Algebraic Shift Register Sequences

Mark Goresky Andrew Klapper

October 14, 2009

Acknowledgements

Mark Goresky thanks The Institute for Advanced Study. His participation in the writing of this book was partially supported by DARPA grant no. HR0011-04-1-0031.

Andrew Klapper thanks The Fields Institute at the University of Toronto, Princeton University, and The Institute for Advanced Study, where he was a visitor during part of the writing of this book. His participation in the writing of this book was partially supported by NSF grants CCF-0514660 and CCF-0914828, and by DARPA grant no. HR0011-04-1-0031.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

Pseudo-random sequences are ubiquitous in modern electronics and information technology. They are used, for example, as spreading codes in communications systems (such as cellular telephones and GPS signals), as components for generating keystreams for stream ciphers and other cryptographic applications, as sampling data for simulations and Monte Carlo integration, for timing measurements in radar and sonar signals and in GPS systems, as error correcting codes in satellite and other communications, as randomizers of digital signals to eliminate spectral lines, as counters in field programmable gate arrays, and in power on self tests.

In all cases speed in generating the sequences is important — any extra time spent generating sequences impacts the speed of the whole system, and thus impacts throughput, or accuracy, or some other important characteristic. In most cases it is also important that the sequence be able to be replicated (for example, so a receiver can undo some encoding by a transmitter). But many desirable characteristics of sequences are more specific to the application. In cryptography we want sequences that are unpredictable from short prefixes. In CDMA we want families of sequences with low pairwise correlations. Error correcting codes require families of sequence with large pairwise Hamming distance and efficient decoding algorithms. Many applications require sequences with uniform distributions of fixed size patterns. Some, such as Monte Carlo methods, require that many other statistical tests be passed.

For all of these applications there are solutions that rely on sequences generated by *linear feedback shift registers* (LFSRs). As the reader shall see, LFSRs are very high speed generators of sequences. There is a rich algebraic theory of LFSR sequences that often allows those generators that have the most desirable properties to be selected. For some applications there are also solutions that rely on other types of sequence generators such as *feedback with carry shift registers* (FCSRs). These are also very high speed generators of sequences. They have an algebraic theory that parallels that of LFSRs.

The goal of this book is to provide algebraic tools for the design of pseudo-random sequences. It is meant to be both a text book and a reference book. We present a unified approach based on algebraic methods, which allows us to simultaneously treat linear feedback shift registers, feedback with carry shift registers, and many other analogous classes of sequence generators. Part of the purpose of this book is to provide engineers who have not had a formal education in abstract algebra with the tools for understanding the modern study of shift register sequences. The requisite algebraic tools are developed in Chapters 2 through 5, and the main body of the book begins with Chapter 6.

Although they were preceded by many important papers and technical reports [44, 45, 52, 81, 124, 176, 195, 196], the publication of "Error correcting codes" [158] by W. W. Peterson

and "Shift Register Sequences" [53] by S. Golomb were milestones in the development of linear feedback shift register (LFSR) techniques for the generation of pseudorandom sequences. These books explained and exploited the deep connection between the architecture of the shift register and the mathematics of Galois theory in a way that makes for exciting reading, even today, forty years later. In Golomb's book, each "cell" of the shift register is a vacuum tube that can be either ON or OFF, and the output of the shift register is a pseudo-random sequence of zeroes and ones. Reading this book, one is tempted to run out and buy the parts to build one of these machines and watch it run.

One of the most fascinating aspects of this theory concerns the design of shift registers that produce maximal length sequences, or *m-sequences*, and the remarkable statistical and correlation properties of these sequences, which we describe in Chapter 13: besides having maximal length, each m-sequence $\mathbf{a}$ of rank k is also a (punctured) de Bruijn sequence[1] and its autocorrelation function is optimal. It is an amazing fact that the design problem can be completely solved using the Galois theory of finite fields. Although this fact was known already to L. E. Dickson [37] (in a slightly different language, of course, since electronic shift registers did not exist in 1919), it was rediscovered in the 1950's by the engineering community, and it remains one of the most compelling illustrations of how an abstract mathematical theory can unexpectedly become the key to understanding a complex physical system. Similarly, the distribution and correlation properties of m-sequences turn out to be related to a variety of mathematical abstractions including finite fields, difference sets (a subject that developed independently, but around the same time [72]) and even elliptic curves (see Section 13.5.d).

