a} \neq 0$ and $b = (\mathcal{O}_{n+1,a})^*$ then $\delta\mu_{n,a} > \delta\mu_{n+1,a}$ implies (a) $\delta\mu_{n,a} = 1 - n - \delta\mu_{n+1,b}$ and (b) $[\mathcal{O}_{n+1,b}] = a^*$.

By construction, $m+1 < \alpha \leq 0$ and $\delta\mu_{m+1,c} = \delta\mu_{\alpha-1,c} > \delta\mu_{\alpha,c}$. Thus $\delta\mu_{\alpha-1,c} = 1-(\alpha-1)-\delta\mu_{\alpha,d}$ and $[\mathcal{O}_{\alpha,d}] = c^* = (\mathcal{O}_{m+1,r})^{**} = \mathcal{O}_{m+1,r}$ either by the inductive hypothesis (if $\delta\mu_{\alpha,d} > \kappa_{0,d}$) or by Theorem 6.2 (if $\delta\mu_{\alpha,d} = \kappa_{0,d}$). If $\mathcal{O}_{\alpha,d}$ is a unit, then $\mathcal{O}_{\alpha,d}|\mathcal{O}_{m+1,r}$; otherwise $\log \mathcal{O}_{\alpha,d} = \log \mathcal{O}_{m+1,r}$ and so $\mathcal{O}_{\alpha,d}|\mathcal{O}_{m+1,r}$. Thus $\mu_{m,r}$ is well–defined and $\mu_{m,r} \in \mathrm{Ann}(s|m, r)$ by Proposition 3.6.

We now prove part (ii). If $\delta\mu_{m,r} > \delta\mu_{m+1,r}$ then

$$\delta\mu_{m,r} \quad = \quad -m + \alpha - 1 + \delta\mu_{\alpha,d}$$

$$
\begin{aligned}
&= -m + \alpha - 1 + (2 - \alpha - \delta\mu_{\alpha-1,c}) \\
&= 1 - m - \delta\mu_{\alpha-1,c} \\
&= 1 - m - \delta\mu_{m+1,c} \text{ by construction, which proves (ii)(a).}
\end{aligned}
$$

To prove (ii)(b), we first show that $\delta\mu_{m+1,r} + \delta\mu_{m+1,c} \leq -m$. Our hypothesis implies that the left–hand side is at most $(\delta\mu_{m,r} - 1) + \delta\mu_{m+1,c}$ which we have just seen is equal to $-m$. As before, it follows from Lemma 4.3(ii) that $r \cdot \mathcal{O}_{m+1,c} = c \cdot \mathcal{O}_{m+1,r} = u\ell$, u a unit of R, which yields (ii)(b).

We now prove part (iii). If $\delta\mu_{m,r} = \delta\mu_{m+1,r}$, then $\delta\mu_{m,r}$ is already minimal, so suppose that $\delta\mu_{m,r} > \delta\mu_{m+1,r}$. Part (ii)(b) implies that $r \cdot \mathcal{O}_{m+1,c} \neq 0$ and so Lemma 4.3(i) gives $\kappa_{m,r} \geq 1 - m - \delta\mu_{m+1,c}$. The latter is $\delta\mu_{m,r}$ by part (ii) (a). Hence $\mu_{m,r} \in \mathrm{Min}(s|m, r)$, which completes the proof.

□

Remark 6.5 *When R is a finite field, we can suppress the subscripts r, c, d since* [1] *is the only associate class. There is also no need to divide by s_0 in Theorem 6.2 or $\mathcal{O}_{\alpha,d}$ in Theorem 6.4. In this case, Theorems 6.2 and 6.4 reduce to Propositions 5.3 and 5.4 of [12].*

6.2 A naive algorithm

Our first minimal polynomial algorithm is an immediate consequence of Theorems 6.2, 6.4. For $i + 1 < 0$, $\alpha_{i+1,r}$ is an integer satisfying $i + 1 < \alpha_{i+1,r} \leq 0$. Thus we can assume that the minimal polynomials $\mu_{\alpha_{i+1,r},r}$ are either $f_{0,r}$ (following Definition 3.1) or have been obtained inductively (according to Theorems 6.2 or 6.4). In this preliminary version, we do not update these polynomials as in [11, 12]; instead we assume that the values of $\alpha_{i+1,r}$ and the polynomials $\mu_{\alpha_{i+1,r},r}$ are available as needed.

