On Shortest Linear Recurrences.*

GRAHAM H. NORTON †
Algebraic Coding Research Group
Centre for Communications Research, University of Bristol, England

July 19, 2001

Abstract

This is an expository account of a constructive theorem on shortest linear recurrences over an arbitrary integral domain R. A generalisation of rational approximation, which we call 'realization', plays a key role throughout the paper. We also give the associated 'minimal realization' algorithm, which has a simple control structure and is division-free. It is easy to show that the number of R-multiplications required is $O(n^2)$, where n is the length of the input sequence.

Our approach is algebraic and independent of any particular application. We view a linear recurring sequence as a torsion element in a natural $R[X]$-module. The standard $R[X]$-module of Laurent polynomials over R underlies our approach to finite sequences. The prerequisites are nominal and we use short Fibonacci sequences as running examples.

Contents

*Dedicated to Christine Byrne 1944-1997, *in memoriam*.

†Current addresses: Dept. Mathematics, University of Queensland, Brisbane 4072; ghn@maths.uq.edu.au. Copyright Academic Press, 1999.

1 Introduction.

The problem of finding a shortest linear recurrence satisfied by a given finite sequence is important in view of its manifold applications. Before citing some of these, we formulate the problem in the notation of Massey (1969).

Let K be a field and let n be a strictly positive integer. Suppose we are given a sequence $S_0, \ldots, S_{n-1}$ over K (i.e. $S_i \in K$ for $0 \leq i \leq n-1$) and $l \geq 0$, we say that $c = (c_0, \ldots, c_l) \in K^{l+1}$ *defines a linear recurrence of length l for $S_0, \ldots, S_{n-1}$* if $c_0 = 1$ and

$$c_0 S_i + \cdots + c_l S_{i-l} = 0 \text{ for } l \leq i \leq n-1. \tag{1}$$

For example, 1 defines a linear recurrence of length 0 for $S_0, \ldots, S_{n-1}$ if, and only if $S_i = 0$ for $0 \leq i \leq n-1$. If $c_0, \ldots, c_l$ defines a linear recurrence of length l for $S_0, \ldots, S_{n-1}$, we say that it

is a *shortest linear recurrence* for $S_0, \ldots, S_{n-1}$ if whenever $b_0, \ldots, b_k$ defines a linear recurrence of length k for $S_0, \ldots, S_{n-1}$, we have $l \leq k$.

Any $1, c_1, \ldots, c_n$ defines a linear recurrence of length n for $S_0, \ldots, S_{n-1}$ since Equation (1) is then vacuously satisfied. Thus the following problem always has a solution:

PROBLEM 1.1 *Find a* shortest *linear recurrence for* $S_0, \ldots, S_{n-1}$ *over* K.

We can rewrite Equation (1) as $S_i = -\sum_{j=1}^{l} c_j S_{i-j}$ for $l \leq i \leq n-1$, where by the usual convention, a sum over the empty set is zero. Then $S_0, \ldots, S_{n-1}$ is said to be *'generated by a linear feedback shift-register (LFSR) of length* $l \geq 0$ with feedback coefficients $c_1, \ldots, c_l$', Massey, *loc. cit.*

The LFSR Synthesis Algorithm, *loc. cit.* solves Problem 1.1. In fact, Massey simplified Berlekamp's method for decoding BCH codes (see Section 7.3 of Berlekamp (1968), 'Heuristic solution of the key equation'). See also Trench (1964). The LFSR Synthesis Algorithm is commonly known as the Berlekamp-Massey Algorithm; see Blahut (1983) for an exposition of it in terms of LFSR's.

The uniquely defined length of a shortest linear recurrence for $S_0, \ldots, S_{n-1}$ is called the *linear complexity* of $S_0, \ldots, S_{n-1}$. We remark that Massey's LFSR Synthesis Algorithm applies to an arbitrary finite sequence, whereas in Berlekamp's original decoding algorithm, the sequence of length $2t$ is known to have linear complexity at most t, where $t \geq 1$. For some applications of linear complexity to Algebraic Coding Theory, see Matt & Massey (1980), Massey & Schaub (1988) and Dür (1992), Chan & Norton (1995).

The linear complexity of a finite (binary) sequence is widely used in Cryptography, van Tilborg (1988). It is related to continued fractions, Welch & Scholz (1979), Mills (1975) and Padé approximation Brent *et. al.* (1980). It can be used for data compression, Massey (1969), for computing growth functions, Brazil (1993), for solving sparse linear equations, Wiedemann (1986), for sparse interpolation, Dür & Grabmeier (1993), for computing canonical binary forms, Dür (1989) and as a method (attributed to D.H. Lehmer) for computing composed products of polynomials, Brawley *et al.* (1999). LFSR's are known as filters in Digital Signal Processing, and the Berlekamp-Massey Algorithm is used to design them, Blahut (1985). The Berlekamp-Massey Algorithm also solves the partial realization problem of Control (Mathematical Systems) Theory, Kalman *et al.*, (1969).

If $S_0, \ldots, S_{n-1}$ is a sequence over an *integral domain* R, we say that $c_0, \ldots, c_l \in R$ defines a *linear recurrence of length* $l \geq 0$ *for* $S_0, \ldots, S_{n-1}$ if $c_0 \neq 0$ and Equation (1) is satisfied. As above, the following problem always has a solution:

PROBLEM 1.2 *Find a* shortest *linear recurrence for* $S_0, \ldots, S_{n-1}$ *over* R.

A division-free solution to Problem 1.2 was given in Norton (1995a), which also contains applications to Linear Algebra, to computing symbolic enumerators (*cf.* Mason's rule for signal-flow

graphs, Mason & Zimmermann (1960)) and to solving the *parametrized* partial realization problem of Control Theory. An application to Yule-Walker and Wiener-Hopf equations appears in Gueret (1996). This theory also applies to a sequence over R with missing or unknown terms $\{S_i : i \in I\}$; we may regard them as indeterminates and the original sequence as being over the integral domain $R[S_i : i \in I]$. The division-free algorithm can also be used to conduct experiments on the shortest linear recurrences of a finite sequence over R; see e.g. Conjecture 3.22 of Norton (1995a).

Here, we have also simplified the theory developed in Norton (1995a). Our concepts and proofs are not difficult, but we have motivated them and given complete details for expository reasons. The set of sequences over R is an $R[X]$-module in a natural way and we view the subset of linear recurring sequences as its torsion submodule. For finite sequences, the standard $R[X]$-module of Laurent polynomials over R underlies our approach. In other words, for finite sequences, we exploit $R[X^{-1}, X]$ with the subring $R[X]$ acting by multiplication in $R[X^{-1}, X]$. Another $R[X]$-module, $R[X]^2$, also enters naturally, see Section 7. Thus our approach to Problem 1.1 is algebraic rather than application-oriented. This paper can be read independently of Berlekamp (1968) and Massey (1969).

Our iterative algorithm has a simpler control structure than the algorithms on page 184, Berlekamp (1968) and Massey (1969), and it is easy to implement and to analyse. Thus we also address the conceptual *v.* practical conflict (*sic*) of p. *vii*, Berlekamp (1968).

Our algebraic approach extends to the study of multiple sequences, Section 8 of Norton (1995b). It generalizes to shortest linear recurrences of a finite sequence over a finite chain ring A, e.g. a Galois ring. Our algorithm can be modified to find a minimal-degree solution of a congruence in $A[X]$ usually solved by the extended Euclidean algorithm when A is a field. See Norton (1999) for details.

This paper is an expanded version of a short note submitted to this Journal. The author would like to thank Tim Blackmore, the referees and the Editor-in-Chief for their useful comments and suggestions.

2 A reformulation of the shortest linear recurrence problem.

2.1 Minimal polynomials of a finite sequence.

Suppose $c_0, \ldots, c_l$ defines a linear recurrence of length l for $S_0, \ldots, S_{n-1}$ over a field K as in Section 1. In Section III of Massey (1969), the 'connection polynomial' of this recurrence, viz. $c(X) = 1 + c_1 X + \cdots + c_l X^l$ is defined, where the connection polynomial of 'the LFSR of length 0' is taken to be $c(X) = 1$. Thus $c_0 = 1$ and $\deg c \leq l$.

We will see that the polynomial $f(X) = \sum_{i=0}^{l} f_i X^i$, where $f_i = c_{l-i}$ for $0 \leq i \leq l$ is more convenient. (Recall that the *reciprocal* of c is $X^{\deg\, c} c(X^{-1})$. We easily have $f(X) = X^{l-\deg\, c} c^*(X)$ and $c = f^*$. However, we prefer not to use reciprocals as we are dealing with a linear problem and it is easy to find polynomials g, h with $(g+h)^* \neq g^* + h^*$.)

Now f is monic and deg $f = l$. (By a *monic* polynomial, we always mean a *non-zero* polynomial with leading coefficient 1.) If we relabel the given sequence as

$$T_0 = S_0, T_{-1} = S_1, \ldots, T_{1-n} = S_{n-1},$$

put $m = 1 - n \leq 0$ and reverse the order of summation in Equation (1), we see that Equation (1) is equivalent to

$$f_0 T_{l-i} + f_1 T_{l-i-1} + \cdots + f_l T_{-i} = 0 \text{ for } l + m \leq l - i \leq 0. \tag{2}$$

Set $T = T(X^{-1}) = T_0 + \cdots + T_m X^m$. Then the left-hand side of Equation (2) is $(f \cdot T)_{l-i}$ the $(l-i)^{th}$ coefficient of the product $f \cdot T$. Thus we can rewrite Equation (1) as

$$(f \cdot T)_j = (f(X) \cdot T(X^{-1}))_j = 0 \text{ for } m + \deg\, f \leq j \leq 0. \tag{3}$$

Again, the polynomial 1 satisfies Equation (3) if and only if $T_i = 0$ for $m \leq i \leq 0$.

Notice that the sequence $T_0, \ldots, T_m$ has $1 - m \geq 1$ terms and that any polynomial f of degree $1 - m$, e.g. $f(X) = X^{1-m}$ satisfies Equation (3) since $m + \deg\, f \leq i \leq 0$ is vacuously satisfied; *cf.* '$1, c_1, \ldots, c_n$ defines a linear recurrence of length n for $S_0, \ldots, S_{n-1}$'. Thus there is always a monic polynomial which satisfies Equation (3).

Let $-\infty$ denote a symbol satisfying $-\infty < d$, $-\infty + d = -\infty$ for *any integer* d and put deg $0 = -\infty$. Then 0 satisfies Equation (3), but we are interested in *non-zero* solutions in $R[X]^* = R[X] \setminus \{0\}$, where R is an integral domain.

DEFINITION 2.1 *We will call a solution $f \in R[X]^*$ of Equation (3) a* minimal polynomial for $T_0, \ldots, T_m$, *written $f \in \mathrm{Min}(T_0, \ldots, T_m)$, if for any $g \in R[X]^*$ which satisfies Equation (3),* deg $f \leq$ deg g.

Notice that in the previous definition, we do not insist that a solution of Equation (3) be monic. If R is a field and $f \in \mathrm{Min}(T_0, \ldots, T_m)$, then $1, f_{l-1} f_l^{-1}, \ldots, f_0 f_l^{-1}$ defines a shortest linear recurrence for $T_0, \ldots, T_{-m}$ and deg f is the linear complexity of $T_0, \ldots, T_{-m}$. We are now ready to reformulate Problem 1.2:

PROBLEM 2.2 *Given $m \leq 0$ and a finite sequence $T_0, \ldots, T_m$ over R, find an $f \in \mathrm{Min}(T_0, \ldots, T_m)$.*

EXAMPLE 2.3 *We show that $\phi(X) = X^2 - X - 1$ is a minimal polynomial for the first three terms 1,1,2 of the Fibonacci sequence. In the notation of the previous section, we have $S_2 = S_1 + S_0$ i.e.*

$l = 2 = n - 1$, $c_0 = 1$ and $c_1 = c_2 = -1$ for Equation (1). So the trial solution is $X^2 - X - 1 = \phi$. Now $m = -2$, $m + \deg \phi = 0$ and the 0^{th} coefficient of

$$\phi \cdot (1 + X^{-1} + 2X^{-2}) = -2X^{-2} + X^{-1} + X^2$$

vanishes. Thus ϕ satisfies Equation (3). It is clear that 1,1,2 does not have linear complexity 0. To see that it cannot have linear complexity 1, consider

$$(aX + b) \cdot (1 + X^{-1} + 2X^{-2}) = 2bX^{-2} + (2a + b)X^{-1} + (a + b) + aX,$$

where $a \neq 0$. If $aX + b$ satisfies Equation (3), we must have $2a + b = a + b = 0$, which contradicts $a \neq 0$. Thus $\phi \in \mathrm{Min}(1, 1, 2)$.

The reader may wish to show that for any integer k, $\phi(X) + k(X - 1) \in \mathrm{Min}(1, 1, 2)$ and that $\phi \in \mathrm{Min}(0, 1, 1, 2)$.