The explosive development of code division multiple access (CDMA) communications, especially with cellular telephones, has created considerable interest in finding families of pseudorandom sequences with low cross-correlation between distinct sequences in a given family. In Chapter 14 we discuss some of the more common families, including Gold codes, Kasami sequences, GMW sequences and Legendre sequences, as well as several more exotic variations on these themes. Our analysis is more "geometric" than the standard approach. We have only scratched the surface of this fascinating topic and the reader interested in a more complete study of correlation questions may wish to consult Golomb and Gong's recent book [55].

In the years since the publication of Golomb's book [53], the basic design of shift registers has been enhanced in several different directions. The "Galois" and "Fibonacci" modes were developed [159], many ways of interconnecting shift registers were analyzed, and perhaps most significantly, the binary state vacuum tubes were eventually replaced by cells with many possible states. Engineers were led, for example, to consider N-ary shift registers, whose cell contents are taken from the integers modulo N, or from a finite Galois field, or more generally from an algebraic ring. It turns out that much of the analysis of shift register sequences goes through in this more

[1] Each subsequence of length k, except the all-zero subsequence, occurs exactly once in each period of $\mathbf{a}$.

general setting. In Chapters 6, 13, and 14 we present this general analysis. Although the material is not new, it is derived from many disparate sources.

It is possible to enhance the basic shift register architecture in yet another way, by the addition of a small amount of *memory*. The memory is used as a "carry" in the calculations. For example, when two sequences of ones and zeroes are added, they can be added as elements of $\mathbb{Z}/(2)$ (or XOR addition) in which 1+1=0, or they can be added as "integers", in which 1+1 = 2 = 0 + a carry of 1. The difference is illustrated by the following example, where carries go to the right:

1	1	0	0	0	1	1	0	1	0
1	0	1	0	1	1	0	1	1	0
0	1	1	0	1	0	1	1	0	0

mod 2

1	1	0	0	0	1	1	0	1	0
1	0	1	0	1	1	0	1	1	0
0	0	0	1	1	0	0	0	1	1

with carry.

In fact, the *summation combiner* [167] does exactly the latter: it combines two binary sequences into a third, using addition with carry. It was originally proposed as a method for creating a difficult-to-predict bit stream from two relatively easy-to-predict bit streams, for cryptographic applications.

In an effort to analyze the summation combiner, the authors decided to incorporate this addition-with-carry into the architecture of a linear feedback shift register. The result was the *feedback with carry* shift register (FCSR) [101, 102, 106], described in Chapters 7, 16, and 19.

Around the same time a similar idea began circulating in the random number generation community, perhaps initiated by G. Marsaglia [130] and A. Zaman [132], and more fully developed by R. Couture and P. l'Écuyer [28, 29] and others, where the method became known as "add with carry" and "multiply with carry" (MWC) random number generators. These approaches appeared to be different at first because the add-with-carry generators involved only two "cells", each of which stores a very large integer, while the FCSR used many cells, each of which stores only a single bit. But in later papers, architectures were considered which involve (possibly) many cells, each storing (possibly) large integers, and in this setting the two methods are seen to be identical. The generation and analysis of FCSR and MWC sequences is covered in Chapters 7 and 16. Galois and Fibonacci versions of FCSR generators also exist, see Chapter 10.

As in the case of LFSRs, the FCSR architecture can be enhanced by considering cell contents taken from $\mathbb{Z}/(p)$, or a finite field, or even an arbitrary ring. But in these cases, the analysis becomes considerably more difficult (and interesting). The natural setting for all these architectures is the *algebraic feedback shift register* or AFSR, for which the general theory is developed in Chapter 8. In this generality, the theory of AFSRs includes both that of FCSRs (with cell contents in an arbitrary ring) and LFSRs. But the AFSR architecture also contains a number of new and interesting special cases as well, some of which are studied in some detail in Chapters 8, 9, 15, 17, and 20.

For example, another way in which the FCSR architecture can be enhanced is to delay the "carry" by a certain number of steps. Here is an example of addition in which the carry is delayed by one step:

$$\begin{array}{cccccccccc} 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 01 \end{array} \qquad \text{carry delay } +1 \tag{1.1}$$

It is possible to build hardware FCSR generators that implement this sort of addition; we call them d-FCSRs[2] and they are described in Chapter 9. The d-FCSR is a special case of an AFSR, and although the analysis of the d-FCSR is more difficult than that of the FCSR, it is somewhat simpler than the general AFSR and is still surprisingly complete.

