Algebraic Feedback Shift Registers

Andrew Klapper [a] Jinzhong Xu [b]

[a] *Dept. of Computer Science, 763H Anderson Hall, University of Kentucky, Lexington, KY, 40506-0046, klapper@cs.uky.edu.* *Project sponsored by the National Science Foundation under grant number NCR-9400762.*

[b] *Dept. of Computer Science, 763H Anderson Hall, University of Kentucky, Lexington, KY, 40506-0046, abc@ms.uky.edu.*

Abstract

A general framework for the design of feedback registers based on algebra over complete rings is described. These registers generalize linear feedback shift registers and feedback with carry shift registers. Basic properties of the output sequences are studied: relations to the algebra of the underlying ring; synthesis of the register from the sequence (which has implications for cryptanalysis); and basic statistical properties. These considerations lead to security measures for stream ciphers, analogous to the notion of linear complexity that arises from linear feedback shift registers. We also show that when the underlying ring is a polynomial ring over a finite field, the new registers can be simulated by linear feedback shift registers with small nonlinear filters.

Key words: cryptography; feedback shift register; complete ring; stream cipher; pseudo-random number generator.

1 Introduction

Linear Feedback Shift Registers (LFSRs) [3] have long been the basis of most research on stream ciphers. Their theory is used both for cryptanalysis [14,7] and for the design of (hopefully) secure keystream generators [15,16]. The importance of LFSRs comes from two facts. They are extremely fast and simple from an engineering point of view, and they have associated with them a number of algebraic structures that make the analysis of their properties tractable. These structures are based on the algebra of power series in one indeterminate over the field with two elements.

Recently, a new class of feedback registers, called Feedback with Carry Shift Registers (FCSRs) has been discovered [9,13]. These registers are nearly as

fast as LFSRs. They have an algebraic theory that parallels that of LFSRs, in this case based on the 2-adic numbers. In a series of papers, Mark Goresky and the first author considered the basic algebraic and statistical properties of FCSR sequences, as well as their use in cryptanalysis [9–13]. This construction was then extended to registers defined over certain extensions of the 2-adic numbers (purely ramified and purely unramified extensions), and some of the basic properties were outlined [8,9].

In this paper we extend FCSRs to a much more general setting, resulting in sequences over an arbitrary finite field. The registers we construct, called Algebraic Feedback Shift Registers (AFSRs), can be based in the abstract on any ring R with a principal prime ideal (π). These rings have analogues of power series, and this provides a setting for an algebraic theory that parallels the theory of LFSRs. An outline of the theory of such rings is given in Section 2, and the definitions of AFSRs are given in Section 3.

There are two principal cases of interest to us: the number field or characteristic zero case, when R is a subring of a finite extension of the rational numbers, and the function field case, when R is a polynomial ring over a finite field. In these cases the residue field $R/(\pi)$ is finite so the registers we construct generate sequences over a finite alphabet. We concentrate on the former case since, as shown in Section 9, the registers that arise in the function field case can be replaced by ordinary LFSRs with output filters that depend only on the ring R.

In Sections 4, 5, and 6 we derive the basic algebraic properties of AFSR sequences. We show that for a reasonable ring R, AFSR sequences correspond to elements α of the completion of the underlying ring (analogous to the generating function associated with an LFSR sequence); that these elements have rational representations; and that the structure of the AFSR can be determined from the denominator in the rational representation. More specifically, associated with each AFSR is an element q in R, called the connection element, that corresponds to the taps in the feedback function in the AFSR in a manner analogous to the connection polynomial associated with a LFSR. We show that the element α is rational with denominator q, $\alpha = u/q$. The numerator determines the initial state of the AFSR. We give explicit conditions on R and π under which the memory in every AFSR is bounded throughout its infinite execution. We further show that there is often an exponential representation of strictly periodic AFSR sequences. Such a sequence is of the form $a_i = (\delta\gamma^i \bmod q) \bmod \pi$ for some δ, where γ is the inverse of π modulo q. This is similar to the trace of a power of a primitive element representation of LFSR sequences.

As in the case of FCSRs and LFSRs, one can ask whether there is an algorithm which, given part of a binary sequence A, synthesizes a (minimal length)

AFSR that generates A. In the case of FCSRs, it was shown that the existence of such an algorithm implies that it is possible to crack Massey and Ruepell's summation combiner [11,13]. It was further argued that the *2-adic span*, the length of the smallest FCSR that generates a given sequence, is thus an important measure of security. A sequence must have large 2-adic span in order to be secure (though this of course does not guarantee security). In this paper we discuss two approaches to generalizing this attack to AFSRs. One generalizes the 2-adic rational approximation algorithm presented previously. This generalization only works for registers defined over rings with particularly nice structure (Euclidean domains). The second approach involves considering an AFSR sequence over R as an interleaving of sequences over a subring and using a rational approximation algorithm over this smaller ring. In general, however, this approach does not find the minimal size AFSR over R.

Many of the results in this paper parallel the theory of FCSRs as developed by the first author and Mark Goresky. We have endeavored to point out where these parallels occur.

2 Algebraic Background

In this section we recall the basics of algebra over completions of rings. We assume a basic knowledge of the theory of rings and fields [1,5,6]. To make the ideas clearer, we describe three examples in parenthetical comments throughout this section. A summary of the 2-adic numbers can be found in [13].

Let R be a commutative ring which is an integral domain (no zero divisors). Let F be its field of fractions. Let $\pi \in R$ be a prime element. The principal ideal generated by π is denoted $I = (\pi)$. (Example 1: given a finite field L, $R = L[x]$, the polynomial ring in one variable over L; $\pi = x$; $I = \{f(x) : f(0) = 0\}$; $F = L((x))$, the field of Laurent series. Example 2: $R = \mathbf{Z}$, the integers; $\pi = p$, a prime integer; $I = \{n : p|n\}$; $F = \mathbf{Q}$, the rational numbers. Example 3: Let $\pi^2 + 2\pi = 2$. $R = \mathbf{Z}[\pi] = \mathbf{Z} + \pi\mathbf{Z}$; $I = \pi\mathbf{Z} + 2\mathbf{Z}$; $F = \mathbf{Q}[\pi] = \mathbf{Q}[\sqrt{3}]$, a quadratic number field. Note that R is a *Euclidean domain* in this case.)

We are principally interested in the case when the quotient $K = R/(\pi)$ is finite. In this case K is a field called the *residue field of* (R, π). More generally, K is a field if (π) is a maximal ideal, and we assume this throughout. (Example 1: $K = L$. Example 2: $K = \mathbf{Z}/(p) = F_p$, the finite field with p elements. Example 3: $K = (\mathbf{Z} + \pi\mathbf{Z})/(\pi) = \mathbf{Z}/(2) = \{0, 1\}$.)

Any such π defines a topology on R with respect to which the operations of addition and multiplication are continuous. The set $\{(\pi^i)\}$ forms a basic set of neighborhoods of zero. This topology is known as the π-adic topology on R

and extends to F with the same basic set of neighborhoods of zero. (Example 1: Two polynomials $f(x) = \sum a_i x^i$ and $g(x) = \sum b_i x^i$ are close in the x-adic topology if $a_i = b_i$ for all but large values of i. Example 2: Integers f and g are close if they are congruent modulo a large power of p. Example 3: $f_0 + \pi f_1$ and $g_0 + \pi g_1$ are close if f_0 is congruent to g_0 and f_1 is congruent to g_1 modulo a large power of 2.)

A *completion* of the π-adic topology on R is a topological ring $\hat{R}$ containing R that is complete (every Cauchy sequence converges) and is a minimal completion containing R. The same notion of completion applies to F.

The set of power series

$$\sum_{i=0}^{\infty} a_i \pi^i, \quad a_i \in R, \tag{1}$$

is a completion of R with the π-adic topology if $\cap_n (\pi)^n = (0)$ (e.g. if R is Noetherian). Two such power series $\sum a_i \pi^i$ and $\sum b_i \pi^i$ are identified if for every n,

$$\sum_{i=0}^{n-1} (a_i - b_i)\pi^i \in (\pi)^n.$$

Addition and multiplication can be defined naturally. The resulting ring is called *the* completion of R and is denoted by $\hat{R}$. The ring $\hat{R}$ has a unique prime ideal $\hat{I}$, the set of such power series with $a_0 = 0$. We have $(\pi) = \hat{I} \cap R$.

It is often convenient to have a standard representation for $\hat{R}$. Let S be a set of elements which is mapped one-to-one and onto the residue field K by reduction modulo π. Such a set is called a *complete set of residues*. (More generally, if J is any ideal in R, then a *complete set of residues modulo* J is a subset of R that maps one-to-one and onto R/J.) It can be shown that every element of $\hat{R}$ can be written uniquely in the form of equation (1) with every a_i in S. A critical observation here is that this representation identifies an element of $\hat{R}$ with a sequence of elements of S. This in turn can be identified with a sequence of elements of K by reduction modulo π. (Example 1: We can let the complete set of residues be L. Then the x-adic completion of $L[x]$ is $L[[x]]$, the set of power series in x over L. Addition is term by term addition. Multiplication is defined by: if $f(x) = \sum a_i x^i$ and $g(x) = \sum b_i x^i$, then

$$f(x)g(x) = \sum_{i=0}^{\infty} (\sum_{j=0}^{i} a_j b_{i-j}) x^i.$$

Example 2: The p-adic completion of R is the set of so-called p-adic numbers

$\mathbf{Z}_p$. We can let the complete set of residues be $\{0, 1, \cdots, p-1\}$. Then $\mathbf{Z}_p$ is the set of expressions of the form

$$\sum_{i=0}^{\infty} a_i p^i,$$

with $a_i \in \{0, 1, \cdots, p-1\}$. Addition and multiplication are with carry. Thus for example if $p = 3$, then

$$\begin{aligned}
&(1 + 0 \cdot 3^1 + 0 \cdot 3^2 + 2 \cdot 3^3 + 3^4 + 2 \cdot 3^5 + \cdots) \\
&+ (2 + 3^1 + 0 \cdot 3^2 + 0 \cdot 3^3 + 2 \cdot 3^4 + 3^5 \cdots) \\
&= (0 + 2 \cdot 3 + 0 \cdot 3^2 + 2 \cdot 3^3 + 0 \cdot 3^4 + 3^5).
\end{aligned}$$

Also, $-1 = (p-1) + (p-1)p + (p-1)p^2 + \cdots$. Example 3: We can let the complete set of residues be $\{0, 1\}$. Thus $\hat{R}$ is the set of expressions of he form

$$\sum_{i=0}^{\infty} a_i \pi^i,$$

with $a_i \in \{0, 1\}$. Note that $2 = \pi^2 + 2\pi$ implies

$$2 = \frac{\pi^2}{1 - \pi} = \pi^2 + \pi^3 + \pi^4 + \cdots.$$

Thus, for example,

$$(1 + \pi + \pi^4 + \cdots) + (1 + \pi^3 + \pi^4) = \pi + \pi^2 + \pi^4 + \cdots$$

while

$$(1 + \pi + \pi^4 + \cdots) \cdot (1 + \pi^3 + \pi^4) = 1 + \pi + \pi^3 + \pi^4 + \cdots.)$$

Any element of $\hat{R}$ with a_0 not divisible by π is invertible in $\hat{R}$. Hence any element of $R - (\pi)$ is invertible in $\hat{R}$. It also follows that the field of fractions $\hat{F}$ of $\hat{R}$ can be identified with the set of Laurent series

$$\sum_{i=t}^{\infty} a_i \pi^i \tag{2}$$

with $t \in \mathbf{Z}$ and $a_i \in S$.