2.2 Minimal realizations of a finite sequence.

Let $m \leq 0$ and $T_0, \ldots, T_m$ be a finite sequence over R as above and put $T = T(X^{-1}) = T_0 + \cdots + T_m X^m$. For $f \in R[X]^*$, it is worthwhile looking at the product $f \cdot T$ occurring in Equation (3) in more detail. The product is

$$f \cdot T = \left(\sum_{i=m}^{m-1+\deg f} + \sum_{i=m+\deg f}^{0} + \sum_{i=1}^{\deg f} \right) (f \cdot T)_i X^i \in R((X^{-1})) \tag{4}$$

i.e. $f \cdot T$ is a Laurent series in X^{-1} over R. *As in Norton (1995a), the third summand will appear often and we will use the notation $\beta(f, T)$ for this polynomial product:*

$$\beta(f, T) = \sum_{i=1}^{\deg f} (f \cdot T)_i X^i \in XR[X].$$

Thus if f satisfies Equation (3) then $f \cdot T - \beta(f, T) = \sum_{i=m}^{m-1+\deg f} (f \cdot T)_i X^i$.

At this stage, it is convenient to introduce a common generalization of the order function $R[[X^{-1}]] \to \{-\infty\} \cup \{\ldots, -1, 0\}$ and $\deg : R[X] \to \{-\infty\} \cup \{0, 1, \ldots\}$.

DEFINITION 2.4 *We define $\delta : R((X^{-1})) \to \{-\infty\} \cup \{\ldots, -1, 0, 1, \ldots\}$ by $\delta(0) = -\infty$ and*

$$\delta(F) = \max\{i : \ F_i \neq 0\} \text{ if } F \in R((X^{-1})) \setminus \{0\}.$$

Thus $\delta(X^{-1} + 1) = 0$ and for any integer d, $\delta(F) \leq d \Leftrightarrow F_i = 0$ for $i \geq d + 1$. As for the order function and deg , $\delta(F \cdot G) = \delta(F) + \delta(G)$ and $\delta(F + G) \leq \max\{\delta(F), \delta(G)\}$ for all $F, G \in R((X^{-1}))$.

PROPOSITION 2.5 *Suppose that $m \leq 0$, $T_0, \ldots, T_m$ is a finite sequence over R, $T = T_0 + \cdots + T_m X^m$ and $f \in R[X]^*$. If f satisfies Equation (3), then $\delta(f \cdot T - \beta(f, T)) \leq m - 1 + \deg f$. Conversely, if $\delta(f \cdot T - g) \leq m - 1 + \deg f$ for some $g \in XR[X]$, then f satisfies Equation (3) and $g = \beta(f, T)$.*

PROOF. For the converse, let $F = (f \cdot T - g) - \sum_{i=m}^{m-1+\deg\ f} (f \cdot T)_i X^i$. The hypothesis implies that $\delta(F) \leq m - 1 + \deg f$ i.e. $F_i = 0$ for $i \geq m + \deg f$. On the other hand, $F = \sum_{m+\deg\ f}^{0} (f \cdot T)_i X^i + \beta(f,T) - g$. Thus $F = 0$ and $\sum_{i=m+\deg\ f}^{0} (f \cdot T)_i X^i = g - \beta(f,T) \in R[[X^{-1}]] \cap XR[X] = \{0\}$. Hence f satisfies Equation (3) and $g = \beta(f,T)$. □

The previous proposition allows us to simplify Definition 2.5 of Norton (1995a) slightly:

DEFINITION 2.6 *Let $m \leq 0$, $T_0, \ldots, T_m$ be a finite sequence over R and $T = T_0 + \cdots + T_m X^m$. We call $(f, \beta(f,T)) \in R[X]^* \times XR[X]$ a* realization of $T_0, \ldots, T_m$ *or say that $(f, \beta(f,T))$ realizes $T_0, \ldots, T_m$ if $\delta(f \cdot T - \beta(f,T)) \leq m - 1 + \deg f$. If further* deg f *is minimal, we call $(f, \beta(f,T))$ a* minimal realization (MR) *of $T_0, \ldots, T_m$.*

For example, if $T_0, \ldots, T_m$ is the sequence $0, \ldots, 0, 1$ (where there are $-m$ zeroes) then (X^{1-m}, X) realizes $T_0, \ldots, T_m$ and $(X^2 - X - 1, X^2)$ is an MR of $1, 1, 2$. We will see in Proposition 5.3 that (X^{1-m}, X) is actually an MR of $0, \ldots, 0, 1$.

It follows from Proposition 2.5 that f is a minimal polynomial of $T_0, \ldots, T_m$ if, and only if $(f, \beta(f,T))$ is an MR of $T_0, \ldots, T_m$. Thus to obtain an MR of $T_0, \ldots, T_m$, it suffices to find a minimal polynomial f and to compute the polynomial product $\beta(f,T)$. It turns out that our iterative solution of Problem 2.2 extends naturally to computing $\beta(f,T)$ as well, so that we will also obtain an MR of $T_0, \ldots, T_m$ iteratively.

2.3 Rational approximation.

We will see that Equation (3) is intimately related to rational approximation in $R[[X^{-1}]]$, where R as usual denotes an integral domain. First we establish when a rational function g/f belongs to $R[[X^{-1}]]$:

LEMMA 2.7 *If $f \in R[X]$ is monic, then f is a unit in $R[[X^{-1}]]$ and if $g \in R[X]$ satisfies* deg $g \leq$ deg f, *then $g/f \in R[[X^{-1}]]$. In fact, $\delta(g/f) =$ deg $g -$ deg f.*

PROOF. Let $d = \deg f, e = \deg g$. Firstly, if f is monic, then $1/f \in X^{-d} R[[X^{-1}]]$. For $f = X^d(1 + f_{d-1}X^{-1} + \cdots + f_0 X^{-d}) = X^d(1 - h)$ say, and so $1/f = X^{-d}(1-h)^{-1} = X^{-d} \sum_{i=0}^{\infty} h^i \in X^{-d} R[[X^{-1}]]$. Thus if $e \leq d$, $g/f \in X^{e-d} R[[X^{-1}]] \subseteq R[[X^{-1}]]$. Finally, $d + \delta(g/f) = e$ since $f \cdot (g/f) = g$, which yields $\delta(g/f)$ since $d \neq -\infty$. □

DEFINITION 2.8 *Let f be monic and* deg $g \leq$ deg f. *We say that the rational function g/f is* a rational approximation of $T = T_0 + \cdots + T_m X^m$ *if $T_i = (g/f)_i$ for $m \leq i \leq 0$.*

PROPOSITION 2.9 *Let f be monic and $T = T_0 + \cdots + T_m X^m$. Then $\beta(f,T)/f$ is a rational approximation of T if, and only if $(f, \beta(f,T))$ realizes $T_0, \ldots, T_m$.*

PROOF. From Lemma 2.7, we have $T_i = (g/f)_i$ for $m \leq i \leq 0 \Leftrightarrow (T - g/f)_i = 0$ for $m \leq i \leq 0 \Leftrightarrow \delta(T - g/f) \leq m-1 \Leftrightarrow \delta(fT - g) \leq m-1+\deg f$ since for $F, G \in R((X^{-1}))$, $\delta(F \cdot G) = \delta(F) + \delta(G)$. □

Combining the results so far, we deduce:

PROPOSITION 2.10 *Suppose that $m \leq 0$, $T_0, \ldots, T_m$ is a finite sequence over R and $T = T(X^{-1}) = T_0 + \cdots + T_m X^m$. If f is a monic solution of Equation (3), then in $R[[X^{-1}]]$,*

$$T \equiv \beta(f,T)/f \pmod{X^{m-1}}. \tag{5}$$

PROOF. We know that $(f, \beta(f,T))$ is a realization of $T_0, \ldots, T_m$ and that $\beta(f,T)/f \in R[[X^{-1}]]$ by Lemma 2.7. Hence by Proposition 2.9, $\delta(T - \beta(f,T)/f) \leq m-1 \leq -1$ i.e. $T - \beta(f,T)/f = F$ for some $F \in X^{m-1}R[[X^{-1}]]$. □

Thus Equation (3) and rational approximation are intimately related. Over a field, an algorithm which computes a minimal realization $(f, \beta(f,T))$ of a finite sequence can be used to compute a ('minimal') rational approximation since f can be made monic.

3 Sequences and annihilating polynomials.

We give some basic definitions and examples for linear recurring and finite sequences.

3.1 Linear recurring sequences.

We continue the conventions of Sections 1 and 2: $\sum_{\emptyset} = 0$, R is a commutative integral domain with $1 \neq 0$ and we let $f,\ g,\ h \in R[X]$ denote polynomials with coefficients from R. For a set E, E^* denotes $E \setminus \{0\}$.

It will be convenient to work with the ring $R((X^{-1}))$ of Laurent series in X^{-1} with coefficients from R. For $-\infty < i < \infty$, F_i denotes the i^{th} coefficient of $F \in R((X^{-1}))$ and a typical element of $R((X^{-1}))$ is $F = \sum_{i \leq d} F_i X^i$ for some integer $d < \infty$. Both $R[X]$ and $R[[X^{-1}]]$ are subrings of $R((X^{-1}))$ and $R[X]$ acts on $R((X^{-1}))$ in the standard way (by multiplication in $R((X^{-1}))$; we let $\cdot$ denote this multiplication).

Let $\mathrm{Seq}(R)$ denote the (additive) abelian group of (negatively-indexed) infinite sequences over R i.e. functions $\{\ldots, -1, 0\} \to R$. The value of $S \in \mathrm{Seq}(R)$ at $i \leq 0$ is written as S_i. Thus for $S,\ T \in \mathrm{Seq}(R)$, $S_i,\ T_i \in R$ and $(S+T)_i = S_i + T_i$ for all $i \leq 0$. The *generating function of* $S \in \mathrm{Seq}(R)$ is $\Gamma(S) = \sum_{i \leq 0} S_i X^i \in R[[X^{-1}]]$.

We define $\circ : R[X] \times \mathrm{Seq}(R) \to \mathrm{Seq}(R)$ by

$$(f \circ S)_i = (f \cdot \Gamma(S))_i = \sum_{j=0}^{\deg f} f_j S_{i-j} \text{ for } i \leq 0.$$

We always have $f \cdot \Gamma(S) \in R((X^{-1}))$ and $f \cdot \Gamma(S) - \Gamma(f \circ S) \in XR[X]$.

PROPOSITION 3.1 *The mapping $\circ$ makes $\mathrm{Seq}(R)$ into a unitary $R[X]$-module.*

PROOF. We verify that for $S \in \mathrm{Seq}(R)$, $(f \cdot g) \circ S = f \circ (g \circ S)$, the other axioms being trivially satisfied. By linearity, we can assume that $f(X) = X^d$ and $g(X) = X^e$. If $i \leq 0$, then $((X^{d+e}) \circ S)_i = S_{i-d-e} = (X^e \circ S)_{i-d} = (X^d \circ (X^e \circ S))_i$, as required. □

The *annihilator ideal* of $S \in \mathrm{Seq}(R)$ is by definition $\mathrm{Ann}(S) = \{f : f \circ S = 0\}$. A sequence S will be called *linear recurring* if it is a torsion element of $\mathrm{Seq}(R)$ i.e. if $\mathrm{Ann}(S) \neq (0)$. Clearly $f \in \mathrm{Ann}(S)$ if, and only if $\Gamma(f \circ S) = 0$ if, and only if $f \cdot \Gamma(S) \in XR[X]$. Thus if $f \in \mathrm{Ann}(S)$ is monic, then $\Gamma(S)$ is the *rational function* $(f \cdot \Gamma(S))/f \in R[[X^{-1}]]$ by Lemma 2.7.

Let S be a linear recurring sequence. We say that f is a *minimal polynomial* of S, written $f \in \mathrm{Min}(S)$, if $f \in \mathrm{Ann}(S)^*$ and for any $g \in \mathrm{Ann}(S)^*$, $\deg f \leq \deg g$. We will call the degree of an element of $\mathrm{Min}(S)$ the *linear complexity* of S. Since R is an integral domain, S has linear complexity 0 if, and only if $S = 0$ if, and only if $\mathrm{Ann}(S) = R[X]$.

EXAMPLE 3.2 *Let $\mathcal{F} = 1, 1, 2, 3, \ldots$ be the Fibonacci sequence, where each term is the sum of the previous two:*

$$\mathcal{F}_{i-2} = \mathcal{F}_{i-1} + \mathcal{F}_i \text{ for } i \leq 0.$$

We show that $\phi = X^2 - X - 1 \in \mathrm{Min}(\mathcal{F})$. We have

$$\begin{aligned} \phi \cdot \Gamma(\mathcal{F}) &= \sum_{i \leq 0} (\mathcal{F}_i X^{i+2} - \mathcal{F}_i X^{i+1} - \mathcal{F}_i X^i) \\ &= \sum_{j \leq 0} (\mathcal{F}_{j-2} - \mathcal{F}_{j-1} - \mathcal{F}_j) X^j + \mathcal{F}_{-1} X + \mathcal{F}_0 X^2 - \mathcal{F}_0 X = X^2. \end{aligned}$$

Thus $\phi \in \mathrm{Ann}(\mathcal{F})$, as expected.