For another example, the cells of an FCSR, or of a d-FCSR, could be polynomials. This gives rise to the function field FCSR, as described in Section 8.3.b and Chapter 15. There appears to be an almost unlimited number of variations on this theme.

As in the case of LFSR sequences, the maximal length FCSR sequences (which we refer to as ℓ-sequences) have several remarkable properties. Such a sequence is as equi-distributed as possible, given its period. Although the autocorrelation function of such a sequence is not known in general, it nevertheless has perfect *arithmetic autocorrelation* (a function that is the direct arithmetic analog of the usual autocorrelation function). These properties are investigated in Chapter 16. In Chapter 17 the analogous questions for maximal length d-FCSR sequences are considered. The case of AFSR sequences based on a function field is especially surprising: the ones of maximal length turn out to be punctured de Bruijn sequences with ideal auto-correlations, but (in the non-binary case) they are not necessarily m-sequences. See Chapter 15.

In Part IV of the book we change gears and consider the "synthesis" problem: given a periodic sequence $\mathbf{a}$, how can we construct a device (such as an LFSR) that will generate the sequence? The length of the smallest such LFSR is called the *linear span* or *linear complexity* of the sequence $\mathbf{a}$. The optimal result in this direction is the *Berlekamp-Massey algorithm* which predicts later terms of the sequence using the minimum possible amount of data, and it does so very efficiently. It is well known that this algorithm is "essentially" the same as the continued fraction expansion in the field of formal power series. In Section 18.2.d the exact relation between these two procedures is made explicit.

The Berlekamp-Massey algorithm therefore provides a possible technique for uncovering the keystream in a stream cipher, based only on the knowledge of the plaintext. It is therefore desirable that such a keystream should have an enormous linear span. Several standard techniques for

[2]Our original intention was that the d in d-FCSR was an integer indicating the amount of delay, but it has been since claimed that d stands fro "delay".

constructing sequences with large linear span are described in Section 18.5, but the reader should be aware that this is a rapidly changing field and none of these techniques is considered "secure" by modern standards.

The continued fractions approach to FCSR synthesis (that is, the construction of an FCSR that produces a given sequence) simply does not work. However, two successful approaches are described in Section 19.3. The first is based on the mathematical theory of *lattice approximations*, and the second is based on the extended Euclidean algorithm. These techniques were used to "break" the summation cipher described above. The problem of AFSR synthesis is even more difficult, but in Chapter 20 we describe Xu's algorithm which converges in many cases.

The mathematics behind all this material can be a daunting obstacle. Although we have included the appropriate mathematical background in Part I of this book, most of the chapters in the main part of the book have been written so as to be independent of this material, as much as possible. The reader is invited to start in the middle, on a topic of interest, and to read as far as possible until it becomes necessary to refer to Part I, using the index as a guide. It may even be possible to delay the retreat to Part I indefinitely, by skipping the proofs of the theorems and propositions.

Part I

Algebraic Preliminaries

Chapter 2 Abstract Algebra

Abstract algebra plays a fundamental role in many areas of science and engineering. In this chapter we describe a variety of basic algebraic structures that play roles in the generation and analysis of sequences, especially sequences intended for use in communications and cryptography. This include groups (see Section 2.1), rings (see Section 2.2), and polynomials over rings (see Section 2.4). We also explore characters and Fourier transforms, basic tools for understanding structures based on groups and rings (see Section 2.3).

2.1 Group theory

Groups are among the most basic building blocks of modern algebra. They arise in a vast range of applications, including coding theory, cryptography, physics, chemistry, and biology. They are commonly used to model symmetry in structures or sets of transformations. They are also building blocks for more complex algebraic constructions such as rings, fields, vector spaces, and lattices.

2.1.a Basic properties

Definition 2.1.1. *A* group *is a set G with a distinguished element e (called the* identity*) and a binary operation $*$ satisfying the following axioms:*

1. *(Associative law) For all $a, b, c \in G$, $(a * b) * c = a * (b * c)$.*
2. *(Identity law) For all $a \in G$, $a * e = e * a = a$.*
3. *(Inverse law) For all $a \in G$, there exists $b \in G$ such that $a * b = e$. The element b is called an* inverse *of a.*

A group G is commutative *or* Abelian *if it also satisfies the following axiom:*

4. *(Commutative law) For all $a, b \in G$, $a * b = b * a$.*

The *order* of a group G, denoted $|G|$, is its cardinality as a set.