The following result, well known to number theorists [1, p. 100, Lemma 1], is used later to find conditions under which the memory of an AFSR is finite. Let $||(x_1, \cdots, x_k)|| = (\sum_i x_i^2)^{1/2}$ be the Euclidean norm on $\mathbf{R}^k$.

Theorem 1 *If $L \subseteq \mathbf{R}^k$ is an integer lattice of rank at most k, and U is a subset of L contained in $\{x : ||x|| < c\}$, then U is finite.*

3 Definitions

The ingredients we use to define algebraic feedback shift registers are as follows:

(1) A domain R with fraction field F, principal maximal prime ideal I generated by an element π, and finite residue field $K = R/I$.
(2) A pair of complete sets of residues $S, T \subseteq R$.

There is a well defined notion of the reduction of an element $\alpha \in \hat{R}$ modulo π relative to a particular complete set of residues. If the expansion of α is

$$\alpha = \sum_{i=0}^{\infty} a_i \pi^i,$$

then the *reduction of α modulo π* is a_0. We also refer to

$$\sum_{i=0}^{\infty} a_{i+1} \pi^i$$

as the *integral quotient* of α by π, denoted $\mathrm{quo}(\alpha, \pi)$. Thus in general

$$\alpha = (\alpha \bmod \pi) + \pi \mathrm{quo}(\alpha, \pi).$$

Note that if $\alpha \in R$, then $\mathrm{quo}(\alpha, \pi) \in R$.

Linear feedback shift registers can be interpreted as outputting the power series (or x-adic) expansion of a rational function $u(x)/q(x)$. Generalizing this, we want a class of registers that outputs the π-adic expansion with coefficients in S of every R-rational element u/q of $\hat{R}$. The structure of the register should depend on the π-adic expansion of q with coefficients in T. A similar construction was used by Klapper and Goresky to define FCSRs [13, Definition 3.1, p. 118].

Definition 2 *An* algebraic feedback shift register *(or AFSR) over (R, π, S, T) of length r is specified by $r+1$ elements $q_0, q_1, \cdots, q_r \in T$ called the* taps*, with*

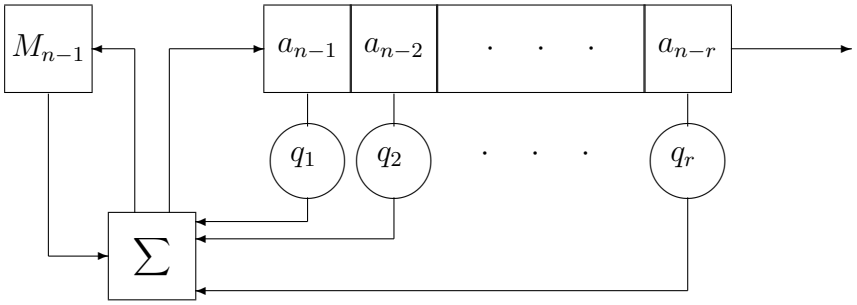

Fig. 1. Diagram of an AFSR after $n - r$ iterations, $n \geq r$.

$q_0 \not\equiv 0 \bmod \pi$. *It is an automaton each of whose states consists of* r *elements* $a_0, a_1, \cdots, a_{r-1} \in S$ *and an element* $m \in R$. *The state is updated by the following steps.*

(1) Compute

$$\tau = \sum_{i=1}^{r} q_i a_{r-i} + m.$$

(2) Find $a_r \in S$ *such that* $q_0 a_r \equiv \tau \bmod \pi$.
(3) Replace $(a_0, \cdots, a_{r-1})$ *by* $(a_1, \cdots, a_r)$ *and replace* m *by* $\mathrm{quo}(\tau - q_0 a_r, \pi)$.

Note that a_r in step (2) can be computed efficiently by reducing q_0 and τ modulo π, dividing in K, and lifting the result to S.

The element $-q_0 + \sum_{i=1}^{r} q_i \pi^i$ plays a central role in the analysis of AFSRs and is referred to as the *connection element*. A diagram of an AFSR is given in Figure 1.

Such a register outputs an infinite sequence over S. By reduction modulo π, this can be identified with an infinite sequence A over K. On the other hand, it can be identified with a power series α in π with coefficients in S, i.e., an element of $\hat{R}$. We generally reserve upper case letters near the beginning of the alphabet for sequences over K and Greek letters near the beginning of the alphabet for elements of $\hat{R}$. At times we move freely between these representations.

Treating the output of an AFSR as a sequence over K, one may ask what effect the choice of the complete sets of residues S and T has. There are several ways to ask this. First, consider the choice of T. We might fix a particular choice of the reductions modulo π of the coefficients q_i, then ask how the choice of T affects the output. Note that the connection element q will depend on the choice of T. It will follow from Theorem 3 that even the period of the output is strongly affected by the choice of T in this case.

Alternatively, we might choose a particular connection element q and construct an AFSR with that connection element. The structure of the resulting AFSR will be strongly affected by the choice of T. Even the length of the AFSR may be affected. However, it also will follow from Theorem 3 that the output is largely unaffected by the choice of T.

Finally, we can consider the effect of the choice of S. In Subsection 7.4 we see that the choice of S can have a strong effect on the output. In particular, it can even affect the period of the output.

We can realize LFSRs over a field K by this construction as follows. We let $(R, \pi, S, T) = (K[x], x, K, K)$. If we initialize the memory of an AFSR in this setting to zero, then it remains zero throughout the infinite execution (there is no carry when multiplying and adding elements of K). The connection element is just the classical connection polynomial which has been used widely in the analysis of LFSRs [3].

We can also realize FCSRs by letting $(R, \pi, S, T) = (\mathbf{Z}, 2, \{0, 1\}, \{0, 1\})$. In this case, if the carry m starts out as a finite sum of powers of π with coefficients in S, then this remains true forever [13].

Two other special cases have been considered. The case where $R = \mathbf{Z}[p^{1/d}]$ was described in [4]. The case where R is the ring of integers in a number field and π is unramified over $\mathbf{Z}$ was considered in [8].

4 Properties of AFSRs

Throughout this section we assume R is a ring, $\pi \in R$ is prime, and S and T are complete sets of residues modulo π. We show that the sequences that are the outputs of AFSRs over (R, π, S, T) are precisely the coefficient sequences of elements of $\hat{R}$ of the form u/v with $u, v \in R$ and π not dividing v.

Theorem 3 *(Generalizes [13, Theorem 4.2, p. 121]) The output, A, of an AFSR with connection element q, initial memory value m_{r-1}, and initial loading $a_0, a_1, \cdots, a_{r-1}$, is the coefficient sequence of the π-adic representation of an element of F*

$$\alpha = \frac{\sum_{n=0}^{r-1} \sum_{i=1}^{n} q_i a_{n-i} \pi^n - q_0 \sum_{n=0}^{r} a_n \pi^n - m_{r-1} \pi^r}{q} = \frac{u}{q}.$$

PROOF. Let

$$\alpha = \sum_{i=0}^{\infty} a_i \pi^i, \tag{3}$$

with $a_i \in S$. Let us consider the transition from one state of the shift register to the next. Suppose that, for some given state, the value of the memory is m_{n-1} and that the state of the register is given by the r elements $a_{n-r}, a_{n-r+1}, \cdots, a_{n-1} \in S$, with a_{n-r} the leftmost and a_{n-1} the rightmost, and where the register shifts towards the left. The next state is determined by calculating

$$\begin{aligned} \tau_n &= m_{n-1} + \sum_{i=1}^{r} q_i a_{n-i} \\ q_0 a_n &= \tau_n \bmod \pi, \end{aligned}$$

writing the new memory contents as

$$m_n = \mathrm{quo}(\tau_n - q_0 a_n, \pi),$$

and using a_n as the new contents of the rightmost cell. (The remaining terms are shifted once to the left.) These equations may be combined into the expression

$$\tau_n = \pi m_n + q_0 a_n,$$

with $a_n \in S$. It follows that

$$q_0 a_n = \sum_{i=1}^{r} q_i a_{n-i} + (m_{n-1} - \pi m_n), \tag{4}$$

provided that $n \geq r$. Suppose the initial loading of the register consists of memory m_{r-1} and with register values $a_0, a_1, \cdots, a_{r-1}$. Now substitute (4) into the expression (3) for α to obtain

$$\begin{aligned} q_0 \alpha &= q_0 (a_0 + a_1 \pi + \cdots + a_{r-1} \pi^{r-1} + \sum_{n=r}^{\infty} a_n \pi^n) \\ &= q_0 x + \sum_{n=r}^{\infty} \left(\sum_{i=1}^{r} q_i a_{n-i} \right) \pi^n + \sum_{n=r}^{\infty} (m_{n-1} - \pi m_n) \pi^n, \end{aligned} \tag{5}$$

where

$$x = a_0 + a_1 \pi + \cdots + a_{r-1} \pi^{r-1}$$

is the element represented by the initial loading of the register. The second summation in equation (5) cancels except for the first term, m_{r-1}, leaving

$$\begin{aligned} q_0\alpha &= q_0x + m_{r-1}\pi^r + \sum_{n=r}^{\infty}\sum_{i=1}^{r} q_i\pi^i a_{n-i}\pi^{n-i} \\ &= q_0x + m_{r-1}\pi^r + \sum_{i=1}^{r} q_i\pi^i \left(\sum_{n=r}^{\infty} a_{n-i}\pi^{n-i}\right) \\ &= q_0x + m_{r-1}\pi^r + \sum_{i=1}^{r} q_i\pi^i(\alpha - (a_0\pi^0 + a_1\pi^1 + \cdots + a_{r-i-1}\pi^{r-i-1})) \\ &= q_0x + m_{r-1}\pi^r + \alpha\sum_{i=1}^{r} q_i\pi^i - \sum_{i=1}^{r-1}\sum_{j=0}^{r-i-1} q_i\pi^i a_j\pi^j \end{aligned}$$

(where the inner sum is empty, hence zero, when $i = r$ in the third line). These equations give

$$\begin{aligned} \alpha &= \frac{q_0x + m_{r-1}\pi^r - \sum_{i=1}^{r-1}\sum_{j=0}^{r-i-1} q_i\pi^i a_j\pi^j}{q_0 - \sum_{i=1}^{r} q_i\pi^i} \\ &= \frac{\sum_{n=0}^{r-1}(\sum_{i=1}^{n} q_i a_{n-i})\pi^n - q_0\sum_{n=0}^{r-1} a_n\pi^n - m_{r-1}\pi^r}{q}. \end{aligned} \tag{6}$$

□

Thus the denominator of α is equal to the connection element q of the shift register.