We now show that $\phi \in \mathrm{Min}(\mathcal{F})$. It is clear that the linear complexity of $\mathcal{F}$ is at least 1. If for some a, b with $a \neq 0$, $a\mathcal{F}_{i-1} = b\mathcal{F}_i$ for all $i \leq 0$, then $(aX - b) \cdot \Gamma(\mathcal{F}) = aX$ as before, so that $(aX - b)X^2 = (aX - b)\phi \cdot \Gamma(\mathcal{F}) = aX\phi$ and $aX = 0$, which is impossible.

Similarly, if $\mathcal{F}'$ denotes the Fibonacci sequence $0, 1, 1, 2, \ldots$, then $\phi \cdot \Gamma(\mathcal{F}') = X$, $\phi \in \mathrm{Ann}(\mathcal{F}')$ and $\phi \in \mathrm{Min}(\mathcal{F}')$.

Moreover X^2/ϕ is a rational approximation of $1 + X + \cdots + \mathcal{F}_m X^m$ for all $m \leq 0$:

$$\delta(1 + X + \cdots + \mathcal{F}_m X^m - X^2/\phi) \leq m - 1 \text{ for all } m \leq 0 \tag{6}$$

and likewise

$$\delta(X + X^2 + \cdots + \mathcal{F}'_m X^m) - X/\phi) \leq m - 1 \text{ for all } m \leq 0. \tag{7}$$

As noted above, $f \in \mathrm{Ann}(S)$ implies that $f \cdot \Gamma(S) \in XR[X]$. We give a simple application of this fact to *impulse response sequences*. Recall that if f is monic and $d = \deg f \geq 1$, the linear recurring sequence $S^{\langle f \rangle}$ with $S_i^{\langle f \rangle} = 0$ for $2 - d \leq i \leq 0$, $S_{1-d}^{\langle f \rangle} = 1$ and $f \in \mathrm{Ann}(S)$ is called an impulse response sequence; see p.402, Lidl & Niederreiter (1983). Clearly $\delta(\Gamma(S^{\langle f \rangle})) = 1 - d$. For example, $\mathcal{F}' = S^{\langle \phi \rangle}$.

PROPOSITION 3.3 *We have* $\mathrm{Ann}(S^{\langle f \rangle}) = fR[X]$ *and* $f \in \mathrm{Min}(S^{\langle f \rangle})$.

PROOF. Clearly $fR[X] \subseteq \mathrm{Ann}(S^{\langle f \rangle})$. Conversely, let $g \in \mathrm{Ann}(S^{\langle f \rangle})$ and put $\Gamma = \Gamma(S^{\langle f \rangle})$. Then $g \cdot \Gamma = Xh$ for some $h \in R[X]$, and $f \cdot \Gamma = X$ from the definition of $S^{\langle f \rangle}$. Thus $g = (X^{-1} f \cdot \Gamma) \cdot g = f \cdot (X^{-1} g \cdot \Gamma) = fh$. The second assertion follows from the first. □

Again, $\phi \in \mathrm{Min}(\mathcal{F}')$. Since f is monic and $f \cdot \Gamma(S^{\langle f \rangle}) = X$, $1/f = X^{-1}\Gamma(S^{\langle f \rangle})$, so $1/f \in X^{-d}R[[X^{-1}]]$ and $\delta(1/f) = -1 + \delta(\Gamma(S^{\langle f \rangle})) = -d$, as already seen.

3.2 The set of annihilating polynomials of a finite sequence.

For finite sequences, we will work with the ring $R[X^{-1}, X]$ of Laurent *polynomials* with coefficients from R, using F, G, H to denote typical elements; thus $F = \sum_{i=e}^{d} F_i X^i$ for some integers $-\infty < e, d < \infty$. The product of $f \in R[X]$ and $F \in R[X^{-1}, X]$ will be written as $f \cdot F$.

From now on, *the letters* m, n *always denote* m, $n \leq 0$ and $S|m$ denotes a typical finite sequence $S|m : \{m, \ldots, 0\} \to R$, with $1 - m \geq 0$ terms $S_i \in R$ for $m \leq i \leq 0$ and last term S_m. For example, $S|m$ could be $S|\{m, \ldots, 0\}$, the restriction of $S \in \mathrm{Seq}(R)$ to $\{m, \ldots, 0\}$. We write

$$\Gamma(S|m) = S_0 + \cdots + S_m X^m \in R[X^{-1}]$$

for the *'generating polynomial'* of $S|m$. We define

$$(f \circ S|m)_i = (f \cdot \Gamma(S|m))_i = \sum_{j=0}^{\deg f} f_j S_{i-j} \text{ if } m + \deg f \leq i \leq 0.$$

When m is understood and $m + \deg f \leq 0$, we will write $(f \circ S)_i$ for $(f \circ S|m)_i$.

What we have seen so far suggests

DEFINITION 3.4 *The* set of annihilating polynomials *of* $S|m$ *is defined to be*

$$\mathrm{Ann}(S|m) = \{f \ : \ (f \circ S|m)_i = 0 \text{ for } m + \deg f \leq i \leq 0\}.$$

For the Fibonacci sequence $\mathcal{F}$, $\phi \in \mathrm{Ann}(\mathcal{F}|m)$ for all $m \leq 0$. Two trivial cases are (i) $S|m$ is the all-zero sequence if, and only if $1 \in \mathrm{Ann}(S|m)$ and (ii) if $\deg f \geq 1 - m$, then $f \in \mathrm{Ann}(S|m)$ since we are summing over an empty index set. It is clear that for any m and $S \in \mathrm{Seq}(R)$, $\mathrm{Ann}(S) \subseteq \mathrm{Ann}(S|\{m, \ldots, 0\})$. The reader can easily check:

PROPOSITION 3.5 $\mathrm{Ann}(S|m-1) \subseteq \mathrm{Ann}(S|m)$.

It is straightforward to see that $\mathrm{Ann}(S|m)$ is not an ideal in general. Consider for example the sequence with $S_0 = 0$, $S_{-1} = 1$ and $m = -1$: X^2, $X^2 - X \in \mathrm{Ann}(S|m)$ (trivially), but their difference $X \notin \mathrm{Ann}(S|m)$ since $(X \circ S|m)_0 = S_{0-1} \neq 0$. The following however, will be sufficient for our purposes:

PROPOSITION 3.6 *(i)* $((f+g) \circ S)_i = (f \circ S)_i + (g \circ S)_i$ *for* $m + \max\{\deg f, \deg g\} \leq i \leq 0$ *and (ii)* $((r \cdot X^j f) \circ S)_i = r \cdot (f \circ S)_{i-j}$ *for* $r \in R$, $j \geq 0$, $m + j + \deg f \leq i \leq 0$.

PROOF. Straightforward verification. □

Problem 2.2 is now: find an $f \in \mathrm{Min}(S|m)$. It is immediate that if

$$\mu_0 = \begin{cases} 1 & \text{if } S_0 = 0 \\ X & \text{otherwise,} \end{cases}$$

then $\mu_0 \in \mathrm{Min}(S|0)$ since zero is the only zero-divisor in R.

EXAMPLE 3.7 *Let* $m < 0$, $S_i = 0$ *for* $m + 1 \leq i \leq 0$ *and* $S_m \neq 0$. *Then certainly* $1 \notin \mathrm{Ann}(S|m)$, $1 \in \mathrm{Min}(S|m+1)$ *and* $X^{1-m} \in \mathrm{Ann}(S|m)$.

4 Two constructions of annihilating polynomials.

We use μ_0 (defined in Section 3.2) as the inductive basis for constructing polynomials in $\mathrm{Ann}(S|m)$. This induction splits naturally into two subcases, depending the non-existence (Section 4.1) or the existence (Section 4.2) of previous annihilating polynomials.

Suppose that $m < 0$, $f \in \mathrm{Ann}(S|m+1)^*$ and $\deg f \leq -m$. More often than not, $f \notin \mathrm{Ann}(S|m)$; This is the case precisely when $(f \circ S)_{m+\deg f}$ is non-zero. Thus we define the *obstruction* $\mathcal{O}f \in R$ by

$$\mathcal{O}f = \begin{cases} (f \circ S)_{m+\deg f} & \text{if } \deg f \leq -m \\ 0 & \text{otherwise.} \end{cases}$$

(We conventionally put $\mathcal{O}f = 0$ if $\deg f > -m$ since $f \in \mathrm{Ann}(S|m)$ vacuously.) Now for $m < 0$ and $f \in \mathrm{Ann}(S|m+1)^*$, we have $f \in \mathrm{Ann}(S|m)$ if, and only if the obstruction $\mathcal{O}f$ vanishes; $\mathcal{O}f$ is called the *discrepancy* of f in the LFSR literature.

We have already seen an instance of $\mathcal{O}f$ in Example 3.7: if $S|m$ has precisely $m+1$ leading zeroes, then $1 \notin \mathrm{Ann}(S|m)$ since $\mathcal{O}1 = S_{m+0} \neq 0$.

4.1 First steps.

We begin with the case $S_0 \neq 0, m < 0$ and $f \in \mathrm{Ann}(S|m+1)^* \setminus \mathrm{Ann}(S|m)$:

EXAMPLE 4.1 *Suppose that* $S_0 \neq 0,\ m+1 \leq 0$ *and* $\mathcal{O}f = (f \circ S)_{m+\deg\ f} \neq 0$, *so that* $\deg\ f \leq -m$. *Since* R *is commutative,*

$$\begin{aligned} 0 &= S_0 \cdot \mathcal{O}f - \mathcal{O}f \cdot S_0 = S_0 \cdot \mathcal{O}f - (\mathcal{O}f \circ S)_0 = (S_0 \cdot f \circ S)_{m+\deg\ f} - (\mathcal{O}f \circ S)_0 \\ &= ((S_0 \cdot X^{-m-\deg\ f} f - \mathcal{O}f) \circ S)_0. \end{aligned}$$

We claim that

$$h = S_0 X^{-m-\deg\ f} f - \mathcal{O}f \in \mathrm{Ann}(S|m).$$

Certainly $h \in R[X]$ *and* $\deg\ h = -m \geq 1$ *since* R *has no zero-divisors. Also,* $(h \circ S)_0 = 0 : (h \circ S)_i = 0$ *for* $m + \deg\ h \leq i \leq 0$. *Thus if we began with* $f \in \mathrm{Ann}(S|m+1)^* \setminus \mathrm{Ann}(S|m)$, *we have produced an* $h \in \mathrm{Ann}(S|m)^*$.

Thus we have used $S_0 \neq 0$ and the commutativity of R to construct an annihilating polynomial for $S|m$.

EXAMPLE 4.2 *Consider 1,1,2, the first three terms of the Fibonacci sequence* $\mathcal{F}$. *Certainly* $X \in \mathrm{Ann}(1)$, *so we begin with* $m+1 = 0$ *and compute* $\mathcal{O}X = (X \circ \mathcal{F})_{-1+1} = \mathcal{F}_{-1} = 1$ *i.e.* $X \notin \mathrm{Ann}(1,1)$. *We apply Example 4.1 with* $m+1 = 0$ *and* $f = X$, *obtaining*

$$h = \mathcal{F}_0 \cdot X^{1-\deg\ f} f - \mathcal{O}f = 1 \cdot X^{1-1} X - 1$$

i.e. $X - 1 \in \mathrm{Ann}(1,1)$.

We continue with $m+1 = -1$ *and* $X - 1$: $\mathcal{O}(X-1) = ((X-1) \circ \mathcal{F})_{-2+1} = \mathcal{F}_{-2} - \mathcal{F}_{-1} = 1$ *i.e.* $X - 1 \notin \mathrm{Ann}(1,1,2)$. *We apply Example 4.1 again with* $f = X - 1$, *obtaining*

$$h = \mathcal{F}_0 \cdot X^{2-\deg\ f} f - \mathcal{O}f = 1 \cdot X^{2-1}(X-1) - 1 = \phi$$

i.e. $\phi \in \mathrm{Ann}(1,1,2)$.

We now combine and summarise Examples 3.7 and 4.1.

PROPOSITION 4.3 *Let* $m < 0,\ f \in \mathrm{Ann}(S|m+1)$. *For* $\mathcal{O}f \neq 0$ *and*

$$h = \begin{cases} X^{1-m} & \text{if } \deg\ f = \deg\ f_0 = 0 \\ S_0 X^{-m-\deg\ f} f - \mathcal{O}f & \text{if } \deg\ f_0 = 1, \end{cases}$$

$h \in \mathrm{Ann}(S|m)^*$.