Algorithm 6.6 *(Minimal polynomial—naive version)*

Input: $m \leq 0$, *a finite chain ring* R, $s_0, \ldots, s_m \in R$.

Output: $\mu_{m,r} \in \mathrm{Min}(s|m, r), r \in [R \setminus \{0\}]$.

for $r \in [R \setminus \{0\}]$ *do*
  { *if* $(r \cdot s_0 = 0)$ *then* $\mu_{0,r} := r$;
    else $\mu_{0,r} := r \cdot X$;
  $\kappa_{0,r} := \delta\mu_{0,r}$;}

for $i := -1$ *to* m *do*
  for $r \in [R \setminus \{0\}]$ *do*

/* *Compute* $\mu_{i,r}$ */

{ $\kappa = \delta\mu_{i+1,r}$;

$\mathcal{O} := (\mu_{i+1,r} \circ s)_{i+\delta\mu_{i+1,r}}$;

if $(\mathcal{O} = 0)$ *then* $\mu_{i,r} := \mu_{i+1,r}$;

else { $c := \mathcal{O}^*$;

case:

$(\kappa = \kappa_{0,c} = 0)$: $\mu_{i,r} := rX^{1-i}$;

$(\kappa = \kappa_{0,c} = 1)$: $\mu_{i,r} := X^{-i-\delta\mu_{i+1,r}}\ \mu_{i+1,r} - \mathcal{O}/s_0$;

$(\kappa > \kappa_{0,c})$: { $\alpha := \alpha_{i+1,c}; d := (\mathcal{O}_{\alpha,c})^*$;

$\mu_{i,r} := [\mu_{i+1,r},\ \mu_{\alpha,d}]/\mathcal{O}_{\alpha,d}$; } *endcase*}}

for $r \in [R \setminus \{0\}]$ *do return* $\mu_{m,r}$.

Example 6.7 *Algorithm 6.6 yields the following table of minimal polynomials for the* $\mathbb{Z}_9$*–sequence* $(s| - 4) = 6, 3, 1, 5, 6$ *of [14, p 512]:*

i	$\mathcal{O}_{i,1}$	$\mu_{i,1}$	$\alpha_{i,1}$	$\mathcal{O}_{i,3}$	$\mu_{i,3}$	$\alpha_{i,3}$
0	6	X	–	0	3	–
-1	3	$X + 4$	–	0	3	–
-2	4	X^3	-1	3	$3X^2 - 5$	-1
-3	5	$X^3 + X^2 + 4X$	-1	0	$3X^2 - 5$	-1
-4	6	$X^3 + X^2 + 7X$	–	4	$3X^3 + 3X + 5$	–

Remarks 6.8 *(i) The* $\mu_{i,r}$ *are obtained without computing the border polynomials* $\beta(\mu_{i,r}, s|i)$ *at each iteration. (ii) We obtain minimal polynomials at each iteration of the algorithm. (iii) If* R *is a finite field, all the* r *indices,* c *and* d *can be suppressed, and Algorithm 6.6 reduces to the monic version of Algorithm MP of [11] (cf. [14, Algorithm]).*

Suppose that $R \cong R_1 \times R_2$ and each R_i is a finite chain ring. A sequence $s|m$ over R gives rise to sequences $s^{(i)}|m$ and corresponding monic minimal polynomials $\mu^{(i)}_{m,1}$ in each factor, $i = 1, 2$. This readily yields a monic minimal polynomial for $s|m$. The details are elementary and omitted. Since $F_q[X]$ is factorial, it now follows from Proposition 5.3 that we can also compute a monic minimal polynomial of a sequence over $F_q[X]/(f)$ for any $f \in F_q[X]$.

7 Minimal Realizations

7.1 Realizations

Definition 7.1 *Let $r \in [R \setminus \{0\}]$. For $f \neq 0$, we say that $(f,g) \in R[X] \times XR[X]$* is an *$r$–realization of $s|m$ if $\lambda f = r$, $\delta g \leq \delta f$ and $\delta(f\Gamma(s|m) - g) < m + \delta f$. In addition, (f,g) is a* minimal *r–realization of $s|m$ (written $(f,g) \in MR(s|m,r)$) if (f,g) is an r–realization of $s|m$ and $f \in \mathrm{Min}(s|m,r)$.*