Corollary 4 *(Generalizes [13, Corollary 4.3, p. 122]) Adding b to the memory adds*

$$\frac{-b\pi^r}{q}$$

to the output.

Corollary 5 *For any $u, q \in R$, with $q \not\equiv 0 \bmod \pi$, there is at most one AFSR over R, π, and S with connection element q, whose output corresponds to u/q.*

PROOF. Suppose that both $m, (a_0, \cdots, a_{r-1})$ and $m', (a'_0, \cdots, a'_{r-1})$ give rise to the sequence corresponding to u/q. Then $a_0, \cdots, a_{r-1}$ and $a'_0 \cdots, a'_{r-1}$ are the first r elements of this sequence. Hence $a_i = a'_i$, $i = 0, \cdots r-1$. It follows from Theorem 3 that

$$\frac{(m - m')\pi^r}{q} = 0.$$

Hence $m = m'$. $\square$

The converse of Theorem 3, that an element u/q of F can be realized as the output of an AFSR, is true as well. To see this, we show how to construct the initial loading of an AFSR for certain u, and then use Corollary 4 to obtain initial loadings for other AFSRs. Let

$$u = \sum_{i=0}^{r-1} u_i \pi^i$$

with $u_i \in S$. Every element of R differs from some such element u by a multiple of π^r. This follows directly from the fact that every element of $\hat{R}$ can be written as a power series in π with coefficients in S. Thus if we can construct initial loadings for u/q with u of this type, then we can construct initial loadings for all u/q.

Theorem 6 *(Generalizes [13, Section 5, p. 123]) Given a connection element*

$$q = -q_0 + \sum_{i=1}^{r} q_i \pi^i \tag{7}$$

with $q_0, \cdots, q_r \in T$, *and*

$$u = \sum_{i=0}^{r-1} u_i \pi^i$$

with $u_i \in S$, *define* $a_0, \cdots, a_{r-1}$ *and* m_{r-1} *by the following procedure:*

1. *Set* $m_{-1} = 0$ *and* $\sigma_0 = 0$.
2. *For each* $i = 0, 1, \ldots, r-1$ *compute the following elements:*

$$\tau_i = \sum_{k=0}^{i-1} q_{i-k} a_k + m_{i-1} - u_i \in R.$$

The empty sum in τ_0 *is interpreted as zero.*
3. *Find* $a_i \in S$ *and* $m_i \in R$ *such that*

$$\tau_i = q_0 a_i + \pi m_i.$$

If $(a_0, a_1, \ldots, a_{r-1})$ *is used as the initial loading and* m_{r-1} *is used as the initial memory in an AFSR with connection element* q, *then the output sequence will correspond to the element* $u/q \in F$.

Note that for some choices of T not every element q can be written in the form in equation (7).

5 Finite Memory

In order to implement an AFSR it is necessary that the memory remain bounded throughout an infinite execution. There are two quite general types of ring for which we have been able to determine when this happens. The first case is when the field of fractions F is a number field (a finite extension of the rational numbers). In this case we can use well known results from number theory to determine those R for which the memory always remains bounded. The second case is when R is a polynomial ring over a finite field – a function field. In this case there is no carry from lower degree terms to higher degree terms when addition and multiplication are carried out, so the degree of the memory always remains bounded.

5.1 The Number Field Case

We first assume that the fraction field F is a number field. Such a field can be embedded in the complex numbers. In general there are several embeddings of F in the real numbers (real embeddings), and several that are not in the real numbers (complex embeddings). The complex embeddings always occur in conjugate pairs. If we let r_1 denote the number of real embeddings, and $2r_2$ denote the number of complex embeddings, then $r_1 + 2r_2 = [F : \mathbf{Q}]$, the degree of the extension $F/\mathbf{Q}$ [1, p. 95].

Having fixed an embedding, we denote by $|x|$ the complex norm of a complex number x. If m is the memory of an AFSR, we want to consider the growth of $|m|$ over an infinite execution. Suppose $a_0, a_1, \cdots$ is the output sequence, and m_n is the memory at the nth state (i.e., when the register contains $(a_{n+r-1}, \cdots, a_n)$). Then

$$\pi m_{n+1} + q_0 a_{n+r} = m_n + \sum_{i=1}^{r} q_i a_{n+r-i}.$$

It follows that

$$\begin{aligned} |m_{n+1}| &\geq \frac{|m_n| - \sum_{i=0}^{r} |q_i a_{n+r-i}|}{|\pi|} \\ &\geq \frac{|m_n| - (r+1)BC}{|\pi|} \end{aligned}$$

where $B = \max\{|t| : t \in T\}$ and $C = \max\{|s| : s \in S\}$. Suppose $|\pi| < 1$.

Then for

$$|m_n| \geq \frac{(r+1)BC+1}{1-|\pi|}$$

we have $|m_{n+1}| > |m_n| + 1$. Thus the memory increases unboundedly, and in particular takes infinitely many values in an infinite execution. We have shown the following.

Proposition 7 *If there is an embedding of F in the complex numbers such that $|\pi| < 1$, then there is an AFSR whose memory grows unboundedly from some initial state.*

Now suppose that for a given embedding of F we have $|\pi| > 1$. By similar reasoning we see that

$$|m_{n+1}| < \frac{|m_n| + (r+1)BC}{|\pi|}.$$

If

$$|m_n| \leq \frac{(r+1)BC}{|\pi|-1}, \tag{8}$$

then the same inequality holds for $|m_{n+1}|$. If equation (8) does not hold, then $|m_{n+1}| < |m_n|$. In either case, the complex norm of the memory is bounded throughout the infinite execution of the AFSR. To guarantee that it only takes on finitely many values, however, we need a stronger condition.

Proposition 8 *If for every embedding of F in the complex numbers we have $|\pi| > 1$, then the memory in the infinite execution of any AFSR over F takes on only finitely many values. The output is therefore eventually periodic.*

PROOF. Let $k = r_1 + 2r_2$. Suppose $\sigma_1, \cdots, \sigma_{r_1+r_2}$ is a set of embeddings of F in the complex numbers that includes all the real embeddings and one complex embedding from each conjugate pair. Consider the map $\psi : F \to \mathbf{R}^k$ defined by

$$\psi(x) = (\sigma_1(x), \cdots, \sigma_{r_1+r_2}(x)).$$

The image of R under ψ is an integer lattice of rank k, and ψ is injective [1, p. 95-99]. By Theorem 1, any set of points in $\psi(R)$ is finite if it is bounded in Euclidean norm.

Let U be the image under ψ of the set of memory values in one infinite of an AFSR. By the preceding argument, for each i, we have that the set of $|\sigma_i(m)|$ is bounded. It follows that $||(\sigma_1(m), \cdots, \sigma_k(m))||$ is bounded. The proposition follows. $\Box$

5.2 The Function Field Case

Suppose R is a polynomial ring over K. Then the degree of the memory is always bounded.

Proposition 9 *Let $U = \max\{deg(u) : u \in S\}$, $V = \max\{deg(u) : u \in T\}$. Suppose that at some state the AFSR has memory m, and let m' be the memory at the next state. Then*

$$\deg(m') \leq \max(U + V, \deg(m)) - d.$$

PROOF. If the state of the AFSR is $(a_0, \cdots, a_{r-1})$, $a_i \in S$, then we have

$$\sigma = \sum_{i=0}^{r-1} a_i q_{r-i} + m = m'\pi + a_r$$

with $q_i \in T$. Thus

$$\deg(m') + d \leq \max(U + V, \deg(m), \deg(a_r)).$$

The proposition follows. $\Box$

It follows that the memory eventually has degree at most $U+V-d$. Since there are finitely many states with this property, the output is eventually periodic. Also, any strictly periodic sequence can be generated by an AFSR where the degree of the memory is bounded by $U+V-d$ throughout its execution. Note that this bound is independent of the length of the AFSR.

6 Exponential Representation and Period of AFSR Sequences

One of the most powerful techniques for the analysis of a shift register sequence is its exponential representation. Suppose $A = a_0, a_1, a_2, \cdots$ is a periodic se-

quence over $K = GF(p^n)$ obtained from a LFSR of length r, with connection polynomial $q(X)$. If $q(X)$ is irreducible and if

$$\gamma \in GF(p^{nr})$$

is a root of $q(X)$ in the finite field with p^{nr} elements, then for $i = 0, 1, 2, \cdots$ we have

$$a_i = Tr(c\gamma^i)$$

for some $c \in GF(p^{nr})$ (which corresponds to the choice of initial loading of the shift register). Here,

$$Tr : GF(p^{nr}) \rightarrow GF(p^n)$$

denotes the trace function. In this section we derive a similar representation for periodic sequences obtained from an AFSR.

We concern ourselves only with strictly periodic sequences $A = a_0, a_1, a_2, \cdots$ that are generated by AFSRs with a given connection element q. We have seen that the element of $\hat{R}$ associated with such a sequence is in F, and can be written in the form u/q, with $u \in R$.