If $m < 0$, then $\mathrm{Ann}(S|m+1) \subseteq \mathrm{Ann}(S|0)$. Thus if $f \in \mathrm{Ann}(S|m+1)$ and $S_0 \neq 0$, then $\deg f \geq 1$. Theorem 2.11(a), (b) of Norton (1995a) is the special case $f = X - r$ where $r \in R^*$ is the 'common ratio of $S|(m+1)$', $S_0 \neq 0$, and $f \in \mathrm{Min}(S|m+1)$ since $\deg f = 1$. We will see that h of Proposition 4.3 is a minimal polynomial whenever $\deg f = 1$.

4.2 Second steps.

Suppose now that $m+1 < n \leq 0$ for some n, and $\mathcal{O}f = (f \circ S)_{m+\deg f} \neq 0$, $\mathcal{O}g = (g \circ S)_{n-1+\deg g} \neq 0$ (so that $\deg f \leq -m$ and $\deg g \leq 1-n$). In similar vein,

$$\begin{aligned} 0 &= \mathcal{O}g \cdot \mathcal{O}f - \mathcal{O}f \cdot \mathcal{O}g = \mathcal{O}g \cdot (f \circ S)_{m+\deg f} - \mathcal{O}f \cdot \mathcal{O}g \\ &= ((\mathcal{O}g \cdot X^{d-\deg f} f) \circ S)_{m+d} - \mathcal{O}f \cdot \mathcal{O}g \end{aligned}$$

where d is any integer satisfying $\deg f \leq d \leq -m$. If we also choose $d \geq \epsilon(g) = -m+n-1+\deg g$, then $\epsilon(g) + m = n - 1 + \deg g$ and $((X^{d-\epsilon(g)} g) \circ S)_{m+d} = (g \circ S)_{n-1+\deg g} = \mathcal{O}g$. Thus we can rewrite the previous identity as

$$\Big((\mathcal{O}g \cdot X^{d-\deg f} f - \mathcal{O}f \cdot X^{d-\epsilon(g)} g) \circ S \Big)_{m+d} = 0.$$

Note that $\mathcal{O}g \cdot X^{d-\deg f} f - \mathcal{O}f \cdot X^{d-\epsilon(g)} g \neq 0$ since R is an integral domain, for otherwise $d = d - \epsilon(g) + \deg g$ and $-m+n-1 = 0$, contradicting $m+1 < n$. Since we are ultimately interested in minimality, we take the smallest such d i.e. $d = \max\{\deg f, \epsilon(g)\} \leq -m$. This discussion motivates our second construction:

DEFINITION 4.4 *Let $m+1 < n \leq 0$, $f \in \mathrm{Ann}(S|m+1)$, $g \in \mathrm{Ann}(S|n)$ and $\mathcal{O}f$, $\mathcal{O}g \neq 0$. We define $\epsilon(g) = -m+n-1+\deg g$ and*

$$[f,g] = \mathcal{O}g \cdot \ X^{\max\{0, \epsilon(g) - \deg f\}} f - \mathcal{O}f \cdot \ X^{\max\{\deg f - \epsilon(g), 0\}} g \in R[X].$$

We next show that $\deg [f,g] = \max\{\deg f, \epsilon(g)\}$, $[f,g] \in \mathrm{Ann}(S|m+1)$ and $\mathcal{O}[f,g] = [\mathcal{O}g, \mathcal{O}f]$, a zero commutator. This justifies our use of the [,] notation.

PROPOSITION 4.5 *Let $m+1 < n \leq 0$, $f \in \mathrm{Ann}(S|m+1)$, $g \in \mathrm{Ann}(S|n)$ and $\mathcal{O}f, \mathcal{O}g \neq 0$. Then* $\deg [f,g] = \max\{\deg f, \epsilon(g)\} \leq -m$ *and* $[f,g] \in \mathrm{Ann}(S|m)$.

PROOF. Put $h = [f,g]$ and $d = \max\{\deg f, \epsilon(g)\}$. The degree of the first summand of h is d and the degree of the second is $d + m - n + 1 < d$ and by construction, $d \leq -m$. We have already seen that $(h \circ S)_{m+d} = 0$ and so we need only check that $(h \circ S)_i = 0$ for $m + 1 + d \leq i \leq 0$. Let $\epsilon(g) = -m + n - 1 + \deg g$ as before. Then for $i \leq 0$

$$(h \circ S)_i = \mathcal{O}g \cdot (f \circ S)_{i-d+\deg f} - \mathcal{O}f \cdot (g \circ S)_{i-d+\epsilon(g)}.$$

Now $m+1+d \le i \le 0$ easily implies that $m+1+\deg f \le i-d+\deg f \le 0$ and $n+\deg g = m+1+\epsilon(g) \le i-d+\epsilon(g) \le 0$, so that $(h \circ S)_i = 0$ for $m+d \le i \le 0$ and $[f,g] \in \mathrm{Ann}(S|m)$. $\square$

EXAMPLE 4.6 *We can now produce* $\phi \in \mathrm{Ann}(0,1,1,2) = \mathrm{Ann}(\mathcal{F}'|-3)$. *We begin with* $1 \in \mathrm{Ann}(0)$, $m+1=0$ *and* $\mathcal{O}1 = \mathcal{F}'_{-1+0} = 1$. *So* $1 \notin \mathrm{Ann}(0,1)$, *but* $X^2 \in \mathrm{Ann}(0,1)$ *from Proposition 4.3. We continue with* $m+1=-1$ *and test whether* $X^2 \in \mathrm{Ann}(0,1,1)$: $\mathcal{O}X^2 = (X^2 \circ \mathcal{F}')_{-2+2} = \mathcal{F}'_{-2} = 1$ *i.e.* $X^2 \notin \mathrm{Ann}(0,1,1)$.

So we apply the [,] *construction with (a)* $m+1=-1$, $f=X^2$, $\mathcal{O}f=1$ *and (b)* $n=0$, $g=1$ *and* $\mathcal{O}g=1$. *We have* $\epsilon(g) = -m+n-1+\deg g = 2+0-1+0=1$ *and*

$$\begin{aligned}[f,g] &= \mathcal{O}g \cdot X^{\max\{0,\epsilon(g)-\deg f\}} f - \mathcal{O}f \cdot X^{\max\{\deg f-\epsilon(g),0\}} g \\ &= 1 \cdot X^{\max\{0,1-2\}} X^2 - 1 \cdot X^{\max\{2-1,0\}} 1 = X^2 - X.\end{aligned}$$

Thus $X^2 - X \in \mathrm{Ann}(0,1,1)$. *We continue with* $m+1=-2$ *and test whether* $X^2-X \in \mathrm{Ann}(0,1,1,2)$: $\mathcal{O}(X^2-X) = ((X^2-X)\circ\mathcal{F}')_{-3+2} = \mathcal{F}'_{-3} - \mathcal{F}'_{-2} = 1$. *So* $X^2-X \notin \mathrm{Ann}(0,1,1,2)$.

We apply the [,] *construction with (a)* $m+1=-2$, $f=X^2-X$, $\mathcal{O}f=1$ *and (b)* $n=0$, $g=1$ *and* $\mathcal{O}g=1$. *We have* $\epsilon(g) = -m+n-1+\deg g = 3+0-1+0=2$ *and*

$$\begin{aligned}[f,g] &= \mathcal{O}g \cdot X^{\max\{0,\epsilon(g)-\deg f\}} f - \mathcal{O}f \cdot X^{\max\{\deg f-\epsilon(g),0\}} g \\ &= 1 \cdot X^{\max\{0,2-2\}} (X^2-X) - 1 \cdot X^{\max\{2-2,0\}} 1 = \phi.\end{aligned}$$

Thus we have constructed $\phi \in \mathrm{Ann}(0,1,1,2)$.

5 The minimality of certain annihilating polynomials.

We show how certain choices in the inductive constructions of Section 4 lead to minimal polynomials.

We generically let $\mu(S|m)$ (or μ_m for short) denote a minimal polynomial of $S|m$. We always have $\deg \mu_n \le \deg \mu_m \le 1-m$ if $n<m$.

The *linear complexity* of $S|m$ is $\kappa(S|m) = \deg \mu(S|m)$, where $\mu \in \mathrm{Min}(S|m)$. Then $\kappa(S|0)=0$ if $S_0=0$ and $\kappa(S|0)=1$ otherwise. Also, either $\kappa(S|m-1)=\kappa(S|m)$ or $\kappa(S|m-1) > \kappa(S|m)$ by Proposition 3.5. We will often abbreviate $\kappa(S|m)$ to κ_m when $S|m$ is an arbitrary finite sequence with last term S_m.

As in Section 2.2, $\beta(f,S|m) \in XR[X]$ is

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

where $f_j = 0$ for $j > \deg f$. For example, $\beta(X^{1-m}, S|m) = \sum_{i=1}^{1-m} S_{i-1+m} X^i$. We have $\deg \beta(f, S|m) \leq \deg f$, $\beta(r, S|m) = 0$ if $r \in R$ and $\beta(\ , S|m)$ is R-linear i.e. $\beta(uf + vg, S|m) = u\beta(f, S|m) + v\beta(g, S|m)$ for any $u,\ v \in R$.

We begin by restating Equation (4) using $\Gamma(S|m)$:

PROPOSITION 5.1 *$f \cdot \Gamma(S|m) = F + \sum_{i=m+\deg f}^{0} (f \cdot \Gamma(S|m))_i X^i + \beta(f, S|m)$ where $\delta(F) \leq m - 1 + \deg f$.*

The proof of the following Minimality Lemma is similar in spirit to that of Corollary 3.25 of Norton (1995a), and was partly suggested by Lemma 1 of Reeds & Sloane (1985):

LEMMA 5.2 *(cf. Massey (1969), Proposition 3.5 and Theorem 3.7 of Norton (1995a).) If $m < 0$ and $g \in \mathrm{Ann}(S|m+1)$, $\mathcal{O}g \neq 0$, then for all $f \in \mathrm{Ann}(S|m)^*$, $\deg f \geq 1 - m - \deg g$.*

PROOF. We expand $h = g\beta(f, S|m) - f\beta(g, S|m) \in XR[X]$ using Proposition 5.1. Since $f \in \mathrm{Ann}(S|m)$, $f \cdot \Gamma(S|m) = F + \beta(f, S|m)$ where $\delta(F) < m + \deg f$. Also, $g \in \mathrm{Ann}(S|m+1)$ and $\mathcal{O}g = (g \circ S)_{m+\deg g}$ imply that

$$g \cdot \Gamma(S|m) = G + \mathcal{O}g \cdot X^{m+\deg g} + \beta(g, S|m)$$

where $\delta(G) < m + \deg g$. This gives

$$\begin{aligned} h &= g(f \cdot \Gamma(S|m) - F) - f(g \cdot \Gamma(S|m) - G - \mathcal{O}g \cdot X^{m+\deg g}) \\ &= \mathcal{O}g \cdot X^{m+\deg g} f + fG - gF. \end{aligned}$$

Since $\delta(fG - gF) < m + \deg f + \deg g$ and R is an integral domain, $\mathcal{O}g \cdot f_{\deg f} \neq 0$ is the non-zero leading coefficient of h and $m + \deg f + \deg g = \deg h \geq 1$. □

We first treat the case $\kappa_{m+1} = \kappa_0$ by exhibiting polynomials which attain the lower bound of the previous result, and which are therefore minimal.

PROPOSITION 5.3 *Let $m < 0$ and $\mu_{m+1} \in \mathrm{Min}(S|m+1)$. If $\deg \mu_{m+1} = \deg \mu_0$, $\mathcal{O}\mu_{m+1} \neq 0$ and*

$$\mu(S|m) = \begin{cases} X^{1-m} & \text{if } \deg \mu_0 = 0 \\ S_0 X^{-m-1} \mu_{m+1} - \mathcal{O}\mu_{m+1} & \text{if } \deg \mu_0 = 1, \end{cases}$$

then $\deg \mu(S|m) = \max\{\deg \mu_{m+1}, 1 - m - \deg \mu_{m+1}\}$ and $\mu(S|m) \in \mathrm{Min}(S|m)$.

PROOF. We have already seen that $\mu(S|m) \in \mathrm{Ann}(S|m)^*$, and $\deg \mu_m = 1 - m - \deg \mu_{m+1} = \max\{\deg \mu_{m+1}, 1 - m - \deg \mu_{m+1}\}$ is trivial to verify. Hence by Lemma 5.2, $\mu_m \in \mathrm{Min}(S|m)$. □

We remark that the first case of Proposition 5.3 is valid for *any* polynomial of degree $1 - m$; in Section 7 we will see why X^{1-m} is a particularly good choice.

It now follows from Proposition 5.3 that $X \in \mathrm{Min}(1)$, $X-1 \in \mathrm{Min}(1,1)$ and $\phi \in \mathrm{Min}(1,1,2)$. (Also, since $\phi \in \mathrm{Ann}(1,1,2,3)$, $\phi \in \mathrm{Min}(1,1,2,3)$.)