Proposition 2.1.2. *Let G be a group. Then the following statements hold.*

1. *If $a, b \in G$ and $a * b = e$ then $b * a = e$.*
2. *Every $a \in G$ has a unique inverse.*
3. *The identity $e \in G$ is unique.*

Proof. To prove the first claim, suppose $a*b = e$. Let c be an inverse of b. By associativity we have $(b*a)*b = b*(a*b) = b*e = b$. Therefore $e = b*c = ((b*a)*b)*c = (b*a)*(b*c) = (b*a)*e = b*a$.

To prove the second claim, suppose $a * b = e = a * c$. Then $b = e * b = (b * a) * b = b * (a * b) = b * (a * c) = (b * a) * c = e * c = c$.

To prove the third claim, suppose e and f are both identities in G. Then $e * f = e$ since e is an identity, and $e * f = f$ since f is an identity. Thus $e = f$. □

Sometimes we use *multiplicative notation* and write a^{-1} to denote the inverse of a, ab for $a * b$, $a^0 = e$, and $a^n = aa^{n-1}$ for n a positive integer. Then $a^n a^m = a^{n+m}$ and $(a^n)^m = a^{nm}$. If G is Abelian, it is common to use *additive notation* in which we write $+$ instead of $*$, $-a$ instead of a^{-1}, $a - b$ for $a + (-b)$, and 0 instead of e. Also, $0a = 0$ and $na = a + (n-1)a$ for n a positive integer. Then $na + ma = (n + m)a$ and $n(ma) = (nm)a$. We sometimes write $e = e_G$ when considering several different groups.

Examples:

1. The integers $\mathbb{Z}$ with identity 0 and addition as operation is an Abelian group.

2. The rational numbers $\mathbf{Q}$ with identity 0 and addition as operation is an Abelian group.

3. The nonzero rational numbers $\mathbf{Q} - \{\mathbf{0}\}$ with identity 1 and multiplication as operation is an Abelian group.

4. If S is any set, the set of permutations of S is a (non-Abelian if $|S| \geq 3$) group with composition as operation and the identity function as identity. The order of the permutation group of S is $|S|!$.

5. For any $n \geq 1$, the set of invertible $n \times n$ matrices (that is, with nonzero determinant) with rational entries is a (non-Abelian if $n \geq 2$) group with multiplication as operation and the $n \times n$ identity matrix as identity.

6. If $N \geq 2$, a, and b are integers, then a is congruent to b modulo N, written $a \equiv b \pmod{N}$, if N divides $a - b$. This is an equivalence relation on $\mathbb{Z}$. Let $\mathbb{Z}/(N)$ denote the set of equivalence classes for this relation. That is, $\mathbb{Z}/(N)$ is the set of sets of the form

$$a + N\mathbb{Z} = \{a + Nb : b \in \mathbb{Z}\}.$$

Then $\mathbb{Z}/(N)$ is an Abelian group with the operation $(a + N\mathbb{Z}) + (b + N\mathbb{Z}) = (a + b) + N\mathbb{Z}$ and $0 + \mathbb{Z}$ as identity. To prove this it suffices to show that this definition of addition is independent of the choice of representatives a and b (that is, if $a + N\mathbb{Z} = c + N\mathbb{Z}$ and $b + N\mathbb{Z} = d + m\mathbb{Z}$, then $(a + b) + N\mathbb{Z} = (c + d) + N\mathbb{Z}$) and that the group axioms for $\mathbb{Z}/(N)$ follow immediately from the

group axioms for $\mathbb{Z}$. We have $|\mathbb{Z}/(N)| = N$. The elements of $\mathbb{Z}/(N)$ are sometimes referred to as *residues* mod N.

The set of equivalence classes of elements that are relatively prime to m, denoted $(\mathbb{Z}/(N))^\times$, is also an Abelian group, with multiplication as operation and 1 as unit. We denote the order of this group by $\phi(N)$, Euler's totient function (or "ϕ" function). That is, $\phi(N)$ is the number of positive integers less or equal to than m and relatively prime to m. We also define $\phi(1) = 1$. We say more about Euler's totient function in Section 2.2.d.

Following is a basic fact about groups that we use later.

Theorem 2.1.3. *If G is a finite group and $a \in G$, then $a^{|G|} = e$.*

Proof. First suppose that G is Abelian. Let us define a function from G to itself by $f(b) = ab$. This function is one-to-one (if $ab = ac$ then multiplying by a^{-1} on the left gives $b = c$), so it is also a permutation of G. Therefore

$$\prod_{b\in G} b = \prod_{b\in G} ab = a^{|G|} \prod_{b\in G} b.$$

Multiplying by the inverse of

$$\prod_{b\in G} b$$

gives the result of the theorem.