Theorem 10 *(Generalizes [13, Theorem 6.1, p. 125]) Let* $q = -q_0 + \sum_{i=1}^{r} q_i \pi^i$, $q_i \in T$, $q_0 \not\equiv 0 \bmod \pi$. *Let* V_q *be the set of elements* $u \in R$ *such that* u/q *has a strictly periodic expansion* $u/q = \sum_{i=0}^{\infty} a_i \pi^i$ *with* $a_i \in S$ *and* u *and* q *relatively prime. Suppose that no two elements of* V_q *are congruent modulo* π, *and let* U_q *be a complete set of residues modulo* q *that contains* V_q. *Let* $A = a_0, a_1, a_2, \cdots$ *be a periodic sequence generated by an AFSR with connection element* q. *Suppose* $\alpha = \sum a_i \pi^i = u/q$ *with* u *and* q *relatively prime. Let*

$$\gamma = \pi^{-1} \in R/(q)$$

be the (multiplicative) inverse of π *in the ring* $R/(q)$. *Then for all* $i = 0, 1, 2, \cdots$ *we have,*

$$a_i = \Big(u\gamma^i \ (\bmod q)\Big) \ (\bmod \pi).$$

Here the notation (mod q)(mod π) means that first the element $\delta\gamma^i$ should be reduced modulo q to give an element of U_q, and then that element should be reduced modulo π to give an element of K.

PROOF. Suppose the AFSR is in a state X, meaning the memory has some value m and the register is loaded with $a_0, a_1, \cdots, a_{r-1}$. Let

$$T = \mathrm{ord}_q(\pi)$$

denote the period of this sequence (which may be less than the order of the multiplicative group of $R/(q)$). Let $\alpha(X) \in \hat{R}$ denote the element associated with the output sequence from the state X. By Theorem 3, $\alpha(X)$ is an element of F of the form

$$\alpha(X) = \frac{u}{q} = \sum_{i=0}^{\infty} a_i \pi^i,$$

with $0 \leq p \leq q-1$. Now let Y denote the next state of the FCSR, so

$$\alpha(Y) = \frac{v}{q} = \sum_{i=0}^{\infty} a_{i+1} \pi^i.$$

Thus, $u, v \in U_q$ and

$$\pi \frac{v}{q} + a_0 = \frac{u}{q},$$

or $u = \pi v + a_0 q \in R$. If we read this equation modulo π, we see

$$u \equiv a_0 \ (\mathrm{mod}\ \pi).$$

Reading this equation modulo q we obtain

$$v \equiv \gamma u (\mathrm{mod}\ q).$$

This shows that the sequence of numerators $(u, v, \cdots)$ is obtained by multiplying by γ and reducing mod q, and that the sequence $(a_0, a_1, \cdots)$ over K is obtained by reducing the numerators modulo π. Finally, the initial state is arbitrary and given by the choice of some $A \in R/(q)$. $\square$

Corollary 11 *Under the hypotheses of Theorem 10, the period of A is the order of π modulo q.*

PROOF. The period of A equals the period of the sequence of numerators $(u, v, \cdots)$ as in the proof of Theorem 10. The ith element in the sequence of numerators is $u\gamma^i \bmod q$. The period of this sequence is the least i such that

$u \equiv u\gamma^i \bmod q$. Since u is relatively prime to q, this period is exactly the order of γ modulo q, which is the same as the order of π modulo q. $\square$

The hypotheses of Theorem 10 do not always hold. For example, suppose $\pi^2 = 2$, $R = \mathbf{Z}[\pi]$, and $S = \{0, 1\}$. Then the periodic sequence $11101110\cdots$ has the corresponding π-adic number

$$\frac{u}{q} = -\frac{1 + \pi + \pi^2}{\pi^4 - 1} = -\frac{3 + \pi}{3}$$

while the periodic sequence $01000100\cdots$ has the corresponding π-adic number

$$\frac{v}{q} = -\frac{\pi}{\pi^4 - 1} = -\frac{\pi}{3} = \frac{u}{q} + 1.$$

That is, we have $q = 3$ and $u \equiv v \bmod q$ both give rise to periodic sequences. The congruence class modulo q does not uniquely determine a periodic sequence.

Suppose the hypotheses of Theorem 10 indeed do not hold. We can still give a bound on the period. For any $u \in R$, the *coset of u modulo q* is the finite set $\{u\pi^i\}$.

Proposition 12 *Suppose $A = a_0, a_1, a_2, \cdots$ is a sequence generated by an AFSR with connection element q and associated element*

$$\alpha = \sum a_i \pi^i = \frac{u}{q}.$$

Then the eventual period of A is a multiple of the order of the coset of u modulo q.

PROOF. We can write $\alpha = b + \pi^k \beta$ with $b = \sum_{i=0}^{k-1} a_i \pi^i \in R$ and the coefficient sequence of β strictly periodic with period t equal to the eventual period of A. Then β can also be generated by an AFSR with connection element q, so $\beta = u'/q$ for some $u' \in R$. It follows that $u = qb + \pi^k u'$, so that u and u' have the same coset modulo q. Thus we may assume α is strictly periodic with period t.

We can write

$$\alpha = \frac{u}{q} = \frac{v}{\pi^t - 1},$$

for some $u, v \in R$. Thus $u(\pi^t - 1) = vq$, and it follows immediately that t is a multiple of the order of the coset of u. $\square$

In particular, if u and q are relatively prime, then the period is a multiple of the order of π modulo q.

Now suppose we are given u/q and are free to choose S. How close can we come to the bound given in Proposition 12?

Proposition 13 *Let u/q be an R-rational element, with q relatively prime to π. There is a complete set of representatives S modulo π such that u/q has a strictly periodic π-adic expansion with coefficients in S and period equal to the order t of the coset of u modulo q.*

PROOF. We have $u - \pi^t u = bq$ for some $b \in R$. Let $b = \pi^k c$, with c not divisible by π. If $t = 1$, let S contain b and enough other elements to make a complete set of representatives modulo π.

If $k < t$ and $t > 1$, let S contain 0, c, and enough other elements to make a complete set of representatives modulo π. Then

$$b = 0 + 0 \cdot \pi + \cdots + 0 \cdot \pi^{k-1} + c\pi^k + 0 \cdot \pi^{k+1} + \cdots + 0 \cdot \pi^{t-1}.$$

If $k \geq t > 1$, let S contain π, $v = c\pi^{k-1} - 1 - \pi - \pi^2 - \cdots - \pi^{t-1}$, and enough other elements to make a complete set of representatives modulo π. Then

$$b = \pi + v\pi + \pi \cdot \pi^2 + \cdots + \pi \cdot \pi^{t-1}.$$

In each case we can write

$$u - \pi^t u = \left(\sum_{i=0}^{t-1} a_i \pi^i \right) q$$

with $a_i \in S$. It follows that the π-adic expansion of u/q with coefficients in S is

$$\begin{aligned} \frac{u}{q} &= \frac{\sum_{i=0}^{t-1} a_i \pi^i}{1 - \pi^t} \\ &= a_0 + a_1 \pi + \cdots a_{t-1} \pi^{t-1} + a_0 \pi^t + \cdots, \end{aligned}$$

which is strictly periodic with period t. $\square$

In particular, if u is relatively prime to q, then the period of the π-adic expansion of u/q with coefficients in the set S found in Proposition 13 is precisely

register	mem	i
111	0	0
110	1	1
100	1	2
000	1	3
001	0	4
011	0	5

Table 1
The states of an AFSR with $R = \mathbf{Z}$, $p = \pi = 2$, and $q = 9$.

the order of π modulo q. Examples where this fails (and the hypotheses of Theorem 10 fail) are given in Subsection 7.4.

7 Examples

In this section we illustrate the behavior of AFSRs by several examples.

7.1 $R = \mathbf{Z}$, $p = \pi = 2$

Suppose that $R = \mathbf{Z}$ so $F = \mathbf{Q}$. Let $\pi = p = 2$, so $K = GF(2)$, and $S = T = \{0, 1\}$. This is precisely the setting that gives rise to FCSRs [13]. Suppose

$$q = \pi^2 + \pi - 1 = 9.$$

Then an AFSR with connection element q has three stages, with coefficients 1, 0, and 1. If we start the register in the initial state $(1, 1, 1)$, and with initial memory 0, then the sequence of states of the register is given in Table 1 (where we shift toward the left at each state transition). The output sequence thus has period 6, and one period is

$$A = 111000 \cdots.$$

Note that the memory size never exceeds one bit, so in effect we have a four stage binary feedback register with period 6. Also note that

$$\mathrm{ord}_q(2) = 6,$$

and $2^{-1} \bmod q = 5$. The exponential representation of this sequence is

$$a_i = (5^i \bmod 9) \bmod 2.$$

Finally, since the period is 6 and one period gives the binary representation of 7, the rational representation of the sequence is

$$\frac{7}{1-2^6} = \frac{7}{-63} = \frac{-1}{9}.$$

7.2 $R = \mathbf{Z}[\pi]$ *with* $\pi = 2^{1/2}$

Suppose that $R = \mathbf{Z}[\pi]$ with $\pi^2 = p = 2$, so $F = \mathbf{Q}[\pi]$ is a real quadratic number field and $K = GF(2)$. Let $S = T = \{0, 1\}$. This is an example of what was previously called a d-FCSR, $d = 2$ [4]. Every element in R can be written in the form $a + b\pi$ with a and b integers. Let

$$q = \pi^3 + \pi - 1 = 3\pi - 1.$$

Then an AFSR with connection element q again has three stages, with coefficients 1, 0, and 1. If we start the register in the initial state $(1, 1, 1)$, and with initial memory 0, then the sequence of states of the register is given in Table 2 (where we shift toward the left at each state transition). The output sequence thus has period 16, and one period is

$$A = 1110111100010000 \cdots.$$

Note that the memory size never exceeds two bits, so in effect we have a five stage binary feedback register with period 16. Also note that

$$\pi^{-1} \bmod q = 3$$

and

$$\mathrm{ord}_q(\pi) = 16.$$

One way to see this is to note that if a and b are integers then

$$a + b\pi \equiv (3a + b)\pi \bmod q.$$

Moreover, $17\pi = (3\pi - 1)(\pi + 6)$. Since 17 is prime, 17π must be the smallest integral multiple of π that is congruent to 0 modulo q. It follows that $R/(q)$

register	mem	i
111	0	0
110	π	1
101	1	2
011	π	3
111	1	4
111	π	5
110	$\pi+1$	6
100	$\pi+1$	7
000	$\pi+1$	8
001	1	9
010	π	10
100	1	11
000	π	12
000	1	13
001	0	14
011	0	15

Table 2
The states of an AFSR with $R = \mathbf{Z}[\pi]$, $\pi = 2^{1/2}$ and $q = 3\pi - 1$.

is isomorphic to the integers modulo 17, and every element has multiplicative order dividing 16. Then one checks that $\pi^8 = 16$ is not congruent to one modulo q.