We now show how to attain the lower bound when $\kappa_0 < \kappa_{m+1}$. First an important definition (*cf.* Massey (1969)):

DEFINITION 5.4 *If $\kappa_0 < \kappa_m$ (so that $m < 0$), we define the* antecedent $\alpha(S|m)$ *of κ_m by*

$$\alpha(S|m) = \min_{m<n\leq 0}\{\ n\ :\ \kappa_n < \kappa_m\}.$$

Consider the sequence $0,1$: we have $\kappa(0) = 0$ and $\kappa(0,1) = 2$ by Proposition 5.3, so $\alpha(0,1) = 0$. If $\kappa(0,1,1) = 2$, then $\alpha(0,1,1) = 0$, but if $\kappa(0,1,1) = 3$, then $\alpha(0,1,1) = -1$.

PROPOSITION 5.5 *(cf. Norton (1995a).) Suppose that $m+1 < 0$, $\mu_n \in \mathrm{Min}(S|n)$ for $m+1 \leq n \leq 0$ and $\mathcal{O}\mu_{m+1} \neq 0$. If $\kappa_0 < \kappa_{m+1}$, $a = \alpha(S|m+1)$ and $\mu_m = [\mu_{m+1}, \mu_a]$, then*

$$\deg \mu_m = \max\{\deg \mu_{m+1}, 1-m-\deg \mu_{m+1}\}$$

and $\mu_m \in \mathrm{Min}(S|m)$.

PROOF. We know from Proposition 4.5 that $\mu_m \in \mathrm{Ann}(S|m)^*$. Suppose inductively that the result is true for $S|n$, where $m+1 \leq n < 0$. By choice of a, $m+1 < a \leq 0$, $\deg \mu_{a-1} = \deg \mu_{m+1} > \deg \mu_a$ and $\deg \mu_{a-1} = 1-(a-1)-\deg \mu_a$ either by the inductive hypothesis (if $\deg \mu_a > \deg \mu_0$) or by Proposition 5.3 (if $\deg \mu_a = \deg \mu_0$). We now have $2 - \deg \mu_{m+1} = a + \deg \mu_a$ and so

$$\epsilon(\mu_a) = -m + a - 1 + \deg \mu_a = 1 - m - \deg \mu_{m+1}.$$

If $\deg \mu_m = \deg \mu_{m+1}$, then $\mu_m \in \mathrm{Min}(S|m)$ and $\deg \mu_{m+1} = \deg \mu_m \geq \epsilon(\mu_a) = 1-m-\deg \mu_{m+1}$ i.e. $\deg \mu_m = \max\{\deg \mu_{m+1}, 1-m-\deg \mu_{m+1}\}$.

If $\deg \mu_m > \deg \mu_{m+1}$, then $\deg \mu_m = \epsilon(\mu_a) = 1-m-\deg \mu_{m+1}$, $\mu_m \in \mathrm{Min}(S|m)$ by Lemma 5.2 and $\deg \mu_m = \max\{\deg \mu_{m+1}, 1-m-\deg \mu_{m+1}\}$. □

We can now show that $\phi \in \mathrm{Min}(0,1,1,2)$ using the results of this section. We saw in Example 4.6 that $1 \in \mathrm{Ann}(0)$, $X^2 \in \mathrm{Ann}(0,1)$, $X^2 - X = [X^2, 1] \in \mathrm{Ann}(0,1,1)$ and $X^2 - X - 1 = [X^2 - X, 1] \in \mathrm{Ann}(0,1,1,2)$. We have just seen that $\alpha(0,1) = 0$ i.e. if $a = \alpha(0,1)$, then $\mu_a = \mu(0) = 1$. Also, $X^2 \in \mathrm{Min}(0,1)$ by Proposition 5.3. Hence by Proposition 5.5, $X^2 - X = [X^2, 1] \in \mathrm{Min}(0,1,1)$. This implies that $\alpha(0,1,1) = 0$ and so $\phi = [X^2 - X, 1] \in \mathrm{Min}(0,1,1,2)$ by Proposition 5.5.

As noted in Example 2.3, it is easy to check that for any integer k, $X^2 - X - 1 + k(X-1) \in \mathrm{Min}(1,1,2)$, so that minimal polynomials are not unique in general. On the other hand, one can show that both $X - 1 \in \mathrm{Min}(1,1)$ and $\phi \in \mathrm{Min}(1,1,2,3)$ are unique (up to a non-zero integer multiplier).

We refer the reader to Lemma 8.1 below and to Section 4.3 of Norton (1995a) for some results on $\mathrm{Min}(S|m)$.

6 Polynomial products of annihilating and generating polynomials.

The goal of this section is to express the polynomial product $\beta(\mu_m, S|m)$ recursively. Proposition 6.2 is also applied in Lemma 8.1 below.

It is clear that $\beta(r, S|m) = 0$ if $r \in R$ and $\beta(X, S|m) = S_0 \cdot X$, so that

$$\beta(\mu_0, S|0) = \begin{cases} 0 & \text{if } S_0 = 0 \\ S_0 \cdot X & \text{otherwise.} \end{cases}$$

For $m < 0$, μ_m is obtained using products of polynomials, so we first show how $\beta(\ , S|m)$ behaves with respect to products.

6.1 A product formula.

LEMMA 6.1 *(Product Formula) Let* $d = \deg f$, $e = \deg g$ *and* $\Gamma = \Gamma(S|m)$. *If* $d + e \leq 1 - m$ *and* $G(X^{-1}) = \sum_{i=m+e}^{0} (g \cdot \Gamma)_i X^i$, *then*

$$\beta(fg, S|m) = f\beta(g, S|m) + \sum_{i=1}^{d} (f \cdot G)_i X^i.$$

PROOF. We have

$$\begin{aligned} (fg)\Gamma &= f \cdot \left(\sum_{i=m}^{0} + \sum_{i=1}^{e} \right) (g \cdot \Gamma)_i X^i = f \cdot \sum_{i=m}^{0} (g \cdot \Gamma)_i X^i + f\beta(g, S|m) \\ &= f \cdot \sum_{i=m}^{m-1+e} (g \cdot \Gamma)_i X^i + f \cdot G + f\beta(g, S|m). \end{aligned}$$

The result now follows since $\delta(f \cdot (\sum_{i=m}^{m-1+e} (g \cdot \Gamma)_i X^i)) \leq m - 1 + d + e \leq 0$ and $\delta(f \cdot G) \leq d$. □

PROPOSITION 6.2 *Let* $\deg f + \deg g \leq 1 - m$. *(i) If* $g \in \mathrm{Ann}(S|m)$, *then* $\beta(fg, S|m) = f\beta(g, S|m)$. *(ii) If* $f,\ g \in \mathrm{Ann}(S|m)$, *then* $f\beta(g, S|m) = g\beta(f, S|m)$.

PROOF. (i) The sum G of Lemma 6.1 is zero if either (a) $m + \deg g > 0$ or (b) $m + \deg g \leq 0$ and $(g \circ S|m)_i = 0$ for $m + \deg g \leq i \leq 0$. (ii) By part (i), $f\beta(g, S|m) = \beta(fg, S|m) = \beta(gf, S|m) = g\beta(f, S|m)$. □

When $\deg g \leq -m$, it will simplify the notation to define the sequence $g \circ S|m$: for $\deg g \leq -m$, we define the finite sequence $g \circ S|m$ by

$$(g \circ S|m)_i = (g \cdot \Gamma(S|m))_i \text{ for } m + \deg g \leq i \leq 0.$$

Thus for deg $g \leq -m$, $g \circ S|m$ has last term $(g \circ S)_{m+\deg\, g}$. (As in Proposition 3.1, if deg f+deg $g \leq -m$, then $(f \cdot g) \circ S|m = f \circ (g \circ S|m)$, but we will not need this.)

We will apply the Product Formula as follows:

COROLLARY 6.3 *If* deg f + deg $g \leq -m$, *then* $\beta(fg, S|m) = f\beta(g, S|m) + \beta(f, g \circ S|m)$.

PROOF. We have deg $g \leq$ deg f + deg $g \leq -m$, so $g \circ S|m$ is well-defined. Now $G(X^{-1})$ of Lemma 6.1 is $\Gamma(g \circ S|m)$ and so $\sum_{i=1}^{d}(f \cdot G)_i X^i = \beta(f, g \circ S|m)$. □

6.2 The inductive step.

The reader can easily verify that since deg $\mu_n \leq 1 - n$, $\beta(\mu_n, S|n) = \beta(\mu_n, S|m)$ for any $m \leq n$. *We can therefore simplify the notation by writing* β_n *for any* $\beta(\mu_n, S|m)$ *with* $m \leq n$.

We are now ready for the case $\kappa_{m+1} = \kappa_0$:

PROPOSITION 6.4 *Let* $m < 0$ *and* μ_n $(m \leq n \leq 0)$ *be as in Proposition 5.3. For* $\kappa_{m+1} = \kappa_0$ *and* $\mathcal{O}\mu_{m+1} \neq 0$,

$$\beta_m = \begin{cases} S_m \cdot X & \text{if } \kappa_0 = 0 \\ S_0 \cdot X^{-m-1}\beta_{m+1} & \text{if } \kappa_0 = 1. \end{cases}$$

PROOF. For the first case, $\beta(X^{1-m}, S|m) = \sum_{k=1}^{1-m} S_{k-1+m} \cdot X^k = S_m \cdot X$. For the second, $\beta_m = \beta(S_0 \cdot X^{-m-1}\mu_{m+1} - \mathcal{O}\mu_{m+1}, S|m)$. Linearity and Corollary 6.3 give

$$\beta_m = S_0 \cdot X^{-m-1}\beta_{m+1} + S_0 \cdot \beta(X^{-m-1}, \mu_{m+1} \circ S|m)$$

since $\mathcal{O}\mu_{m+1} \in R$ and deg $\mu_{m+1} \leq 1$. The second summand is:

$$\begin{aligned} S_0 \cdot \beta(X^{-m-1}, \mu_{m+1} \circ S|m) &= S_0 \cdot \sum_{k=1}^{-m-1} (\mu_{m+1} \circ S)_{k+m+1} \cdot X^k \\ &= S_0 \cdot \sum_{j=m+2}^{0} (\mu_{m+1} \circ S)_j \cdot X^{j-m-1}. \end{aligned}$$

Since $(\mu_{m+1} \circ S)_j = 0$ for $m + 1 + \deg\, \mu_{m+1} \leq j \leq 0$ and deg $\mu_{m+1} = \kappa_{m+1} = 1$, the second summand is zero and $\beta_m = S_0 \cdot X^{-m-1}\beta_{m+1}$, as required. □

In Example 4.2, $\beta(1) = \beta(1,1) = X$, $\beta(1,1,2) = X^2$. Now for the case $\kappa_0 < \kappa_{m+1}$:

PROPOSITION 6.5 *Let* $m + 1 < 0$, $\kappa_0 < \kappa_{m+1}$, μ_n *be as in Proposition 5.3 for* $m \leq n \leq 0$ *and* $\mathcal{O}\mu_{m+1} \neq 0$. *If* $a = \alpha(S|m+1)$, *then*

$$\beta_m = \mathcal{O}\mu_a \cdot X^{\max\{0, \epsilon(\mu_a) - \deg\, \mu_{m+1}\}}\beta_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^{\max\{\deg\, \mu_{m+1} - \epsilon(\mu_a), 0\}}\beta_a$$

where $\epsilon(\mu_a) = -m + a - 1 + \deg\, \mu_a$.

PROOF. We know that

$$\mu_m = \mathcal{O}\mu_a \cdot X^{\max\{0,\epsilon(\mu_a)-\deg\ \mu_{m+1}\}}\mu_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^{\max\{\delta\mu_{m+1}-\epsilon(\mu_a),0\}}\mu_a,$$

so linearity gives β_m equal to

$$\mathcal{O}\mu_a \cdot \beta(X^{\max\{0,\epsilon(\mu_a)-\deg\ \mu_{m+1}\}}\mu_{m+1}, S|m) - \mathcal{O}\mu_{m+1} \cdot \beta(X^{\max\{\deg\ \mu_{m+1}-\epsilon(\mu_a),0\}}\mu_a, S|m).$$

We consider two cases: (i) $\deg\ \mu_{m+1} \geq \epsilon(\mu_a)$ and (ii) $\deg\ \mu_{m+1} \leq \epsilon(\mu_a)$. Set $d = \deg\ \mu_{m+1} - \epsilon(\mu_a)$. In case (i), $m+1 < \alpha(S|m+1) = a$ by definition, and this implies that $d + \deg\ \mu_a \leq -m$. So by Corollary 6.3,

$$\begin{aligned}\beta_m &= \mathcal{O}\mu_a \cdot \beta_{m+1} - \mathcal{O}\mu_{m+1} \cdot \beta(X^d\mu_a, S|m) \\ &= \mathcal{O}\mu_a \cdot \beta_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^d\beta_a - \mathcal{O}\mu_{m+1} \cdot \beta(X^d, \mu_a \circ S|m).\end{aligned}$$

The last term is

$$-\mathcal{O}\mu_{m+1} \cdot \sum_{k=1}^{d}(\mu_a \circ S)_{k-d} \cdot X^k = -\mathcal{O}\mu_{m+1} \cdot \sum_{j=1-d}^{0}(\mu_a \circ S)_j \cdot X^{j+d}.$$

Now $1 - d = 1 - \deg\ \mu_{m+1} - m + a - 1 + \deg\ \mu_a \geq a + \deg\ \mu_a$ since $-m \geq \deg\ \mu_{m+1}$ and $\mu_a \in \mathrm{Ann}(S|a)$ by hypothesis, so the last term is zero and β_m is as stated.