We refer the reader to [11] for a discussion of minimal realization. In particular, if $\delta f > -m$, we always have $\delta(f\Gamma(s|m) - \beta(f,s|m)) \leq 0 < m + \delta f$ and *any* $(f,\beta(f,s|m))$ with $\lambda f = r$ and $\delta f > -m$ is an r–realization of $s|m$, e.g. $(rX^{1-m}, \beta(rX^{1-m}, s|m))$. Addition and polynomial multiplication of realizations will be componentwise. The presence of zero–divisors does not affect the validity of the next result ([11, Propositions 2.6, 2.8]):

Proposition 7.2 *Let $\delta f \leq -m$. If (f,g) is an r–realization of $s|m$ then $g = \beta(f,s|m)$, and the following are equivalent:*

(i) $f \in \mathrm{Ann}(s|m,r)$;

(ii) (f,g) is an r–realization of $s|m$ for some g;

(iii) $(f,\beta(f,s|m))$ is an r–realization of $s|m$.

Thus for any $r \in [R \setminus \{0\}]$, an r–realization of $s|m$ can be obtained by finding a $\mu_{m,r} \in \mathrm{Min}(s|m,r)$ and computing $\beta(\mu_{m,r}, s|m)$. From now on, we write $\beta_{m,r}$ for $\beta(\mu_{m,r}, s|m)$. Also, it is easily verified that since $\delta\mu_n \leq 1-n$, $\beta(\mu_{n,r}, s|n) = \beta(\mu_{n,r}, s|m)$ for any $m \leq n$, so we can further simplify the notation by writing $\beta_{n,r}$ for any $\beta(\mu_{n,r}, s|m)$ with $m \leq n$.

It is elementary that

$$\beta(\mu_{0,r}, s|0) = \begin{cases} 0 & \text{if } r \cdot s_0 = 0 \\ r \cdot s_0 \cdot X & \text{otherwise.} \end{cases}$$

Since $\mu_{m,r}$ is obtained using products of polynomials, a product formula (*cf.* [11, Proposition 2.3(e)], [12, Lemma 5.2]) is needed for expressing the polynomials $\beta_{m,r}$ recursively. Recall that for $\delta g \leq -m$, the finite sequence $g \circ s|m$ was defined in Section 3.

Lemma 7.3 *(Product Formula) If $\delta f + \delta g \leq -m$, then $\beta(f \cdot g, s|m) = f \cdot \beta(g, s|m) + \beta(f, g \circ s|m)$.*

Corollary 7.4 *(cf. [14].) Let $m < 0$, $r \in [R \setminus \{0\}]$, $\mu = \mu_{m+1,r}$, $\mathcal{O}_{m+1,r} \neq 0$ and $c = (\mathcal{O}_{m+1,r})^*$. If $\kappa_{m+1,c} = \kappa_{0,c}$ then*

$$\beta_{m,r} = \begin{cases} r \cdot s_m \cdot X & \text{if } \kappa_{0,c} = 0 \\ X^{-m-\delta\mu}\beta_{m+1,r} & \text{otherwise.} \end{cases}$$

PROOF. Suppose that $\kappa_{0,c} = 0$. By definition, $m+1$ is the first index with $\mathcal{O}_{m+1,r} \neq 0$, so $r \cdot s_0 = \ldots = r \cdot s_{m+1} = 0$ and $\beta_{m,r} = \beta(r \cdot X^{1-m}, s|m) = \sum_{i=1}^{1-m} r \cdot s_{i-1+m} \cdot X^i = r \cdot s_m \cdot X$. The other case is an easy consequence of Lemma 7.3 and the fact that $\mu \in \mathrm{Ann}(s|m+1)$. □

For the case $\kappa_{m+1,c} > \kappa_{0,c}$, we have:

Corollary 7.5 *(cf. [14].) Suppose that $m+1 < 0$, $r \in [R \setminus \{0\}]$, $\mu_{n,a} \in \mathrm{Min}(s|n,a)$ for all $a \in [R \setminus \{0\}]$ and $m+1 \leq n \leq 0$. Let $\mu = \mu_{m+1,r}$, $\mathcal{O}_{m+1,r} \neq 0$ and $c = (\mathcal{O}_{m+1,r})^*$.*

If $\kappa_{m+1,c} > \kappa_{0,c}$, $\alpha = \alpha_{m+1,c}$, d are defined as in Theorem 6.4 and $\mu_{m,r} = [\mu_{m+1,r}, \mu_{\alpha,d}]$ then

$$\beta_{m,r} = X^{\delta - \delta\mu} \beta_{m+1,r} - (\mathcal{O}_{m+1,r}/\mathcal{O}_{\alpha,d}) \cdot X^{\delta+m-\alpha+1-\delta\mu_{\alpha,d}} \beta_{\alpha,d}$$

where $\delta = \delta\mu_{m,r}$.