The exponential representation of this sequence is thus

$$a_i = (3^i \bmod (3\pi - 1)) \bmod 2.$$

Finally, since the period is 16 and one period gives the π-adic representation of $15 + 45\pi$, the rational representation of the sequence is

$$\begin{aligned} \frac{15+45\pi}{1-\pi^{16}} &= \frac{15+45\pi}{-255} \\ &= \frac{-1}{-1+3\pi}. \end{aligned}$$

7.3 $R = \mathbf{Z}[\pi, \gamma]$ with $\pi^2 = 2$ and, $\gamma^2 = \gamma + 1$

Suppose that $R = \mathbf{Z}[\pi, \gamma]$ with $\gamma^2 = \gamma + 1$ and $\pi^2 = p = 2$. Here $F = \mathbf{Q}[\pi, \gamma]$ is a degree 4 extension of $\mathbf{Q}$. Also, γ reduces modulo π to a primitive cube root of 1, so $K = GF(4)$. Let $S = T = \{0, 1, \gamma, 1 + \gamma\}$. Every element in R can be written in the form $(a + b\gamma) + (c + d\gamma)\pi$ with a, b, c, and d integers. Let

$$q = (\gamma + 1)\pi^3 + \pi - 1 = (2\gamma + 3)\pi - 1.$$

Then an AFSR with connection element q again has three stages, with coefficients $1 + \gamma$, 0, and 1. If we start the register in the initial state $(1, 1, 1)$, and with initial memory $(1 + 2\gamma)\pi$, then the output sequence has period 400. Each output symbol has 2 bits, so this register outputs 800 bits. Furthermore, it can be shown that each integer in the memory never exceeds 3. Hence this register is, in effect, a 14 stage binary feedback register with period 800. The first few states are shown in Table 3 (where we shift toward the left at each state transition). Each output symbol is of the form $a + b\gamma$, with $a, b \in \{0, 1\}$. In one period, the sequence of as is

11100001101001000100010111110010000010100011000101100110110
00001010110100111111001111011011000100110111000000000010000001
11101111000001110001010101100101110110101111110100011010101 1
11001100111001101011000111100101101110111010000011011111010 1
11001110100110010010111101010010110000001100001001001110110 0
10001111111111011111000010000111110001110101010011010001001 01
00000010111001010100001100110001100101 00.

The sequence of bs is

00010111111001000001010001100010110011011010000101011010011 11
11001111011011000100110111000000000100000111101111000001110 0
01010101100101110110101111110100011010101111001100111001101 0
11000111100101101110111010000011011111010111001110100110010 0
10111101010010110000001100001001001110110010001111111110111 1
10000100001111100011101010100110100010010100000010111001010 1
00001100110001100101001110000110100100 01.

register			mem	i
1	1	1	$(1+2\gamma)\pi$	0
1	1	γ	$(1+2\gamma)+\pi$	1
1	γ	0	$1+(1+2\gamma)\pi$	2
γ	0	γ	$(1+2\gamma)+\pi$	3
0	γ	γ	$1+(1+2\gamma)\pi$	4
γ	γ	$1+\gamma$	$(1+2\gamma)$	5
γ	$1+\gamma$	$1+\gamma$	$(1+2\gamma)\pi$	6
$1+\gamma$	$1+\gamma$	γ	$(1+2\gamma)+(1+\gamma)\pi$	7
$1+\gamma$	γ	1	$(1+\gamma)+(1+3\gamma)\pi$	8
γ	1	0	$(1+3\gamma)+(2+2\gamma)\pi$	9
1	0	γ	$(2+2\gamma)+(1+2\gamma)\pi$	10
0	γ	1	$(1+2\gamma)+(1+2\gamma)\pi$	11
γ	1	0	$(1+2\gamma)+(1+\gamma)\pi$	12
1	0	0	$(1+\gamma)+(1+2\gamma)\pi$	13
0	0	0	$(1+2\gamma)+(1+\gamma)\pi$	14
0	0	1	$(1+\gamma)+\gamma\pi$	15

Table 3
The first 15 states of an AFSR over $\mathbf{Z}[\pi,\gamma]$ with $\pi^2 = 2$, $\gamma^2 = \gamma + 1$, and $q = (2\gamma+3)\pi - 1$.

Note that

$$\pi^{-1} \bmod q = 2\gamma + 3.$$

It can be shown that

$$\mathrm{ord}_q(\pi) = 400.$$

One way to see this is to note that if a and b are integers then

$$a + b\pi \equiv ((2\gamma+3)a + b)\pi \bmod q.$$

Moreover, $401\pi = N(q)\pi$. Since 401 is prime, 401π must be the smallest integral multiple of π that is congruent to 0 modulo q. It follows that $R/(q)$ is isomorphic to the integers modulo 401, and every element has multiplicative order dividing $400 = 2^4 \cdot 5^2$. Then one checks that neither $\pi^{200} = 2^{100}$ nor $\pi^{80} = 2^{40}$ is congruent to one modulo q.

The exponential representation of this sequence is thus

$$a_i = (3^i \bmod (3\pi - 1)) \bmod 2.$$

Finally, by Theorem 3, the rational representation of the sequence is

$$\alpha = \frac{-\gamma\pi^3 - 1}{(\gamma + 1)\pi^3 + \pi - 1}.$$

7.4 Dependence of the Period on S

It is not the case that for every choice of S the period is the order of the coset of u modulo q. For example, let $R = \mathbf{Z}$ and $\pi = 2$. Let $u = -1$ and $q = 3$. Consider the two complete sets of residues $S_1 = \{0, 1\}$ and $S_2 = \{4, 1\}$. With respect to S_1, $-1/3$ has the coefficient sequence $101\underline{01}0\cdots$, with period 2. With respect to S_2, $-1/3$ has the coefficient sequence $40440004\underline{440004}\cdots$, with (eventual) period 6. Now suppose u and q are arbitrary relatively prime integers. It has been shown that the choice of S_1 as complete set of residues always gives rise to a coefficient sequence with eventual period equal to the order of 2 modulo q [13]. Consider the set of residues $S_3 = \{0, k\}$ for k odd and relatively prime to u. Suppose

$$\frac{u}{q} = \sum_{i=0}^{\infty} a_i 2^i = k \sum_{i=0}^{\infty} b_i 2^i,$$

where $a_i \in S_3$ and $b_i = 1$ if $a_i = 4$ and $b_i = 0$ if $a_i = 0$. If the period of $a_0, a_1, \cdots$ is t, then

$$\frac{u}{qk} = \frac{c}{2^t - 1}$$

for some integer c. Thus t is the least integer such that qk divides $2^t - 1$, so t is the least common multiple of the orders of 2 modulo q and k. In particular, the period can be arbitrarily large (but the requirements for the memory grow as the size of the elements in S_3 grow).

One can also consider a fixed T, q, and initial state of an AFSR with connection element q. We can then vary the set S and ask what effect there is on the output. To make sense of this, we consider the contents of the register and the output as consisting of elements of the residue field K.

For example, let π be a root of the quadratic equation $x^2 - 2x + 2 = 0$. Then the ring $R = \mathbf{Z}[\pi]$ equals the Gaussian domain $\mathbf{Z}[\sqrt{-1}]$ and π is prime.

register	mem	i	register	mem	i
3 3 3	$1-\pi$	0	3 0 3	$1-2\pi$	12
3 3 3	$3-2\pi$	1	3 3 0	$-\pi$	13
3 3 3	$4-3\pi$	2	3 3 3	-1	14
0 3 3	$7-5\pi$	3	3 3 3	$2-\pi$	15
0 0 3	$5-5\pi$	4	0 3 3	$7-4\pi$	16
0 0 0	$3-4\pi$	5	0 0 3	$6-5\pi$	17
3 0 0	-4	6	3 0 0	$1-3\pi$	18
3 3 0	$-4+2\pi$	7	0 3 0	$1-2\pi$	19
3 3 3	$-2+2\pi$	8	3 0 3	$-4+\pi$	20
0 3 3	$6-2\pi$	9	0 3 0	$3-\pi$	21
3 0 3	$4-3\pi$	10	3 0 3	-1	22
0 3 0	$7-5\pi$	11	3 3 0	$2-\pi$	23
			3 3 3	$1-\pi$	24

Table 4
The states of an AFSR over the Gaussian domain.

We have that $R/(\pi) \cong Z/(2)$, and so may choose complete sets of residues $T = S_1 = \{0, 1\}$ and $S_2 = \{0, 3\}$. Consider an AFSR over (R, π, S_1, T) with connection element $q = \pi^3+\pi-1$ (so the length $r = 3$), initial register contents $(a_2, a_1, a_0) = (1, 1, 1)$ and initial memory $m = 1 - \pi$. Since $\pi^2 = 2\pi - 2$ and $2 = \pi(2-\pi)$, we have $\sigma = a_0q_3 + a_1q_2 + a_2q_1 + m = 2 + (1-\pi) = 1 + (2-\pi) = 1+\pi(1-\pi)$. Therefore the feedback element is $a_3 = 1$ and the updated memory is $m = 1-\pi$, hence unchanged. This shows that the output sequence consists of all 1s and its rational representation is $1/(1-\pi) = \pi - 1$.

We now keep everything the same but replace S_1 by S_2. We then have an AFSR over (R, π, S_2, T). In terms of S_2, the initial register contents are $(3, 3, 3)$. The first 25 iterations are displayed in Table 4. The output sequence has period 24. In terms of K, one period is

$$A = 111110001110101111001010\cdots .$$

By Theorem 3, the rational representation for the output is $(\pi^4 - \pi^3 - 3)/q$, which is a reduced rational representation.