Now suppose that G is arbitrary. Nonetheless,

$$H = \{a^i : i = 0, 1, \cdots\}$$

is an Abelian group, so $a^{|H|} = e$. Thus it suffices to show that $|H|$ divides $|G|$. Consider the cosets bH with $b \in G$. Suppose two of these have a nonempty intersection, $bH \cap cH \neq \emptyset$. Then there are integers i, j so that $ba^i = ca^j$. It follows from this that every ba^k is in cH and every ca^k is in bH. That is, $bH = cH$. This implies that the set of all cosets bH forms a partition of G. Since each bH has cardinality $|H|$, $|G|$ is a multiple of $|H|$ as desired. □

2.1.b Subgroups

In this section we examine subsets of group that inherit a group structure of their own.

Definition 2.1.4. *If G is a group, then a subset $H \subseteq G$ is a subgroup if it is a group with the same operation as G and the same identity as G.*

This means that H is a subset of G such that (1) $e \in H$; (2) if $a, b \in H$, then $a + b \in H$; and (3) if $a \in H$, then $a^{-1} \in H$. Then the group axioms hold in H. Also, if G is Abelian then H is Abelian.

For example, the additive group of integers is a subgroup of the additive group of rational numbers. The set of cyclic permutations of $\{1, 2, \cdots, n\}$ is a subgroup of the group of all permutations.

If G_1 and G_2 are groups with operations $*_1$ and $*_2$ and identities e_1 and e_2, then their *direct product* $G_1 \times G_2 = \{(a, b) : a \in G_1, b \in G_2\}$ is a group with operation $(a, b) * (c, d) = (a * c, b * d)$ and identity (e_1, e_2). More generally, if $\{G_i : i \in I\}$ is any collection of groups, indexed by a set I, then the Cartesian product

$$\prod_{i \in I} nG_i$$

is a group, again called the direct product of $\{G_i : i \in I\}$. The group operation is defined coordinate-wise. If all the groups are Abelian, then so is the product. If $I = \{1, 2, \cdots, n\}$ for some natural number n, then we write

$$\prod_{i \in I} nG_i = \prod_{i=1}^{n} G_i = G_1 \times G_2 \times \cdots \times G_n.$$

If $a \in G$ then we let $\langle a \rangle$ denote $\{a^i : i \in \mathbb{Z}\}$. This set is an Abelian subgroup, called the *subgroup generated by* a. If $|\langle a \rangle| < \infty$ then we say the *order of* a is this cardinality, $|\langle a \rangle|$. Otherwise we say a has infinite order. Equivalently, the order of a is the least $k > 0$ such that $a^k = e$, if such a k exists. A group is *cyclic* if $G = \langle a \rangle$ for some a and then a is called a generator of G. Every cyclic group is Abelian. The group $\mathbb{Z}/(N)$ is cyclic with generator 1. It is sometime referred to as the *cyclic group of order* N.

We need a basic lemma from number theory.

Lemma 2.1.5. *Let a, b be integers. Then there exist integers s, t with $\gcd(x, y) = sx + ty$.*

Proof. First we may assume that x and y are nonnegative since we can negate x or y without changing either $\gcd(x, y)$ or the set of integers of the form $sx + ty$. We can also assume $y \leq x$.

Now we proceed by induction on y. If $y = 0$, then $\gcd(x, y) = x$ and we can take $s = 1$ and $t = 0$. Otherwise, let $x = ay + z$ with $0 \leq z < y$. Then by induction there exist integers u, v with $\gcd(y, z) = uy + vz$. But $\gcd(x, y) = \gcd(y, z) = uy + v(x - ay) = vx + (u - av)y$ as claimed. □

Theorem 2.1.6. *Every subgroup of a cyclic group is cyclic. Suppose $\langle a \rangle$ is a finite cyclic group with order n.*

1. *If k is a positive integer, then $\langle a^k \rangle$ is a subgroup of $\langle a \rangle$ of order $n/\gcd(n, k)$.*
2. *If $d|n$ and $d > 0$, then $\langle a \rangle$ contains one subgroup of order d.*

3. If $d|n$ and $d > 0$, then $\langle a \rangle$ contains $\phi(d)$ elements of order d.
4. $\langle a \rangle$ contains $\phi(n)$ generators.