PROOF. This uses Lemma 7.3 and is omitted since it is similar to the proof of [11, Proposition 4.4]. □

7.2 The main result

The results of the previous two sections can be combined into a compact formula for $(\mu_{m,r}, \beta_{m,r})$. We will use the following notation:

Definition 7.6 *If m, r, $\mu_{m+1,r}$, c, $\mu_{\alpha,d}$ and $\mu_{m,r}$ are as in Theorem 6.4, we set*

$$\overline{\mu}_{m,r} = (\mu_{m,r},\ \beta_{m,r})$$
$$q = q_{m+1,r} = \mathcal{O}_{m+1,r}/\mathcal{O}_{\alpha,d}$$
$$d = d_{m+1,r} = \delta\mu_{m+1,r} + \delta\mu_{m+1,c} + m - 1.$$

We now give a succint formula for $\overline{\mu}_{m,r}$ using Theorem 6.4 and Corollary 7.5:

Corollary 7.7 *Suppose that $m+1 < 0$, $r \in [R \setminus \{0\}]$, $\overline{\mu}_{n,a} \in MR(s|n,a)$ for all $a \in [R \setminus \{0\}]$ and $m+1 \leq n \leq 0$. Let $\mathcal{O}_{m+1,r} \neq 0$ and $c = (\mathcal{O}_{m+1,r})^*$. If $\kappa_{m+1,c} > \kappa_{0,c}$, then*

$$\overline{\mu}_{m,r} = \overline{\mu}_{m+1,r} - q \cdot X^{|d|} \overline{\mu}_{\alpha,d}$$

where $\overline{\mu}_{m+1,r}$ and $-q \cdot \overline{\mu}_{\alpha,d}$ have been interchanged if $d < 0$.

PROOF. This is a straightforward consequence of the definitions, Theorem 6.4 and Corollary 7.5 since $\delta\mu_{m+1,c} - 1 = -\alpha - \delta\mu_{\alpha,d} + 1$. □

To obtain the analogue of Corollary 7.7 for the case $\kappa_{m+1,c} = \kappa_{0,c}$, we extend Definition 7.6 as follows:

Definition 7.8 *If $m < 0,\ r,\ \mu_{m+1,r},\ c,\ \kappa_{m+1,c} = \kappa_{0,c}$ and $\mu_{m,r}$ are as in Theorem 6.2, we set*

$$(\overline{\mu}_{\alpha,d},\ \mathcal{O}_{\alpha,d}) = \begin{cases} ((0,-X),\ 1) & \text{if } \kappa_{0,c} = 0 \\ ((1,0),\ s_0) & \text{if } \kappa_{0,c} = 1 \end{cases}$$

and let $q = q_{m+1,r},\ d = d_{m+1,r}$ be as in Definition 7.6.

We now give the compact formula for $\overline{\mu}_{m,r}$, which is the main result of this paper (combining Corollary 7.7, Theorem 6.2 and Corollary 7.4):

Theorem 7.9 *Suppose that $m < 0,\ r \in [R \setminus \{0\}],\ \overline{\mu}_{n,a} \in MR(s|n,a)$ for all $a \in [R \setminus \{0\}]$ and $m+1 \leq n \leq 0$. Let $\mathcal{O}_{m+1,r} \neq 0$ and $c = (\mathcal{O}_{m+1,r})^*$. If $\overline{\mu}_{\alpha,d},\ q$ and d are as in Definitions 7.6 and 7.8, then*

$$\overline{\mu}_{m,r} = \overline{\mu}_{m+1,r} - q \cdot X^{|d|}\overline{\mu}_{\alpha,d}$$

where $\overline{\mu}_{m+1,r}$ and $-q \cdot \overline{\mu}_{\alpha,d}$ have been interchanged if $d < 0$.

Proof. We only need to verify the case $\kappa_{m+1,c} = \kappa_{0,c}$, which is straightforward since

$$d = \begin{cases} m-1 & \text{if } \kappa_{0,c} = 0 \\ \delta\mu_{m+1,r} + m & \text{if } \kappa_{0,c} = 1. \end{cases}$$

□

7.3 Algorithms MR and MP

The main result yields a compact algorithm to compute $\overline{\mu}_{m,r}$. When $d_{i+1,r} < 0$, the integer $\alpha_{i+1,r}$ is as in the preamble to Algorithm 6.6. However, in this compact version, we update $\mu_{\alpha_{i+1,r},r}$ directly, as in [11, 12].