8 Rational Approximation and Security Measures

In this section we consider the register synthesis problem for AFSRs when R is a subring of a number field. Given a prefix of a sequence $A = a_0, a_1, \cdots$, we want to find a short AFSR over a given (R, π, S, T) that generates A. As we have seen, the output from such an AFSR corresponds to an element $\alpha = \sum a_i \pi^i = u/q$ for some $u, q \in R$. The elements u and q determine the structure and initial state of the AFSR that generates A. Thus the problem of finding an AFSR that generates A amounts to the problem of finding a rational representation for α. The construction of the AFSR from u and q is a straightforward generalization of the same construction for FCSRs [13, Section 5, p. 123] and is not treated here. In this section we describe conditions on R under which a rational approximation algorithm exists. Note that if u and q are relatively prime, then any other AFSR that outputs S corresponds to a pair uv and qv. The problem of choosing v to minimize the size of the AFSR is also not treated here.

The algorithm we describe here is based on one due to de Weger in the p-adic case [17, p. 77]. Our algorithm is more general in that it works over many number fields. It is also an improvement in two regards. First, it is adaptive. That is, the number of bits of the sequence A need not be predetermined. Second, if the input sequence is, in fact, the coefficient sequence of an R-rational element, then we can bound the number of elements of the sequence that are needed for the algorithm to converge.

A similar modification to de Weger's algorithm was described by Klapper and Goresky in the case of FCSRs with $p = 2$ [13]. The situation here, however, is somewhat more complicated. In Klapper and Goresky's version, the updating was able to be performed with a single iteration of algorithm **Improve**. In our more general setting we must allow several iterations and take into account the number of iterations in the complexity analysis.

We assume R has a *norm function* $N : K - \{0\} \to \mathbf{N}$ ($\mathbf{N}$ denotes the natural numbers) that makes it a Euclidean domain. That is,

a. For all $a, b \in K$, $N(ab) = N(a)N(b)$.
b. For all $a, b \in R$, $N(a + b) \leq N(a) + N(b)$.
c. For all $a, b \in R$, there exist $q, r \in R$ so that $a = qb + r$ and either $r = 0$ or $N(r) < N(b)$.

In addition, in order to ensure the algorithm converges rapidly, we need

d. There is a function $\psi : \mathbf{N} \to \mathbf{N}$ such that if $a \equiv b \bmod \pi^{\psi(k)}$, $N(a) < k$, and $N(b) < k$, then $a = b$.

Conditions (a), (b), and (d) hold, for example, if R is an imaginary quadratic extension of the rationals and N is the square root of the usual norm function on field extensions. Condition (c) holds as well if $R = \mathbf{Q}[\sqrt{D}]$ with $D \in \{-1, -2, -3, -7, -11\}$ [1, Chapter 3, Section 2, p. 164-169].

For any pair of elements u and q of R, define

$$\Phi(u, q) = \max(N(u), N(q)).$$

Assume we have consecutive terms $a_0, a_1, \cdots$ of the sequence A. In the algorithm **Rational_Approximation**, given in Figure 2, the symbols $f = (f_1, f_2)$ and $g = (g_1, g_2)$ denote pairs of elements of R. These two pairs form a basis for the set of rational approximations to $\alpha = \sum_{i=0}^{\infty} a_i \pi^i$ that are accurate modulo π^k.

Algorithm **Improve**, given in Figure 3, is used to find a basis for L_k whose Φ-values are as small as possible.

Theorem 14 *Suppose the output from* **Rational_Approximation** *is* $h = (h_1, h_2)$ *when* T *bits* a_i *are used. Then* $\pi \not| h_2$, $\alpha \cdot h_2 - h_1 \equiv 0 \pmod{\pi^T}$, *and any other pair* $h' = (h'_1, h'_2)$ *which satisfies these two conditions has* $\Phi(h') \geq \Phi(h)$.

Theorem 15 *Suppose* $A = a_0, a_1, a_2, \cdots$ *is an eventually periodic sequence with associated* π*-adic number* $\alpha = \sum a_i \pi^i = u/q$, *with* $u, q \in R$. *If* $T \geq \psi(\Phi(u, q)^2)$ *bits* a_i *are used, then* **Rational_Approximation** *outputs* (h_1, h_2) *with* $h_1/h_2 = u/q$.

For example, consider the case of an imaginary quadratic number field. $F = \mathbf{Q}[\sqrt{D}]$, $D \in \{-1, -2, -3, -7, -11\}$. If $D \equiv 3 \bmod 4$, then $R = \mathbf{Z}[(1 + \sqrt{D})/2]$. Otherwise $R = \mathbf{Z}[\sqrt{D}]$ [1, Chapter 3, Section 2, p. 164-169]. If p is a prime integer, then either p is prime in R or is the product of two primes. In the former case we can take $\psi(x) = \lceil \log_p(x) + c \rceil$ and in the latter case we can take $\psi(x) = \lceil 2 \log_p(x) + c \rceil$ for some constant c. In either case the number of bits required for convergence is linear in $\log(\Phi(u, q))$. The quantity $\log(\Phi(u, q))$ is thus a measure of the cryptographic security of the sequence corresponding to u/q.

The proofs of these two optimality results occupy the remainder of this section, and utilize the methods of [17]. Consider the k^{th} *approximation lattice* for the 2-adic number α,

$$L_k = \{h \in R \times R : \alpha \cdot h_2 - h_1 \equiv 0 \pmod{\pi^k}\}.$$

Then $L_k \supset L_{k+1} \supset \cdots$. If $f = (f_1, f_2) \in L_k$ then $\pi f = (\pi f_1, \pi f_2) \in L_{k+1}$. The elements $(f_1, f_2) \in L_k$ with $\pi \not| f_2$ represent fractions f_1/f_2 whose π-adic

Rational_Approximation()

begin
Input a_is until the first $a_{k-1} \neq 0$
Let $b \in S$ satisfy $ba_{k-1} \equiv 1 \pmod{\pi}$
Let d minimize $N(b + d\pi^{k-1})$
$\alpha = a_{k-1}\pi^{k-1}$
$f = (0, \pi)$
$g = (\pi^{k-1}, b + d\pi^{k-1})$
while more input **do**
  input a_k
  $\alpha = \alpha + a_k\pi^k$
  if $\alpha \cdot g_2 - g_1 \equiv 0 \pmod{\pi^{k+1}}$ **then**
    $f = \pi f$
    if $N(g) < N(f)$ **then**
      swap f and g
    fi
    Improve($\langle f, g \rangle$)
  else
    let $f + dg \in L_{k+1}$ with $d \in S$
    $\langle f, g \rangle = \langle \pi g, f + dg \rangle$
    if $N(g) < N(f)$ **then**
      swap f and g
    fi
    Improve($\langle f, g \rangle$)
  fi
  $k = k + 1$
od
if $\pi | f_2$ **then**
  return g
else return f
fi
end

Fig. 2. Algorithm **Rational_Approximation**.

expansion agrees with that of α in the first k places. Two pairs of elements $f, g \in L_k$ form a *basis* for L_k if every element $h \in L_k$ can be written $h = cf + dg$ for some $c, d \in R$. Such bases exist and are described in the following lemma, which is a key observation of [17, Lemma 2.1, p. 72]. Its proof is straightforward:

Lemma 16 *Two pairs of integers $f, g \in L_k$ form a basis for L_k if and only if $f_1g_2 - f_2g_1 = u\pi^k$ for some unit u.*

It follows that at every stage of the algorithm the pair $\langle f, g \rangle$ is a basis for L_k. A basis $\langle f, g \rangle$ for a lattice L is Φ-*minimal* if $\Phi(f)$ is minimal in L and $\Phi(g)$ is

Improve($\langle f, g\rangle$)
begin
while $\min_d(\Phi(g + df)) < \Phi(f)$ **do**
Let d minimize $\Phi(g + df)$
$\langle f, g\rangle = \langle g + df, f\rangle$
od
Let d minimize $\Phi(g + df)$
$g = g + df$
return $\langle f, g\rangle$
end

Fig. 3. Algorithm **Improve**.

minimal for elements of L_k that are independent of f.

Lemma 17 *If L is a lattice with basis $\langle f, g\rangle$, then algorithm* **Improve** *outputs a Φ-minimal basis for L.*

PROOF. Algorithm **Improve** halts eventually because in two steps it always reduces $\max(\Phi(f), \Phi(g))$ by at least one. So suppose $\langle f, g\rangle$ is a basis for L such that $\Phi(f) \le \Phi(g)$ and for all $d \in R$, $\Phi(g) \le \Phi(g + df)$. Suppose that $\Phi(f)$ is not minimal in L. Then there exist $a, b \in R$ such that $\Phi(af + bg) < \Phi(f)$. It is immediate that $a \neq 0$ and $b \neq 0$. Let $a = cb + r$, with $N(r) < N(b)$. Then

$$\Phi(cbf + bg) \le \Phi(af + bg) + \Phi(rf)$$

so

$$\begin{aligned}\Phi(cf + g) &< \frac{1}{N(b)}(\Phi(f) + (N(b) - 1)\Phi(f)) \\ &= \Phi(f),\end{aligned}$$

which is impossible.

If $\Phi(g)$ is not minimal among elements of L that are independent of f, then there exist $a, b \in R$ such that $\Phi(af+bg) < \Phi(g)$ and $b \neq 0$. A similar argument also leads to a contradiction. □

Thus at the end of the main loop of **Rational_Approximation** we have a Φ-minimal basis for L_k. If both $\pi | f_2$ and $\pi | g_2$, then π divides the second coordinate of every element of L_k, which is false. Thus Theorem 14 holds.

Proof of Theorem 15. By assumption, $\alpha = u/q$ with π not dividing q and $(u, q) \in L_k$ for all k. The output from the algorithm is a pair $h = (h_1, h_2) \in L_T$

with π not dividing h_2. There is a Φ-minimal basis $\langle f, g\rangle$ with $\Phi(f) \leq \Phi(g)$, and either $\pi | f$ and $h = g$ or $h = f$. In the former case, any element of L_T whose second component is not divisible by π is independent of f. Thus in either case h is the Φ-minimal element with π not dividing h_2. Thus $\Phi(h_1, h_2) \leq \Phi(u, q)$. Hence $N(h_1 q) = N(h_1)N(q) \leq \Phi(h_1, h_2) \cdot \Phi(u, q) \leq \Phi(u, q)^2$. Similarly, $N(h_2) \leq \Phi(u, q)^2$. However, $\alpha h_2 - h_1 \equiv 0 \pmod{\pi^T}$ so $h_1 q \equiv u h_2 \pmod{\pi^T}$, which implies that $h_1 q = u h_2$. Therefore $h_1/h_2 = u/q$. $\square$

Now let us consider the time complexity. The outer loop is iterated T times if T symbols of A are used. Algorithm **Improve** is called at most $2N(\pi) + 1$ times for each iteration of the outer loop. Thus the elements f and g are built up from at most cT operations in R. Let $\langle f, g\rangle$ be a Φ-minimal basis for the lattice $L = \{(u, v) : \alpha = u/v\}$, and $\lambda = \max(\Phi(f), \Phi(g))$. Then all the inputs and results of the operations in R involve elements h with $\Phi(h) \leq \lambda$. Suppose we have a function $\mu(k)$ such that every operation in R whose inputs and result are elements bounded by $\Phi(h) \leq t$ takes time at most $\mu(t)$. Let $\sigma(t)$ be the time required for the minimization step in **Improve**. Then the overall time complexity of the algorithm is bounded by $O(T(\mu(\lambda) + \sigma(\lambda)))$.