In case (ii), $\deg\ \mu_a \leq 1 - a$, which implies that $-d + \deg\ \mu_{m+1} \leq -m$. Then

$$\begin{aligned}\beta_m &= \mathcal{O}\mu_a \cdot \beta(X^{-d}\mu_{m+1}, S|m) - \mathcal{O}\mu_{m+1} \cdot \beta_a \\ &= \mathcal{O}\mu_a \cdot X^{-d}\beta_{m+1} + \mathcal{O}\mu_a \cdot \beta(X^{-d}\mu_{m+1}, S|m) - \mathcal{O}\mu_{m+1} \cdot \beta_a\end{aligned}$$

by Corollary 6.3. The middle summand is

$$\mathcal{O}\mu_a \cdot \sum_{k=1}^{-d}(\mu_{m+1} \circ S)_{k-d} \cdot X^k = \mathcal{O}\mu_a \cdot \sum_{j=1+d}^{0}(\mu_{m+1} \circ S)_j \cdot X^{j+d}.$$

Now $1 + d = 1 + \deg\ \mu_{m+1} - \epsilon(\mu_a) = 1 + \deg\ \mu_{m+1} + (m - a + 1 - \deg\ \mu_a)$ and as in the proof of Proposition 5.5, $-a + 1 - \deg\ \mu_a = \deg\ \mu_{m+1} - 1 = \kappa_{m+1} - 1 \geq \kappa_0 \geq 0$. Hence $1 - d \geq m + 1 + \deg\ \mu_{m+1}$. Finally, $\mu_{m+1} \in \mathrm{Ann}(S|m+1)$ and so the middle summand is zero, and β_m is as stated. □

For Example 4.6, we obtain $\beta(0) = 0$ and $\beta(0,1) = \beta(0,1,1) = \beta(0,1,1,2) = X$.

7 The minimal realization theorem and algorithm.

Our goal in this section is to derive the main result (Theorem 7.3, which combines Propositions 5.3, 5.5, 6.4, and 6.5) and the associated algorithm.

7.1 The main result.

We write $\overline{\mu}_m$ for $(\mu_m, \beta_m) \in R[X]^2$ with addition and polynomial multiplication by component. (Readers interested in linear recurrences only can ignore the second component.) Recall that $\overline{\mu}_m$ is a minimal realization (MR) of $S|m$. In this case, Proposition 5.1 yields $\delta(\mu_m \cdot \Gamma(S|m) - \beta_m) \leq m - 1 + \deg \mu_m$, so that if μ_m is monic, we have *the rational approximation*

$$\delta(\Gamma(S|m) - \beta_m/\mu_m) \leq m - 1$$

where $\deg \mu_m$ *is minimal, cf.* Sections 2.2, 2.3. Thus X/X, $X/(X-1)$, $X^2/(X^2-X-1)$ and $0/1$, X/X^2, $X/(X^2-X)$, $X/(X^2-X-1)$ are 'minimal rational functions'.

PROPOSITION 7.1 *Let* $m+1 < 0$ *and* $\overline{\mu}_n$ *be an MR for* $S|n$, *where* $m+1 \leq n \leq 0$. *If* $\kappa_0 < \kappa_{m+1}$, $\mathcal{O}\mu_{m+1} \neq 0$, $a = \alpha(S|m+1)$ *and* $d_{m+1} = 2\kappa_{m+1} + m - 1$, *then*

$$(i)\ \overline{\mu}_m = \mathcal{O}\mu_a \cdot \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^{|d_{m+1}|}\overline{\mu}_a$$

is an MR for $S|m$, *where* $\overline{\mu}_{m+1}$, $-\overline{\mu}_a$ *and* $\mathcal{O}\mu_{m+1}$, $\mathcal{O}\mu_a$ *have been interchanged if* $d_{m+1} < 0$, *and with this interchange*

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

PROOF. From Propositions 5.5 and 6.5

$$\overline{\mu}_m = \mathcal{O}\mu_a \cdot X^{\max\{0, \epsilon(\mu_a) - \deg\ \mu_{m+1}\}}\overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^{\max\{\deg\ \mu_{m+1} - \epsilon(\mu_a), 0\}}\overline{\mu}_a,$$

where $\epsilon(\mu_a) = -m + a - 1 + \deg \mu_a$ and $a + \deg \mu_a = 2 - \deg \mu_{m+1}$. Put $d = d_{m+1}$. Then $\epsilon(\mu_a) - \deg \mu_{m+1} = -m + 1 - 2\kappa_{m+1} = -d$. Hence $\max\{0, \epsilon(\mu_a) - \deg \mu_{m+1}\} = \max\{0, -d\}$, $\max\{\deg \mu_{m+1} - \epsilon(\mu_a), 0\} = \max\{d, 0\}$ and

$$\overline{\mu}_m = \begin{cases} \mathcal{O}\mu_a \cdot \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^d\overline{\mu}_a & \text{if } d \geq 0 \\ \\ \mathcal{O}\mu_a \cdot X^{-d}\overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot \overline{\mu}_a & \text{if } d < 0. \end{cases}$$

If $d < 0$ and we interchange $\overline{\mu}_{m+1}$ and $-\overline{\mu}_a$, $\mathcal{O}\mu_{m+1}$ and $\mathcal{O}\mu_a$, we obtain the right-hand side as stated:

$$\mathcal{O}\mu_{m+1} \cdot X^{-d}(-\overline{\mu}_a) - \mathcal{O}\mu_a \cdot (-\overline{\mu}_{m+1}) = \mathcal{O}\mu_a \cdot \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^{-d}\overline{\mu}_a.$$

For the updating, recall that $\deg \mu_m = \max\{\deg \mu_{m+1}, 1 - m - \deg \mu_{m+1}\} = \deg \mu_{m+1}$ if, and only if $\deg \mu_{m+1} \geq 1 - m - \deg \mu_{m+1}$ if, and only if $d \geq 0$. Thus either (i) $d \geq 0$, $\alpha(S|m) = a$ and $\overline{\mu}_{\alpha(S|m)} = \overline{\mu}_a$ or (ii) $d < 0$, $\deg \mu_m > \deg \mu_{m+1}$ (so $\alpha(S|m) = m+1$) and we have interchanged $\overline{\mu}_{m+1}$ and $\overline{\mu}_a$. Thus $\overline{\mu}_{\alpha(S|m)} = \overline{\mu}_a$, as asserted. Finally,

$$d_m = \begin{cases} 2\deg\ \mu_{m+1} + m - 2 = d - 1 & \text{if } d \geq 0 \\ 2(1 - m - \deg\ \mu_{m+1}) + m - 2 = -(2\kappa_{m+1} + m - 1) - 1 = -d - 1 & \text{if } d < 0. \end{cases}$$

□

When $m < 0$ and $\kappa_{m+1} = \kappa_0$, we require analogues of $\mu_{\alpha(S|m+1)}$ and $\mathcal{O}\mu_{\alpha(S|m+1)}$ to rewrite Proposition 5.3 in the form of Proposition 7.1:

DEFINITION 7.2 *If $m \leq 0$ and $\kappa_m = \kappa_0$, let*

$$(\overline{\mu}_{\alpha(S|m)},\ \mathcal{O}\mu_{\alpha(S|m)}) = \begin{cases} ((0,-X),\ 1) & \text{if } \kappa_0 = 0 \\ ((1,0),\ S_0) & \text{if } \kappa_0 = 1. \end{cases}$$

We can now state the main theorem of this paper:

THEOREM 7.3 *(cf. Section 7.3, Berlekamp (1968), Massey (1969)) Let $m < 0$, $\overline{\mu}_n$ be an MR of $S|n$ for $m+1 \leq n \leq 0$, $\mathcal{O}\mu_{m+1} \neq 0$ and $d_{m+1} = 2\kappa_{m+1} + m - 1$. Put $a = \alpha(S|m+1)$. Then*

$$(i)\ \mathcal{O}\mu_a \cdot \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^{|d|}\overline{\mu}_a$$

is an MR of $S|m$, where $\overline{\mu}_{m+1}$, $-\overline{\mu}_a$ and $\mathcal{O}\mu_{m+1}$, $\mathcal{O}\mu_a$ have been interchanged if $d_{m+1} < 0$, and with this interchange

$$(ii)\ \mu_{\alpha(S|m)} = \mu_a,\ d_m = |d_{m+1}| - 1.$$

PROOF. We can assume that $\kappa_{m+1} = \kappa_0$. Put $d = d_{m+1}$. Proposition 5.3 gives

$$d = \begin{cases} m-1 & \text{if } \kappa_0 = 0 \\ m+1 & \text{if } \kappa_0 = 1. \end{cases}$$

When $d \geq 0$, $d = 0,\ m = -1,\ \kappa_0 = 1$, and

$$\begin{aligned} \mathcal{O}\mu_a \cdot \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^d\overline{\mu}_a &= S_0 \cdot \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1}(1,0) \\ &= (S_0\mu_{m+1} - \mathcal{O}\mu_{m+1}, S_0\beta_{m+1}) = \overline{\mu}_m \end{aligned}$$

from Proposition 5.3. If $d < 0$ then after interchanging, $\mathcal{O}\mu_a \cdot X^{-d}\overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot \overline{\mu}_a$ is $1.X^{1-m}(1,0) - S_m(0,-X) = (X^{1-m}, S_mX)$ if $\kappa_0 = 0$ and

$$S_0X^{-1-m}\overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1}(1,0) = (S_0X^{-1-m}\mu_{m+1} - \mathcal{O}\mu_{m+1}, S_0X^{-1-m}\beta_{m+1})$$

if $\kappa_0 = 1$. In either case, we have precisely $\overline{\mu}_m$ by Propositions 5.3, 6.4.

We now prove (ii). If $d \geq 0$, then $m = -1, \deg\ \mu_0 = 1$ and $\deg\ \mu_m = -m = \deg\ \mu_0 = \deg\ \mu_{m+1}$, there is no interchanging and $\overline{\mu}_{\alpha(S|m)} = \overline{\mu}_a$. If $d < 0$, then $\overline{\mu}_a$ and $\overline{\mu}_{m+1}$ have been interchanged. We check that $\deg\ \mu_m > \deg\ \mu_{m+1}$ (which will imply that $\overline{\mu}_{\alpha(S|m)} = \overline{\mu}_a$): if $\kappa_0 = 0$, then $\deg\ \mu_m = 1 - m > 0 = \kappa_0 = \kappa_{m+1}$, whereas if $\kappa_0 = 1,\ \deg\ \mu_m = -m > 1 = \kappa_0 = \kappa_{m+1}$ since $d = m + 1 < 0$.

If $d \geq 0$, then $d_m = 2\deg\ \mu_m + m - 2 = m = |d| - 1$, whereas if $d < 0$,

$$d_m = 2\deg\ \mu_m + m - 2 = \begin{cases} 2(1-m) + m - 2 = -m = |d| - 1 & \text{if } \kappa_0 = 0 \\ 2(-m) + m - 2 = -m - 2 = |d| - 1 & \text{if } \kappa_0 = 1. \end{cases}$$

□

The proof of Theorem 7.3 now shows why X^{1-m} was a good choice in Proposition 5.3. We will see in Proposition 7.5 below that Theorem 7.3 effectively makes $\mathcal{O}\mu_a$ and $\overline{\mu}_a$ of Definition 7.2 unique.

7.2 Algorithm MR.

Theorem 7.3 clearly suggests the following algorithm statements when $m < 0$:

$\mathcal{O} := \sum_{i=0}^{\deg\ (\mu_{m+1})} (\mu_{m+1})_i \cdot S_{(m+\deg\ \mu_{m+1})-i}$;
if $(\mathcal{O} \neq 0)$ { if $(d_{m+1} < 0)$ $\{d_{m+1} := -d_{m+1}; swap(\overline{\mu}_{m+1}, \overline{\mu}'); swap(\mathcal{O}, \mathcal{O}');\}$
$\overline{\mu}_m := \mathcal{O}' \cdot \overline{\mu}_{m+1} - \mathcal{O} \cdot X^{d_{m+1}} \overline{\mu}'$;
$d_m := d_{m+1} - 1;\}$

We have written $\overline{\mu}'$ for $\overline{\mu}_a$ and $\mathcal{O}'$ for $\mathcal{O}\mu_a$, independently of m, since we need neither the actual values $\alpha(S|m+1)$, $\alpha(S|m)$, nor their provenance (*cf.* Norton (1995a)). We have also suppressed the negation in the swap, for if $\overline{\mu}_m$ is an MR for $S|m$, so is $-\overline{\mu}_m$.