(i) initialize $\overline{\mu}_{\alpha,d}$ as in Definition 7.8 and initialize $\overline{\mu}_{r,0}$;

(ii) update $\overline{\mu}_{m,r}$ via Theorem 7.9;

(iii) if $d_{i+1,r} < 0$, set $\overline{\mu}_{\alpha_{i,r},r}$ equal to $\overline{\mu}_{i+1,r}$ and restore $\overline{\mu}_{\alpha_{i+1,c},d}$;

(iv) incorporate the case $\kappa_{m+1,c} = \kappa_{0,c} = 1$ into the body of the main loop, updating $\overline{\mu}_{\alpha,r}$ to $\overline{\mu}_{r,0}$ as appropriate;

(v) suppress the subscripts i and $i+1$ (only the current $\overline{\mu}_{i+1,r}$ and $\overline{\mu}_{\alpha_{i+1,c},d}$ are needed to obtain $\overline{\mu}_{i,r}$ and $\overline{\mu}_{\alpha_{i,r},r}$).

Algorithm MR. *(cf. [14]) Input: $m \leq 0$, a finite chain ring R, $s_0, \ldots, s_m \in R$.*

Output: $\overline{\mu}_r \in MR(s|m,r), r \in [R \setminus \{0\}]$.

for $r \in [R \setminus \{0\}]$ *do* $\{\alpha_r := 1;\ \overline{\mu}_{1,r} := (0, -X);\ \mathcal{O}_{1,r} := 1;\ \overline{\mu}_r\ := (r, 0);\}$

for $i := 0$ *to* m *do*
  for $r \in [R \setminus \{0\}]$ *do*
    $\{\mathcal{O} := (\mu_r \circ s)_{i+\delta\mu_r};$
     if $\mathcal{O} \neq 0$ *then* $\{\ c := \mathcal{O}^*;\ d := (\mathcal{O}_{\alpha_c,c})^*; q := \mathcal{O}/d;$
                $d := \delta\mu_r + \delta\mu_c + i - 1;$
                if $d < 0$ *then* $\{\ \overline{\mu}_{\alpha_r,r} := \overline{\mu}_r;\ d := -d;$
                  $\overline{t} := \overline{\mu}_{\alpha_c,d};\ swap(\overline{\mu}_r, q \cdot \overline{\mu}_{\alpha_c,d});\}$
                $\overline{\mu}_r := \overline{\mu}_r - q \cdot X^d \overline{\mu}_{\alpha_c,d};$
                if $d < 0$ *then* $\overline{\mu}_{\alpha_c,d} := \overline{t};\}\}$

for $r \in [R \setminus \{0\}]$ *do return* $\overline{\mu}_r$.

Example 7.10 *Algorithm MR computes the following table of border polynomials for the* $\mathbb{Z}_9$*–sequence* $6, 3, 1, 5, 6$ *of [14, p 512]:*

i	$\beta_{i,1}$	$\beta_{i,3}$
0	$6X$	0
-1	$6X$	0
-2	$6X^3 + 3X^2 + X$	0
-3	$6X^3 + X$	0
-4	$6X^3 + X$	$3X$

Remarks 7.11 *(i) If* $\mathcal{O} = \mathcal{O}_{i+1,r} = 0$ *then* $\overline{\mu}_r$ *is unchanged. (ii) We have suppressed the negation in the swap (-1 is a unit). (iv) There does not seem to be a simple recursive formula for updating* $d_{m+1,r}$ *(as in Theorem 8.4, where* R *is a field).*

If we omit the second component of the $\overline{\mu}_r$ in Algorithm MR, we obtain an improved version of Algorithm 6.6, which we call **Algorithm MP**. Our next example is an application of Algorithm MP to a finite sequence over a Galois ring.

Example 7.12 *(cf. [6, Example 2, p. 1019]) Since* $y^2 + y + 2$ *is irreducible over* $GF(3)$*, it is irreducible over* $\mathbb{Z}_9$ *and we may use* $\mathbb{Z}_9[y]/(y^2 + y + 2)$ *as* $R = GR(9, 2)$.