The minimization step is left unspecified and depends on the particular ring R and norm N. For example, suppose R is the ring of integers in an imaginary quadratic number field $\mathbf{Q}[\sqrt{D}]$ with $D \equiv 1 \bmod 4$. Thus $R = \mathbf{Z} + \mathbf{Z}\sqrt{D}$ and $N(a + b\sqrt{D}) = a^2 + Db^2$. If $d = x + y\sqrt{D}$, then $N(f_i + dg_i)^2$ can be written in the form

$$F_i(x, y) = (a_i + b_i x + c_i y)^2 + D(d_i + e_i x + h_i y)^2$$

with integer coefficients. Note that in the minimization we may work with the square of the norm.

The minimum of $\max(F_1, F_2)$ must occur either at a critical point of F_1 or F_2 or at a critical point of the intersection of F_1 and F_2. The minimum must exist because the surface is bounded below. By an affine change of coordinates F_i is equivalent to $x^2 + Dy^2$, so it has a single critical point which can be found by solving $a_i + b_i x + c_i y = 0$ and $d_i + e_i x + h_i y = 0$. (these linear equations are always independent).

The critical points of the intersection can be found by equating the derivative of $F_1(x, y)$ with respect to x to zero, and differentiating the constraint $F_1(x, y) = F_2(x, y)$. This leads to a pair of (inhomogeneous) quadratic equations in x and y, which can be solved. There are at most four solutions.

The value of d can then be found by considering the (at most six) critical points we have found and checking the nearest integer points. The entire procedure takes a constant number of operations in R. Thus in this case the complexity

of the entire algorithm is $O(T\mu(\lambda))$.

8.1 Rational Approximation by Interleaving

Even when there is no such rational approximation algorithm for AFSRs over a ring R, it may be possible to synthesize an AFSR for a given sequence A by thinking of it as an interleaving of several sequences over subrings. In this subsection we assume that S_0 is a subset of R such that $R = \mathbf{Z}[S_0, \pi]$, with $|S_0| = f$ and $\pi^e = p$, and that

$$S = \left\{ \sum_{\sigma \in S_0} b_\sigma \sigma : b_\in \mathbf{Z}, 0 \le b_\sigma < p \right\}$$

is a complete set of residues for R modulo π. For the AFSRs in this subsection we assume $T = S$. We also assume that

$$Q = \mathbf{Z}[V_0, \rho]$$

is a subring of R, with $|V_0| = h$ and $\rho^g = p$, and that

$$V = \left\{ \sum_{\tau \in V_0} c_\tau \tau : c_\in \mathbf{Z}, 0 \le c_\tau < p \right\}$$

is a complete set of residues for Q modulo π. We assume further that g divides e, h divides f, $\pi^{e/g} = \rho$, and that there is a set $U \subseteq R$ such that $|U| = f/h$ and

$$S_0 = \{\tau\phi : \tau \in V_0, \phi \in U\}.$$

The simplest example of such a subring is $Q = \mathbf{Z}$, where $g = h = 1$.

The idea is to decompose a sequence of elements of S into several sequences of elements of V, find rational approximations for these sequences, and then combine them into a rational approximation for the original sequence. Let $A = a_0, a_1, \cdots$, with $a_i \in S$ and

$$a_i = \sum_{\sigma \in S_0} a_{i,\sigma}\sigma, \quad a_{i,\sigma} \in \mathbf{Z}, \quad 0 \le a_{i,\sigma} < p,$$

and let

$$\alpha = \sum_{i=0}^{\infty} a_i \pi^i$$

be the associated element of $\hat{R}$. Then we can write

$$\sum_{\sigma} a_{i,\sigma}\sigma = \sum_{\tau\in V_0}\sum_{\phi\in U} b_{i,\tau,\phi}\tau\phi,$$

where $b_{i,\tau,\phi} = a_{i,\sigma}$ if $\sigma = \tau\phi$. It follows that

$$\alpha = \sum_{j=0}^{e/g-1}\sum_{\phi\in U}\left(\sum_{k=0}^{\infty}\left(\sum_{\tau\in V_0} b_{j+ke/g,\tau,\phi}\tau\right)\rho^k\right)\phi\pi^j.$$

For each j and ϕ, we let

$$B_{j,\phi} = \sum_{\tau\in V_0} b_{j,\tau,\phi}, \sum_{\tau\in V_0} b_{j+e/g,\tau,\phi}, \sum_{\tau\in V_0} b_{j+2e/g,\tau,\phi}, \cdots,$$

and let

$$\beta_{j,\phi} = \sum_{k=0}^{\infty}\left(\sum_{\tau\in V_0} b_{j+ke/g,\tau,\phi}\tau\right)\rho^k \in \hat{Q}.$$

Then

$$\alpha = \sum_{j=0}^{e/g-1}\sum_{\phi\in U}\beta_{j,\phi}\phi\pi^j.$$

Thus if we can find rational representations $\beta_{j,\phi} = u_{j,\phi}/q_{j,\phi}$, then we can write

$$\begin{aligned}\alpha &= \sum_{j=0}^{e/g-1}\sum_{\phi\in U}\frac{u_{j,\phi}}{q_{j,\phi}} \\ &= \frac{\sum_{j=0}^{e/g-1}\sum_{\phi\in U}\left(\prod_{(\psi,\ell)\neq(\phi,j)} q_{\psi,\ell}\right)}{\prod_{(\phi,k)} q_{\phi,k}}\phi\pi^j,\end{aligned}$$

which is a rational representation of α. If Q is a UFD, then we can improve this by using a denominator which is the least common multiple of the $q_{\phi,k}$.

This representation is, in general, not minimal. It is, however, close to minimal. To see this, let $\alpha = v/q$ be the best rational representation of α in the sense that q is the connection element of an AFSR of minimal length that outputs the sequence of coefficients in the π-adic expansion of α. Also let F and E be the fraction fields of R and Q, respectively, and let N be the norm function from F to E.

Lemma 18 *The element q divides $N(q)$.*

PROOF. Recall that if $R_1 \subseteq R_2$ are commutative rings without zero divisors, then an element u of R_2 is *integral over* R_1 if u is a root of a monic polynomial with coefficients in R_1. It is known that if $R_2 = R_1[u]$ for some u that is integral over R_1, then every element of R_2 is integral over R_1 [5, p. 270]. It follows that the characteristic polynomial of every element u of R_2 has coefficients in R_1 [6, p. 611]. The constant term of the characteristic polynomial is precisely the norm. By taking a series of such integral extensions, this is the situation we are in with Q and R. Hence we have a relation

$$q^k + a_1 q^{k-1} + \cdots + a_{k-1} q + N(q),$$

where each a_i is in Q and k is the degree of the extension of F over E. Therefore

$$N(q) = q(-q^{k-1} - a_1 q^{k-2} - \cdots - a_{k-1}),$$

and the theorem is proved. $\square$

It follows that for some $z \in R$ and $\{z_{j,\phi}\} \subseteq Q$, we can write

$$\begin{aligned} \alpha &= \frac{z}{N(q)} \\ &= \frac{\sum_{j=0}^{e/g-1} \sum_{\phi \in U} z_{j,\phi}}{N(q)} \phi \pi^j. \end{aligned}$$

Therefore, for each j, ϕ,

$$\frac{z_{j,\phi}}{N(q)} = \frac{u_{j,\phi}}{q_{j,\phi}}.$$

If the latter is a minimal rational representation in the sense that the numerator and denominator are relatively prime, then $q_{j,\phi}$ divides $N(q)$. Therefore the least common multiple of the $q_{j,\phi}$ also divides $N(q)$.

Proposition 19 *Suppose a rational representation for α is found as described*

above, and combined over a common denominator r which is the least common multiple of the denominators of the subsequences used. If v/q is any other rational representation for α, then r divides $N(q)$.

The importance of this fact is that the π-adic log of $N(q)$ (the largest power of π appearing in a π-adic expansion of $N(q)$ with coefficients in S) can be bounded in terms of the π-adic log of q.

Proposition 20 *There is a constant d, depending only on R and Q, such that for every q*

$$\log_\pi(N(q)) \leq d + \frac{ef}{gh}\log_\pi(q).$$

PROOF. One way to compute $N(q)$ is as the determinant of the E-linear transformation $x \to qx$ from F to itself. Let us write

$$q = \sum_{\phi \in U} \sum_{j=0}^{e/g-1} \left(\sum_{\tau \in V_0} \sum_{i=0}^{g-1} v_{\tau,i,\phi,j} \tau \rho^i \right) \phi \pi^j,$$

with $v_{\tau,i,\phi,j} \in \mathbf{Z}$. Then $N(q)$ is a linear combination of $\{\tau\rho^i : \tau \in V_0, 0 \leq i < g\}$ whose coefficients are polynomials in $\{v_{\tau,i,\phi,j} : \tau \in V_0, 0 \leq i < g\}$ with integer coefficients and degree $ef/(gh)$. Thus each polynomial is bounded by

$$c \cdot (\max\{|v_{\tau,i,\phi,j}| : \tau \in V_0, 0 \leq i < g\})^{\frac{ef}{gh}},$$

where c is an integer constant (at worst the maximum over all the polynomials of the sum of the absolute values of the coefficients in the polynomial). It follows that

$$\begin{aligned}\log_\pi(N(q)) &\leq \frac{ef}{gh}\max\{\log_\pi(v_{\tau,i,\phi,j}) + \frac{ie}{g}\} \\ &\leq d + \frac{ef}{gh}\log_\pi(q)\end{aligned}$$

for some d, as desired. $\square$

We would like to conclude that $\log_\pi(r)$ is not too large. Unfortunately, we cannot in general bound the π-adic log of a divisor of a number in terms of the log of the number. For example, if we let $\pi^2 = 2$, and $R = \mathbf{Z}[\pi]$, then for any i, $2(1+\pi)^i$ is a divisor of 2, and the π-adic log of $2(1+\pi)^i$ is unbounded as i ranges over all positive integers.