The case $\mathcal{O}\mu_{m+1} = 0$ is easily incorporated into these algorithm statements: evidently we take $\overline{\mu}_m = \overline{\mu}_{m+1}$. Then $\kappa_m = \kappa_0$ or $\kappa_m > \kappa_0$ depending on κ_{m+1}, κ_0. Also, $d_m = 2\kappa_{m+1} + m - 1 = d_{m+1} - 1$ in this case, so we need only factor out the statement $d_m := d_{m+1} - 1$ from the preceding statements.

At this stage, we could initialize at $m = 0$ with

$$\overline{\mu}_0 = \begin{cases} (1,0) & \text{if } S_0 = 0 \\ (X, S_0 X) & \text{otherwise.} \end{cases}$$

Remarkably, the algorithm statements yield a $\overline{\mu}_0$ when $S_0 \neq 0$ if we define '$\overline{\mu}_1$' by $\overline{\mu}_1 = (1,0)$. For then $d_1 = 2\deg\ \mu_1 + m - 1 = -1$, $\mathcal{O} = S_0$, we swap $(1,0)$ and $(0,-X)$, S_0 and 1, to get $\overline{\mu}_0 := -(X, S_0 X)$, and $d_0 = 2\deg\ \mu_0 + m - 2$ assumes its correct value since $|-1| - 1 = 0$!

Finally, only the current values of $\overline{\mu}_{m+1}$ and d_{m+1} are used, so we can suppress their indices too, giving

Algorithm MR *(cf. p.184 of Berlekamp (1968), p.124 of Massey (1969))*

Input: $m \leq 0$, R an integral domain, $S_0, \ldots, S_m \in R$.

Output: $\overline{\mu}$, an MR for $S|m$.

$\overline{\mu} := (1,0)$; $\overline{\mu}' := (0,-X)$; $\mathcal{O}' := 1$; $d := -1$;
for $n := 0$ *downto* m *do*
{ $\mathcal{O} := \sum_{i=0}^{\deg\ \mu} \mu_i \cdot S_{(n+\deg\ \mu)-i}$; $\quad$ /* *compute* $\mathcal{O}$ */
if $(\mathcal{O} \neq 0)$ { *if* $(d < 0)$ $\{d := -d; swap(\overline{\mu}, \overline{\mu}'); swap(\mathcal{O}, \mathcal{O}');\}$ /* *update* $\overline{\mu}', \mathcal{O}'$ */

$\overline{\mu} := \mathcal{O}' \cdot \overline{\mu} - \mathcal{O} \cdot X^d \overline{\mu}'; \}$ $\quad /* \textit{update } \overline{\mu} */$

$d := d - 1; \}$ $\quad /* \textit{update } d */$

return $\overline{\mu}$.

For completeness, we give the values at the end of the iterations for 1,1,2 and 0,1,1,2:

n	$\mathcal{O}\mu_n$	$\overline{\mu}'$	$\mathcal{O}'$	$\overline{\mu}_n$	d_n
0	1	$(1,0)$	1	$(-X,-X)$	0
-1	1	$(1,0)$	1	$(-X+1,-X)$	-1
-2	-1	$(1-X,-X)$	-1	(X^2-X-1,X^2)	0

n	$\mathcal{O}\mu_n$	$\overline{\mu}'$	$\mathcal{O}'$	$\overline{\mu}_n$	d_n
0	0	$(0,-X)$	1	$(1,0)$	-2
-1	1	$(1,0)$	1	$(-X^2,-X)$	1
-2	1	$(1,0)$	1	$(-X^2+X,-X)$	0
-3	-1	$(1,0)$	1	$(-X^2+X+1,-X)$	-1.

The storage requirements of Algorithm MR are modest:

PROPOSITION 7.4 *For* $m < 0$ *(i)* $\deg \beta_{m+1} \leq \deg \mu_{m+1} \leq -m$ *and* $|d_{m+1}+1| \leq -m$ *and (ii) if* $\kappa_{m+1} > \kappa_0$, *then* $\deg \mu_{\alpha(S|m+1)} \leq -m-1$.

PROOF. The first part of (i) is trivial, and since $0 \leq \deg \mu_{m+1} \leq 1-(m+1) = -m$, $m-1 \leq 2\kappa_{m+1}+m-1 = d_{m+1} \leq 2(-m)+m-1 = -m-1$. Part (ii) follows from (i) and the fact that $\deg \mu_{\alpha(S|m+1)} < \deg \mu_{m+1}$. □

A simple counting argument using Proposition 7.4 shows that Algorithm MR computes $\overline{\mu}_m$ in at most $(1-m)(6-5m)/2$ R-multiplications, Proposition 4.8 of Norton (1995a).

Finally, we show that the initialization of Algorithm MR is effectively unique. Note first that if $m < 0$, $a = \alpha(S|m+1)$ and μ_a is as in Definition 7.2, then $\deg \mu_a < 1 - m - \deg \mu_{m+1}$.

PROPOSITION 7.5 *Let* $m < 0$, $\kappa_{m+1} = \kappa_0$, $\mathcal{O}\mu_{m+1} \neq 0$ *and* $d_{m+1} = 2\kappa_{m+1} + m - 1$. *Write* $a = \alpha(S|m+1)$ *and suppose that* $\overline{\mu}_a$, $\mathcal{O}_a$ *are as in Definition 7.2 and* $\overline{\mu}_m$ *is constructed as in Theorem 7.3. For* $\mathcal{O}' \in R^*$, $\overline{\mu}' \in R[X] \times R[X]$, *put*

$$\overline{\mu}'_m = \mathcal{O}' \cdot \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^{|d_{m+1}|} \overline{\mu}'$$

where $\overline{\mu}_{m+1}, -\overline{\mu}'$ *and* $\mathcal{O}\mu_{m+1}, \mathcal{O}'$ *have been interchanged if* $d_{m+1} < 0$. *If* $\deg \mu' < 1-m-\deg \mu_{m+1}$, *then* $\overline{\mu}'_m = \overline{\mu}_m$ *if, and only if* $(\overline{\mu}', \mathcal{O}') = (\overline{\mu}_a, \mathcal{O}_a)$.

PROOF. We know from Theorem 7.3 that

$$\overline{\mu}_m = \mathcal{O}_a \cdot \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot X^{|d_{m+1}|} \overline{\mu}_a$$

where $\overline{\mu}_{m+1}, \overline{\mu}_a$ and $\mathcal{O}_{m+1},\ \mathcal{O}_a$ have been interchanged if $d_{m+1} < 0$.

If $\kappa_0 = 0$, we have $\overline{\mu}_{m+1} = (1,0),\ d_{m+1} = m-1 < 0$ and $\deg \mu' < 1-m$. So $\overline{\mu}'_m = \overline{\mu}_m$ if, and only if $\mathcal{O}' \cdot X^{1-m}(1,0) - \mathcal{O}\mu_{m+1} \cdot \overline{\mu}' = \mathcal{O}_a \cdot X^{1-m}(1,0) - \mathcal{O}\mu_{m+1} \cdot \overline{\mu}_a$ if, and only if $\mathcal{O}' = \mathcal{O}_a$ and $\overline{\mu}' = \overline{\mu}_a$.

If $\kappa_0 = 1$, consider first the case $d_{m+1} \geq 0$. Here $\mathcal{O}' \cdot \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot \overline{\mu}' = \mathcal{O}_a \cdot \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot \overline{\mu}_a$ where $\deg \mu' < 1 = \deg \mu_{m+1}$, so that $\mathcal{O}' = \mathcal{O}_a$ and so $\overline{\mu}' = \overline{\mu}_a$. If $d_{m+1} < 0$, the hypothesis $\deg \mu' < 1 - m - \deg \mu_{m+1} = -d_{m+1} + \deg \mu_{m+1}$ enables us to equate leading coefficients in

$$\mathcal{O}' X^{-d_{m+1}} \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot \overline{\mu}' = \mathcal{O}_a X^{-d_{m+1}} \overline{\mu}_{m+1} - \mathcal{O}\mu_{m+1} \cdot \overline{\mu}_a$$

so that $\mathcal{O}' = \mathcal{O}_a$, whence $\overline{\mu}' = \overline{\mu}_a$. □

8 A transform problem revisited.

We solve a transform problem over an arbitrary field K using Algorithm MR. This requires a result on finite sequences over K (Lemma 8.1), which is of interest in its own right. When $K = GF(q)$, this is the decoding problem solved by Berlekamp's algorithm, which we briefly compare with Algorithm MR.

8.1 Solution of a transform problem.

From now on, R is a field K. The following lemma is of independent interest.

LEMMA 8.1 *(Cf. Massey (1969), Corollary to Theorem 3.)*

(i) Let $f \in \mathrm{Ann}(S|m)^*$. *If* $f \in \mathrm{Min}(S|m)$, *then* $\gcd(f, \beta(f, S|m)/X) \in K^*$. *Conversely, if* $2\deg f \leq 1-m$ *and* $\gcd(f, \beta(f, S|m)/X) \in K^*$, *then* $f \in \mathrm{Min}(S|m)$.

(ii) If $2\kappa(S|m) \leq 1-m$, *then* $S|m$ *has a unique monic MR.*

PROOF. (i) Suppose that $\deg g > 0$ and g divides $f,\ \beta(f, S|m)/X$. Since $f \in \mathrm{Ann}(S|m)^*$, $\delta(f/g \cdot \Gamma(S|m) - \beta(f, S|m)/g) \leq m + \deg f - \deg g = m + \deg (f/g)$. Hence by Proposition 2.5, $f/g \in \mathrm{Ann}(S|m)^*$ and $\deg f$ is not minimal.

Conversely, let $g \in \mathrm{Min}(S|m)$. Then $\deg g + \deg f \leq 2\deg f \leq 1-m$ and so by Proposition 6.2(ii), $f\beta(g, S|m) = g\beta(f, S|m)$. Now $\gcd(f, \beta(f, S|m)/X) \in K^*$, so f divides g and therefore $\deg f \leq \deg g$. Hence $\deg f = \deg g$ and $\deg f$ is minimal.

(ii) If $f,\ g \in \mathrm{Min}(S|m)$ and $\deg f + \deg g \leq 1-m$ then by Proposition 6.2(ii), $f\beta(g, S|m)/X = g\beta(f, S|m)/X$. The first paragraph implies that $\gcd(f, \beta(f, S|m)/X) \in K^*$ and again, f must divide g. Similarly, g divides f and since $f,\ g$ are monic, they are equal. Hence by Proposition 2.5, $S|m$ has a unique MR. □

We let ord α denote the *order* of $\alpha \in K^*$, define the *weight* of $e \in K[X]$ to be wt $e = |\{i : e_i \neq 0\}|$ and adopt the usual convention that $\prod_{\emptyset} = 1$.

PROBLEM 8.2 *Fix $b \geq 0$ and $\alpha \in K^*$. Given $S|m$, find $e \in K[X]$ such that $S|m$ is the sequence $e(\alpha^b), \ldots, e(\alpha^{b-m})$ and* wt $e \leq \min\{\text{ord } \alpha, (1-m)/2\}$.

For fixed $b \geq 0$ and $\alpha \in K^*$, we define $S^{(e)} \in \text{Seq}(K)$ by $S_j^{(e)} = e(\alpha^{b-j})$ for $j \leq 0$. If ord $\alpha < \infty$, then $S^{(e)}$ clearly has period ord α i.e. $X^{\text{ord } \alpha} - 1 \in \text{Ann}(S^{(e)})$.