For the R*–sequence* $3, 3y, 3, 3$*, Algorithm* MP *iterates as follows:*

i	$\mathcal{O}_{i,1}$	$\mu_{i,1}$	$\alpha_{i,1}$	$\mathcal{O}_{i,3}$	$\mu_{i,3}$	$\alpha_{i,3}$
0	3	X	0	0	3	0
-1	$3y$	$X-y$	0	0	3	0
-2	$3y$	X^2-yX-y	-1	0	3	0
-3	0	X^2-yX-y	-1	0	3	0

Thus X^2-yX-y is a monic minimal polynomial for $3, 3y, 3, 3$.

We conclude this section by tabulating the complexity of Algorithms MP and MR:

Proposition 7.13 *Let R be a finite chain ring, let ν be the nilpotency index of R and $l = 1-m$. The algebraic and storage complexity of the Algorithms MP and MR is as follows:*

Algorithm	*R–multiplications*	*ℓ–inverses*	*divisions by associate*
MP	$\leq \nu l^2$	$2\nu l$	νl
MR	$\leq 3\nu l(l-1)/2$	$2\nu l$	νl

Variable	*Type*	*Norm*	*Number*
$\mu_{\alpha_c,d}$	$R[X]$	$\leq l-2$	ν
$\mathcal{O}_{\alpha_c,d}$	$[R\setminus\{0\}]$	–	ν
μ_r	$R[X]$	$\leq l$	ν
$\mathcal{O}$	$[R\setminus\{0\}]$	–	ν
d_r	$\mathbb{Z}$	$\leq l$	ν
$\beta_{\alpha_c,d}$	$XR[X]$	$\leq l-2$	ν
β_r	$XR[X]$	$\leq l$	ν

PROOF. These are easy counting arguments as in [11, Proposition 3.23]. □

If ν is 'large', we can reduce the storage complexity of Algorithm MR by first using Algorithm MP and then computing $\beta(\mu_r, s|m)$ directly from the definition.

8 The key equation

It is well–known that the key equation over a field may be solved using the Euclidean algorithm. For a recent approach using left and right–shift versions of this algorithm, see [5, 10]. When there are zero–divisors however, division–based methods such as the Euclidean method fail. In this section we show how to solve the key equation over a finite chain ring.

We first simplify and generalize [13] to a commutative ring, using the key equation derived in [3].

Definition 8.1 *Let $g,\ S \in R[X],\ 0 \le \delta S \le \delta g - 1$ and g monic. Then $(\sigma, \omega) \in R[X] \times R[X]$, $\sigma \neq 0$, solves the key equation if*

$$\sigma S \equiv \omega \pmod{g},$$

σ is monic and $\delta\omega \le \delta\sigma - 1 \le \delta g - 1$. If $\delta\sigma$ is minimal amongst all non–zero solutions of the key equation, then (σ, ω) is called a minimal solution.

Throughout this section, g and S are as in the definition of the key equation.

Proposition 8.2 *Let $m = 1 - \delta g \le 0$ and define s by $\Gamma(s) = XS/g$. Then*

(i) (σ, ω) solves the key equation iff $(\sigma, \beta(\sigma, s|m))$ realizes $s|m$, σ is monic, $\delta\sigma \le 1 - m$ and $\omega = \sigma S - g\beta(\sigma, s|m)/X$;

(ii) (σ, ω) is a minimal solution iff $(\sigma, \beta(\sigma, s|m)) \in MR(s|m, 1)$ and $\omega = \sigma S - g\beta(\sigma, s|m)/X$;

(iii) $s_0 = S_{-m}$ and $s_i = ((X^{-i}S) \bmod g)_{-m}$ for $m \le i \le 0$.

Proof. Since g is monic, $1/g,\ XS/g \in R[[X^{-1}]]$ are well–defined. Also, $1/g$ is monic and $\delta(1/g) = -\delta g$.

(i) Let (σ, ω) be a solution and $\sigma S - \omega = gh$ for some $h \in R[X]$. Put $d = \delta\sigma$. Since g and σ are monic, $\delta h + 1 - m = \delta(\sigma S - \omega) = \max\{d + \delta S, \delta\omega\} = d + \delta S \le d - m$. Hence $\delta(Xh) \le d \le 1 - m$. Also, $\sigma\Gamma(s|m) - Xh = X\omega/g + \sigma G$ where $\delta G < m$. Hence (σ, Xh) realizes $s|m$ and $Xh = \beta(\sigma, s|m)$ by Proposition 7.2. Thus ω is as stated.