However, the situation is much simpler over the ordinary integers. In this case we have $g = h = 1$, $\rho = p$, $V_0 = \{0, 1\}$, and $U = S_0$. From the fact that r divides $N(q)$, it follows that

$$\log_\pi(r) = e \log_p(r) \leq e \log_p(N(q)) = \log_\pi(N(q)).$$

Also, we can assume that r is a positive integer, and hence is the connection element of an AFSR over R. We have shown the following.

Proposition 21 *Suppose A is an eventually periodic sequence over R, and k is the length of the smallest AFSR over R that outputs A. Then there is a constant d depending only on R such that an AFSR of length $d + efk$ that outputs A can be found by decomposing A into ef interleaved binary sequences, and finding minimal length AFSRs over $\mathbf{Z}$ for each of these sequences.*

9 The Function Field Case

Suppose $R = GF(p^n)[x]$, $\pi \in R$ is irreducible, and $S, T \subseteq R$ are complete sets of representatives for $K = R/(\pi)$. Then K is an extension of $L = GF(p^n)$ of degree $d = \deg(\pi)$, hence is $GF(p^{nd})$. It follows that the cardinality of S and T is p^{nd}. Furthermore, if we choose a basis for K over L, then every element of K can be treated as a d-tuple of elements of L. Thus a sequence of period t over K can be treated as a sequence of period dt over L.

Let A be the output from an AFSR of length r over (R, π, S, T) with initial memory of degree e. In this section we show that there is an LFSR sequence B over L whose linear span is at most rd plus the maximum of e and a constant that depends only on R, π, S, and T such that A can be transformed into B (and vice versa) by a finite state "filter" that also depends only on R, π, S, and T. First we treat a special case.

Proposition 22 *If S is closed under addition and under multiplication by elements of L, then the state transition function of an AFSR over (R, π, S, T) is linear over L. Thus if the length of the AFSR is r, then the linear span over L of the output is at most rd plus the maximum degree of the memory throughout its infinite execution.*

PROOF. We can take $\{1, x, \cdots, x^{d-1}\}$ as a basis for K over L. Addition and multiplication by fixed elements (the q_i) in $L[x]$ are always L-linear operations. By the closure properties of S, if $m_1, m_2, m_1', m_2' \in L[x]$, $a_1, a_2 \in S$, and $u, v \in L$ satisfy $m_i = a_i + \pi m_i'$, then $um_1 + vm_2 = (ua_1 + va_2) + \pi(um_1' + vm_2')$ with $ua_1 + va_2 \in S$. Thus the entire state change operation is linear.

It follows that such an AFSR is equivalent to a linear feedback (not necessarily shift) register. It has been shown, however, that the linear span of the output from such a register is at most its length [2]. The second conclusion of the proposition follows. □

An example of a set of representatives satisfying the closure property is $S_0 = \{t(x) : deg(t) < d\}$.

Now suppose we have an AFSR defined over (R, π, S, T) with S, T arbitrary and with length r and connection element q. In general such an AFSR does not have a linear state change function. Let $A = a_0, a_1, \cdots$ be the output from this AFSR, and $\alpha = \sum_{i=0}^{\infty} a_i \pi^i$ the associated π-adic number with coefficients in S. In the ring $\hat{R}$ of π-adic numbers, α can also be represented by a series $\sum_{i=0}^{\infty} b_i \pi^i$ with the b_i in S_0. We show two things. First, there is a finite state device that depends only on R, π, and S that takes $B = b_0, b_1, \cdots$ as input and outputs A. Second, the sequence B can be generated by an AFSR defined over (R, π, S_0, T) whose length is at most r and whose memory is small. Let $U = \max\{deg(u) : u \in S\}$, $V = \max\{deg(u) : u \in T\}$.

Consider the following finite state device. Its state at any time is an element t of R. At each step it inputs an element $b \in S_0$ and finds $a \in S$ and $t' \in R$ such that $b + t = a + \pi t'$. The device outputs a and changes state to t'. If the state is initially $t = 0$ and the input sequence is B, then the output will be A. Furthermore,

$$d + \deg(t') \leq \max(\deg(a), \deg(b), \deg(t)) \leq \max(\deg(a), d, \deg(t)),$$

so the degree of the state is bounded by $U - d$ during an infinite execution and this is indeed a finite state device. Furthermore, the inverse transformation can be realized by a finite state device constructed in the same way with the roles of S_0 and S reversed. The same bound on the degree of the state holds.

By Theorem 3, if $\alpha = u/q$, then

$$\deg(u) \leq \max((r-1)d + U + V, rd + \deg(m)),$$

where m is the initial memory. Also, we have equality if $\deg(m) > U + V - 1$. Now consider the AFSR over (R, π, S_0, T) that generates B. Since T is unchanged, the length of this AFSR is r. If m' is the initial memory of this AFSR, then

$$u = \sum_{i=0}^{r-1} \sum_{j=0}^{r-i-1} q_i b_j \pi^{i+j} - m' \pi^r.$$

Thus $\deg(m')+rd \leq \max(V+d-1+(r-1)d, (r-1)d+U+V, rd+\deg(m))$ so $\deg(m') \leq \max(V+1, U+V-d, \deg(m))$. Combining this with Proposition 22 we have proved the following.

Theorem 23 *If A can be generated by an AFSR over (R, π, S, T) of length r with initial memory of degree e, then there is a sequence B that has linear span at most $rd + \max(V+1, U+V-d, e)$ over L such that B can be transformed into A by a finite state device depending only on R, π, and S with $p^{n(U-d)}$ states.*

10 Conclusions

We have described a general method for constructing algebraic feedback shift registers over certain rings, R. These registers are analogous to linear feedback shift registers. They can be thought of as generating sequences by carrying out division in the completion $\hat{R}$ of the ring at a principal prime ideal (π). Associated with them are algebraic structures that are similar to those associated with LFSRs.

The cryptographic importance of these registers is twofold. First, they are a potential source of cryptographically secure sequences for stream ciphers. As with LFSR sequences, there are many possible (as yet unexplored) ways to modify these sequences that may make them secure. Second, these registers can be used for cryptanalysis in the cases where we have a rational approximation algorithm. Such an algorithm exists if R is a Euclidean domain with an extra condition on its norm. For a few rings R we have shown that these conditions occur. It remains to be seen whether other rings have these properties and, if not, whether there is a different rational approximation algorithm that works.

We have also shown that the generators that arise when R is a polynomial ring over a finite field are equivalent to certain "filtered" LFSRs and thus give nothing of new cryptographic interest.

We have considered only the case when π is prime. This affects primarily the analysis of the boundedness of the memory (which is critical if AFSRs are to be implemented) and the correctness of the rational approximation algorithm. It can be shown, however, that rational approximation algorithms (using quite different techniques) exist when $R = \mathbf{Z}$ and π is a composite integer. The case when $\pi = 4$ gives rise to sequences over $\mathbf{Z}/(4)$, which have generated a great deal of interest in coding theory recently. The case when π is not prime is the subject of a future paper.

11 Acknowledgement

The authors thank an anonymous referee who made several suggestions that significantly improved the manuscript. This referee also deserves a prize for speedy refereeing.

References

[1] Z. Borevich and I. Shafarevich, *Number Theory.* (Academic Press, New York, 1966).

[2] A. Chan, M. Goresky, and A. Klapper, On the linear complexity of feedback registers, *IEEE Trans. Info. Theory,* **IT-36** (1990) 640-645.

[3] S. Golomb, *Shift Register Sequences.* Aegean Park Press, Laguna Hills CA, 1982.

[4] M. Goresky and A. Klapper, Feedback registers based on ramified extensions of the 2-adic numbers – extended abstract, in: ed. A. de Santis, *Advances in Cryptology - Eurocrypt 1994, LNCS 718* (Heidelberg, Springer Verlag) 215-222.

[5] N. Jacobson, *Basic Algebra I.* (W.H. Freeman, San Francisco, 1974).

[6] N. Jacobson, *Basic Algebra II.* (W.H. Freeman, San Francisco, 1980).

[7] A. Klapper, The Vulnerability of Geometric Sequences Based on Fields of Odd Characteristic, *J. Cryptology*, **7** (1994) 33-51.

[8] A. Klapper, Feedback with carry shift registers over finite fields, in: ed. B. Preneel, *Fast Software Encryption, LNCS 1008* (Heidelberg, Springer Verlag, 1995) 170-178.

[9] A. Klapper and M. Goresky, 2-adic shift registers, in: ed. R. Anderson, *Fast Software Encryption, LNCS 809* (Heidelberg, Springer Verlag, 1994.) 174-178.

[10] A. Klapper and M. Goresky, Large period nearly deBruijn FCSR sequences, in: ed. J. Quisqater, *Advances in Cryptology - Eurocrypt 1995, LNCS 921* (Heidelberg, Springer Verlag, 1995.) 263-273.

[11] A. Klapper and M. Goresky, Cryptanalysis based on 2-adic rational approximation, in: ed. D. Coppersmith, *Advances in Cryptology - Crypto 95, LNCS 963* (Heidelberg, Springer Verlag, 1995.) 262-273.

[12] A. Klapper and M. Goresky, Arithmetic cross-correlation of FCSR sequences, *IEEE Trans. Info. Theory,* **43** (1997) 1342-1346.

[13] A. Klapper and M. Goresky, Feedback Shift Registers, 2-Adic Span, and Combiners with Memory, *J. Cryptology* **10** (1997) 111-147.

[14] J.L. Massey, Shift register sequences and BCH decoding, *IEEE Trans. Info. Theory,* **IT-15**, (1969) 122-127.

[15] J. Massey and R. Rueppel, Method of, and Apparatus for, Transforming a Digital Data Sequence into an Encoded Form, U.S. Patent No. 4,797,922, 1989.

[16] R. Rueppel, *Analysis and Design of Stream Ciphers.* Springer Verlag, New York, 1986.

[17] B. M. M. de Weger, Approximation lattices of p-adic numbers, *J. Num. Th.,* **24** (1986) 70-88.