PROPOSITION 8.3 *We have (i) $\Gamma(S^{(e)}) = X\omega(X)/\sigma(X)$ where*

$$\sigma(X) = \prod_{e_i \neq 0} (X - \alpha^i) \text{ and } \omega(X) = \sum_{e_i \neq 0} e_i \alpha^{bi} \prod_{e_j \neq 0, j \neq i} (X - \alpha^j),$$

(ii) σ is monic, deg $\sigma =$ wt e *and (iii) if* wt $e \leq 1 - m$, *then $\beta(\sigma, S^{(e)}|m) = X\omega$.*

PROOF. (i)

$$\begin{aligned} \Gamma(S^{(e)}) &= \sum_{j \leq 0} \left(\sum_{e_i \neq 0} e_i (\alpha^{b-j})^i \right) X^j = \sum_{e_i \neq 0} e_i \alpha^{bi} \left(\sum_{j \geq 0} (\alpha^i / X)^j \right) \\ &= \sum_{e_i \neq 0} \left(e_i \alpha^{bi} / (1 - \alpha^i / X) \right) = \sum_{e_i \neq 0} \left(X e_i \alpha^{bi} / (X - \alpha^i) \right) = X\omega(X)/\sigma(X). \end{aligned}$$

(ii) This is trivial. (iii) Put $F = \Gamma(S^{(e)}) - \Gamma(S^{(e)}|m)$ and $d = \deg \sigma$, so that $\delta(F) \leq m - 1$ and $\delta(\sigma \cdot F) \leq m - 1 + d \leq 0$. Then $\beta(\sigma, S^{(e)}|m)$ is

$$\sum_{i=1}^{d} (\sigma \cdot \Gamma(S^{(e)}|m))_i X^i = \sum_{i=1}^{d} (\sigma \cdot (\Gamma(S^{(e)}) - F))_i X^i = X\omega - \sum_{i=1}^{d} (\sigma \cdot F)_i X^i = X\omega.$$

□

In particular, $\sigma \in \text{Ann}(S^{(e)}) \subseteq \text{Ann}(S^{(e)}|m)$ and $S^{(e)}$ is always a linear recurring sequence. For the next result, we will need the formal derivative σ' of σ, which is quickly seen to be $\sigma'(X) = \sum_{e_i \neq 0} \prod_{e_j \neq 0, j \neq i} (X - \alpha^j)$.

THEOREM 8.4 *If* wt $e \leq \min\{\text{ord } \alpha, (1-m)/2\}$ *then (i) $(\sigma, X\omega)$ is an MR of $S^{(e)}|m$ and (ii) $S^{(e)}$ has a unique monic MR.*

PROOF. (i) We know that $\sigma \in \text{Ann}(S^{(e)}|m)$ and $\beta(\sigma, S^{(e)}|m) = X\omega$. from the previous result. Thus to see that $\sigma \in \text{Min}(S^{(e)}|m)$, we show that $\gcd(\sigma, \omega) = 1$ and apply Lemma 8.1(i). Evaluation of ω at a root α^k of σ gives

$$\omega(\alpha^k) = e_k \alpha^{kb} \prod_{e_j \neq 0, j \neq k} (\alpha^k - \alpha^j) = e_k \alpha^{kb} \sigma'(\alpha^k).$$

Since deg σ = wt $e \leq$ ord α, σ has distinct roots and $\gcd(\sigma, \sigma') = 1$. In particular, if α^k is a root of σ, then $\sigma'(\alpha^k) \neq 0$. Thus all the terms on the right-hand side of $\omega(\alpha^k)$ are non-zero, $\omega(\alpha^k) \neq 0$ and $\gcd(\sigma, \omega) = 1$.

(ii) Since $\sigma \in \mathrm{Ann}(S^{(e)}|m)^*$, $2\kappa(S^{(e)}|m) \leq 2\deg \sigma = 2\mathrm{wt}\ e \leq 1 - m$. Thus by Lemma 8.1(ii), $S^{(e)}|m$ has a unique monic MR. □

Combining Theorem 8.4 and the results of Section 7 now gives:

COROLLARY 8.5 *Let* $S|m = S^{(e)}|m$ *and* wt $e \leq \min\{\mathrm{ord}\ \alpha, (1-m)/2\}$ *as in Problem 8.2. Then* $(\sigma, X\omega) = \overline{\mu}(S|m)$ *is the monic MR of* $S|m$ *obtained from Algorithm MR.*

PROOF. Since $S^{(e)}|m = S|m$ and $S^{(e)}|m$ has a unique monic MR, so does $S|m$ i.e. $\overline{\mu}(S^{(e)}|m) = \overline{\mu}(S|m)$. The result now follows from Theorem 8.4. □

The determination of e in Problem 8.2 is thus reduced to factorizing σ and evaluating ω in $K[X]$. (If $S|m$ is the all-zero sequence, then $(\sigma, X\omega) = (1, 0)$ and so $e = 0$. Otherwise, if r is a root of σ and $k = \log_\alpha r$, then $e_k \neq 0$ and $e_k = \omega(r)/(r^b\sigma'(r))$.) Finally, since $\omega = \beta(\sigma, S^{(e)}|m)/X$ is required here, we may initialize $\overline{\mu}'$ to $(0, -1)$ (rather than to $(0, -X)$) in Algorithm MR.

8.2 Algebraic decoding.

Problem 8.2 applies to Algebraic Coding Theory. Here $K = GF(q)$ and $\gcd(n, q) = 1$, so that K has an n^{th} root of unity, α. We have integers $b \geq 0$, $d \geq 3$ and a cyclic code $C \subseteq K^n$ with generator polynomial $\prod_{i=0}^{d-2}(X - \alpha^{b+i})$ of designed distance $d = 2t + 1$ to correct up to t symbol errors.

A transmitted codeword $c \in C$ is received as $c + e$ for some $e \in K^n$ with $wt(e) \leq t \leq n = \mathrm{ord}\ \alpha$ and $S|m$ is the sequence of $2t = 1 - m$ known syndromes — for background material and the approach to Problem 8.2 using the Extended Euclidean Algorithm due to Sugiyama *et. al.*(1975), see Chapter 12 of MacWilliams & Sloane (1977).

We remark that in this context, Equation (5) (with the additional requirement that deg f be minimal) is an analogue of the 'key equation' for BCH and Reed-Solomon codes.

Thus Algorithm MR can be used for decoding BCH, Reed-Solomon and other codes. For examples and details, see Norton (1995b), Norton (1999).

8.3 The original application.

Given a t-error correcting BCH code and a codeword corrupted by up to t errors, Berlekamp's algorithm will correct (decode) these errors; see *loc. cit.*, 'Binary BCH codes for correcting multiple

errors', where minimal degree ensures maximum likelihood decoding. It also decodes Reed-Solomon codes, errors and erasures, Lee-metric negacyclic codes, Berlekamp (1968), and classical Goppa codes, Patterson (1975).

We remark that Algorithm MR — like Massey's algorithm — does not require the variable B (defined in Berlekamp (1968), Equation (7.320)) used to resolve an ambiguity (*loc. cit.* Equations (7.306), (7.307)) which can occur in Berlekamp's algorithm.

9 Guide to the notation.

In general, we use Roman letters for elements and Greek letters for functions. When $S|m$ is an arbitrary finite sequence with last term S_m, we abbreviate $\mu(S|m)$, $\kappa(S|m)$ etc. to μ_m, κ_m etc. For any set E, $E^* = E \setminus \{0\}$.

Symbol	Meaning
$\alpha(S\|m)$	See Definition 5.4 if $\kappa(S\|m) > \kappa(S\|0)$ and Definition 7.2 otherwise.
$\mathrm{Ann}(S\|m)$	See Definition 3.4.
$\beta(f, S\|m)$	$\sum_{i=1}^{\deg f}(f \cdot \Gamma(S\|m))_i X^i$.
$\beta(S\|n)$	$\beta(\mu(S\|n), S\|m)$ where $m \leq n$.
$\delta(F)$	$\max_{-\infty<i<\infty}\{i : F_i \neq 0\}$ if $F \neq 0$; $\delta(0) = -\infty$.
$d(S\|m)$	$2\kappa(S\|m) + m - 2$.
$\epsilon(g)$	See Definition 4.4.
f, g, h	Polynomials over R.
$(f \circ S)_i = (f \circ S\|m)_i$	$(f \cdot \Gamma(S\|m))_i$, where $m + \deg f \leq i \leq 0$.
$[f, g]$	See Definition 4.4.
F, G, H	Laurent series or polynomials over R.

F_i	i^{th} coefficient of F, $-\infty < i < \infty$.
$\mathcal{F}, \mathcal{F}'$	Fibonacci sequences.
$\Gamma(S\|m)$	$\sum_{i=m}^{0} S_i X^i$.
$\kappa(S\|m)$	Linear complexity of $S\|m$.
$m,\ n$	Integers $m,\ n \leq 0$.
$\mathrm{Min}(S\|m)$	Set of f in $\mathrm{Ann}(S\|m) \setminus \{0\}$ of minimal degree.
$\mu(S\|m)$	Minimal polynomial of $S\|m$.
$\overline{\mu}(S\|m)$	$(\mu(S\|m), \beta(\mu(S\|m), S\|m)) \in R[X]^2$.
$\mathcal{O}f$	$(f \circ S)_{m+\deg\ f}$ if $m + \deg\ f \leq 0$, otherwise 0.
$S\|m$	Finite sequence $\{m, \ldots, 0\} \to R$.
$\mathrm{Seq}(R)$	Sequences $\{\ldots, -1, 0\} \to R$.

Erratum

Lemma 4.14 of Norton (1995a), where $L \geq 1$ and $s|L = s_0, \ldots, s_{-L+1}$, is incorrectly stated. It should read as follows:

Lemma 4.14 *Let $s|L$ be a sequence over R and either (i) $f = 1$ or (ii) $1 \leq i \leq L-1$ and $f \in \mathrm{Ann}(s|i)$. Put $\beta = \beta(f, s|L-1)$ and $\delta = \delta(f \cdot \Gamma(s|L) - \beta)$. Then either (i) $\delta \leq \deg\ f$ or (ii) $\delta \leq -i + \deg\ f$. If in addition either (i) $f \notin \mathrm{Ann}(s|0)$ or (ii) $i \leq L-2$ and $f \notin \mathrm{Ann}(s|i+1)$, then either (i) $\delta = \deg\ f$ or (ii) $\delta = -i + \deg\ f$.*

The proof is unchanged.

References

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

Blahut, R. (1983). *Theory and Practice of Error Control Codes*. Addison–Wesley, Reading.

Blahut, R. (1985). *Fast algorithms for digital signal processing.* Addison–Wesley. Reading.

Brawley, J.V.,Gao, S., Mills, D. (1999). Computing composed products of polynomials. In *Finite Fields: Theory, Applications and Algorithms, R.C. Mullin and G.L. Mullen (Eds.)*, American Math. Soc. Series in Contemporary Mathematics, 1–16.

Brazil, M. (1993). Growth functions for some one–relator monoids. *Communications in Algebra* **21**, 3135–3146.

Brent, R.P., Gustavson, F.G., Yun, D.Y. (1980). Fast solution of Toeplitz systems of equations and computation of Padé approximants. *J. Algorithms* **1**, 259–295.

Chan, K.Y., Norton, G.H. (1995). A new algebraic algorithm for generating the transfer function of a trellis encoder. *IEEE Trans. on Communications* **43**, 1866–1867.

Dür, A. (1989). On computing the canonical form for a binary form of odd degree. *J. Symbolic Computation*, **8**, 327-333.

Dür, A. (1992). A fast algorithm to determine the burst-correcting limit of cyclic or shortened cyclic codes. *IEEE Trans. Information Theory* **38**, 504–509.

Dür, A., Grabmeier, J. (1993). Applying Coding Theory to sparse interpolation. *S.I.A.M. J. Computing* **22**, 695–704.

Gueret, O. (1996). A new algorithm for solving Yule-Walker and Wiener-Hopf equations. M.Sc. thesis, Department of Electrical Engineering, University of Bristol.

Kalman, R.E., Farb, P.L., and Arbib, M.A. (1969). *Topics in Mathematical Systems Theory.* McGraw-Hill.

Lidl, R., Niederreiter, H. (1983). *Finite Fields. Encyclopedia of Mathematics and its Applications* **20**. Addison-Wesley, Reading.

MacWilliams, F.J., Sloane, N.J.A. (1977). *The Theory of Error-Correcting Codes.* North-Holland Mathematical Library.

Mason, S.J., Zimmermann, H.J. (1960). *Electronic Circuits, Signals and Systems.* John Wiley.

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

Massey, J.L., Schaub, T. (1988). Linear Complexity in Coding Theory. *Lecture Notes in Computer Science* **311**, 19–32. Springer.

Matt, H.J., Massey, J.L. (1980). Determining the burst-correcting limit of cyclic codes. *IEEE Trans. Information Theory* **26**, 289–297.

Mills, W.H. (1975). Continued fractions and linear recurrences. *Mathematics of Computation* **29**, 173–180.

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

Norton, G.H. (1995b). Some decoding applications of minimal realization. *Cryptography and Coding, LNCS* **1025**, 53–62. Springer.

Norton, G.H. (1999). On minimal realization over a finite chain ring. *Designs, Codes and Cryptography* **16**, xx–xx.

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

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

Sugiyama, Y., Kasahara, M., Hirasawa, S., Namekawa, T. (1975). A method for solving the key equation for decoding Goppa codes. *Information and Control*, **27**, 87–99.

Trench, W.F. (1964). An algorithm for the inversion of finite Toeplitz matrices. *J. S.I.A.M.* **12**, 515–522.

van Tilborg, H.C.A. (1988). *An Introduction to Cryptology.* Kluwer, Boston.

Welch, L.R., Scholz, R.A. (1979). Continued fractions and Berlekamp's algorithm. *IEEE Trans. Information Theory* **25**, 19–27.

Wiedemann, D.H. (1986). Solving sparse linear equations over finite fields. *IEEE Trans. Information Theory* **32**, 54–62.