Conversely, let $\beta = \beta(\sigma, s|m),\ (\sigma, \beta)$ realize $s|m$ and $\omega = \sigma S - g\beta/X$. Then $\sigma S = \omega + g\beta/X$ and $\delta\omega = \delta(g(\sigma XS/g - \beta)X^{-1}) < \delta g + (m + \delta\sigma) - 1 = \delta\sigma$, as required.

(ii) Simple consequence of (i).

(iii) For $i \le 0,\ X^{-i}S = X^{-i-1}g\Gamma(s) = g\sum_{j \le 0} s_j X^{-i-1+j} = g\sum_{j < i+1} s_j X^{-i-1+j} + gq$ where $q = \sum_{j=i+1}^{0} s_j X^{-i-1+j} \in R[X]$. Hence $((X^{-i}S) \bmod g)_{-m} = (g\sum_{k<0} s_{k+i+1}X^k)_{-m} = s_i$ since $-m = \delta g - 1$ and g is monic. □

This means we can now solve the key equation over a finite chain ring as follows:

Algorithm 8.3 *(Solve classical key equation)*

Input: R a finite chain ring, $g, S \in R[X],\ \delta g \ge 1,\ 0 \le \delta S \le \delta g - 1,\ g$ monic.

Output: $\overline{\sigma} = (\sigma, \omega)$ with $S\sigma \equiv \omega \pmod{g},\ \delta\omega \le \delta\sigma - 1 \le \delta g - 1$ and $\delta\sigma$ minimal.

0. $m := 1 - \delta g$;

1. For $i := 0$ to m do $s_i = ((X^{-i}S) \bmod g)_{-m}$;

2. Compute $(\mu, \beta(\mu, s|m))$, where $\mu = \mu_{m,1}$.

3. Return $\overline{\sigma} = (\mu, \mu S - g\beta(\mu, s|m)/X)$.

As in [13], we can reduce the complexity of Algorithm 8.3 by setting $\omega = (S\sigma) \bmod g$ and not computing $\beta(\mu_r, s|m)$.

We conclude by showing that over a field, a monic minimal solution may be also computed by re–initializing Algorithm MR of [12]. Over a field, there are no non–trivial zero–divisors, and so we suppress the subscripts r, c and d. (For example, when $m < 0$ and $\kappa_{m+1} = \kappa_0$, $\overline{\mu}_\alpha$ is defined as in Definition 7.8.) We also abbreviate $\mathcal{O}\mu_i$ to $\mathcal{O}_i$ for $i \leq 0$. Theorem 7.9 can then be simplified and improved as in the following result from [12]:

Theorem 8.4 *(cf. [2, Section 7.3], [8].) Let* $m < 0$, $\overline{\mu}_i \in MR(s|i)$ *for* $m + 1 \leq i \leq 0$, $\mathcal{O}_{m+1} \neq 0$ *and* $d = d_{m+1} = 2\delta\mu_{m+1} + m - 1$. *Put* $\alpha = \alpha_{m+1}$ *and* $\mathcal{O}_\alpha = \mathcal{O}_\alpha$. *Then*

$$(i)\ \overline{\mu}_m = \mathcal{O}_\alpha \cdot \overline{\mu}_{m+1} - \mathcal{O}_{m+1} \cdot X^{|d|}\overline{\mu}_\alpha$$

where $\overline{\mu}_{m+1}$, $-\overline{\mu}_\alpha$ *and* $\mathcal{O}_{m+1}$, $\mathcal{O}_\alpha$ *have been interchanged if* $d < 0$, *and with this interchange*

$$(ii)\ \overline{\mu}_{\alpha_m} = \overline{\mu}_{\alpha_{m+1}},\ d_m = |d_{m+1}| - 1.$$

Proposition 8.5 *Let* R *be a field and define* $\overline{\sigma}_{\alpha_0} = (0, g)$, $\overline{\sigma}_0 = (1, S)$, $\mathcal{O}_{\alpha_0} = 1$ *and* $d_0 = -1$. *If for* $m \leq i \leq 0$, $\overline{\sigma}_{\alpha_i}, \overline{\sigma}_i$ *and* d_i *are defined iteratively as in Theorem 8.4, then* $\overline{\sigma}_m/\lambda\sigma$ *is a minimal solution of the key equation.*

Proof. By Proposition 8.2, it suffices to show that $\overline{\nu}_m = (\sigma_m, X(\sigma_m S - \omega_m)/g) \in MR(s|m)$. Let $m \leq i \leq 0$ and let $\overline{\mu}_i \in MR(s|i)$ be obtained from Theorem 8.4. Then $\overline{\nu}_{\alpha_0} = (0, -X) = \overline{\mu}_{\alpha_0}$, $\overline{\nu}_0 = (1, 0) = \overline{\mu}_0$. It is easy to verify that if $m \leq i \leq 0$ and $\overline{\nu}_k = \overline{\mu}_k$ for all $k \geq i + 1$, then $\overline{\nu}_i = \overline{\mu}_i$. Hence it follows inductively that $\overline{\nu}_m = \overline{\mu}_m$ and so $\overline{\sigma}_m$ is a minimal solution of the key equation. □

We have now justified

Algorithm 8.6 *(cf. [13].) (Solve classical key equation.)*

Input: Field F, $g, S \in F[X]$, $\delta g \geq 1$, $0 \leq \delta S \leq \delta g - 1$, g *monic.*

Output: $\overline{\sigma} = (\sigma, \omega)$ *with* $S\sigma \equiv \omega \pmod{g}$, $\delta\omega \leq \delta\sigma - 1 \leq \delta g - 1$ *and* $\delta\sigma$ *minimal.*

$m := 1 - \delta g$;

for $i := 0$ *to* m *do* $s_i = ((X^{-i}S) \bmod g)_{-m}$;

$\overline{\sigma}' := (0, g)$; $\mathcal{O}' := 1$; $\overline{\sigma} := (1, S)$; $d := -1$;

for $i := 0$ *to* m *do* { $\mathcal{O} := (\sigma \circ s)_{i+\delta\sigma}$;

$$\begin{aligned}
&\textit{if } (\mathcal{O} \neq 0) \ \{ \ \textit{if } (d < 0) \ \{d := -d; swap(\overline{\sigma}, \overline{\sigma}'); swap(\mathcal{O}, \mathcal{O}'); \} \\
&\qquad\qquad \overline{\sigma} := \mathcal{O}' \cdot \overline{\sigma} - \mathcal{O} \cdot X^d \overline{\sigma}'; \} \\
&d := d - 1; \}
\end{aligned}$$

return $\overline{\sigma}/\lambda(\sigma)$.

Remarks 8.7 *(i) We have written* $\overline{\mu}'$ *and* $\mathcal{O}'$ *for* $\overline{\mu}_\alpha$ *and* $\mathcal{O}_\alpha$ *since the actual values of* α *and their provenance are not needed. (ii) We have suppressed the negation in the swap, which is possible since we are over a field. (iii) When* R *is a finite chain ring, it is more efficient to proceed as in Algorithm 8.3, and so we will not compute* ω_m *by reinitializing in this case.*

References

[1] Atiyah, M.F., Macdonald, I.G. (1969). Introduction to Commutative Algebra. Addison–Wesley.

[2] Berlekamp, E.R. (1968). *Algebraic Coding Theory.* Mc–Graw Hill, New York.

[3] Berlekamp, E.R. (1973). Goppa Codes. *IEEE Trans. Information Theory* **19**, 590–592.

[4] Calderbank, A.R., Hammons, A.R., Kumar, P.V., Sloane, N.J.A., Solé, P. (1993). A linear construction for certain Kerdock and Preparata codes. *Bull. Amer. Math. Soc.* **29**, 218–222.

[5] Fitzpatrick, P., Nelson, J., Norton, G.H. (1989). A systolic version of the extended Euclidean algorithm. *Proc. International Conference on systolic arrays, Killarney.* 477–487. Prentice–Hall.

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

[7] Gilmer, R. (1972). *Multiplicative Ideal Theory.* Marcel Dekker.

[8] Massey, J.L. (1969). Shift–register synthesis and BCH decoding. *IEEE Trans. Information Theory* **15**, 122–127.

[9] MacDonald, B.R. (1974). Finite rings with identity. Marcel Dekker, Inc.

[10] Norton, G.H. (1989). Precise analyses of the right– and left–shift greatest common divisor algorithms for $GF(q)[X]$. *S.I.A.M. J. Computing* **18**, 608–624.

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

[12] Norton, G.H. (1999). On shortest linear recurrences. *J. Symbolic Computation* **27**, 325–349.

[13] Patterson, N.J.(1975). The algebraic decoding of Goppa codes. *IEEE Trans. IT* **21**, 203–207.

[14] Reeds, J.A., Sloane, N.J.A. (1985). Shift–register synthesis (modulo m). S.I.A.M. J. Computing **14**, 505–513.