Proof. Let H be a nontrivial subgroup of $\langle a \rangle$. H contains some a^k with $k > 0$. Let k be the smallest positive integer with $a^k \in H$ and let $a^m \in H$. Suppose k does not divide m. Then $\gcd(k, m) < k$ and $\gcd(k, m) = sk + tm$ for some integers s, t. Indeed, every common divisor ofThen

$$a^{\gcd(k,m)} = (a^k)^s (a^m)^t \in H,$$

which is a contradiction. Therefore $H = \langle a^k \rangle$. Thus every subgroup of $\langle a \rangle$ is cyclic.

(1) Let $H = \langle a^k \rangle$ and $b = \gcd(n, k)$. We have $(a^k)^r = e$ if and only if $n|kr$. Thus the order of H is the least positive r such that $n|kr$. This is equivalent to $(n/b)|(k/b)r$, and this is equivalent to $(n/b)|r$. That is, the order of H is n/b.

(2) By (1), a subgroup $H = \langle a^k \rangle$ has order $d|n$ if and only if $d = n/\gcd(n, k)$, or, equivalently, $d \cdot \gcd(n, k) = n$. Let $b = \gcd(n, k) = sn + tk$ for some $s, t \in \mathbb{Z}$. Then $e = a^n \in H$, so $a^b \in H$ as above. Since $b|k$, we also have $H = \langle a^b \rangle$. But $b = n/d$ so H is the unique subgroup of order d. Conversely, $\langle a^{n/d} \rangle$ is a subgroup of order d, proving existence.

(3) Let $n = df$. By (1), an element a^k has order d if and only if $\gcd(n, k) = n/d = f$. This holds precisely when $k = gf$ with g relatively prime to $n/f = d$ and $0 < k < n$. That is, $0 < g < d$. The number of such g is $\phi(d)$.

(4) Follows immediately from (3) with $d = n$. □

For example, the group $\mathbb{Z}$ is cyclic (with generator 1) so every subgroup is of the form $m\mathbb{Z} = \{mk : k \in \mathbb{Z}\}$ for some integer m.

2.1.c Homomorphisms

More generally, relationships between groups often arise as functions from one group to another that preserve all the relevant algebraic structures and operations.

Definition 2.1.7. *Let G and H be two groups. A function $\varphi : G \to H$ is a* homomorphism *if it preserves the group operations. That is, if for every $a, b \in G$ we have $\varphi(ab) = \varphi(a)\varphi(b)$. The* image *of φ, denoted by* $\mathrm{Im}(\varphi)$*, is the set of $b \in H$ such that there is $a \in G$ with $\varphi(a) = b$. The* kernel *of φ, denoted by* $\mathrm{Ker}(\varphi)$*, is the set of $a \in G$ such that $\varphi(a) = e_H$. The homomorphism φ is an* endomorphism *if $G = H$. It is an* epimorphism *or is* surjective *if it is onto as a set function. It is a* monomorphism *or is* injective *if it is one-to-one as a set function. It is an* isomorphism *if it is both injective and surjective. It is an* automorphism *if it is an endomorphism and an isomorphism.*

If G is a group and $a \in G$, then we can define the function $\varphi(n) = a^n$. This function is a homomorphism and is a monomorphism if and only if a has infinite order. If a has finite order

m, then φ induces a monomorphism from $\mathbb{Z}/m\mathbb{Z}$ to G. In particular, every infinite cyclic group is isomorphic to the integers $\mathbb{Z}$ and every finite cyclic group is isomorphic to the (additive) group $\mathbb{Z}/(m)$ where m is the order of any generator.

Proposition 2.1.8. *Let $\varphi : G \to H$ be a homomorphism. Then φ preserves identity elements and inverses. Morever* $\mathrm{Ker}(\varphi)$ *is a subgroup of G and* $\mathrm{Im}(\varphi)$ *is a subgroup of H.*

Proof. To see that φ preserves identities observe that $\varphi(e_G) = \varphi(e_G e_G) = \varphi(e_G)\varphi(e_G)$. Multiplying by $\varphi(e_G)^{-1}$ then gives $e_H = \varphi(e_G)$. To see that φ preserves inverses, let $a \in G$. Then $e_H = \varphi(e_G) = \varphi(aa^{-1}) = \varphi(a)\varphi(a^{-1})$ so $\varphi(a)^{-1} = \varphi(a^{-1})$ by uniqueness of inverses. The remaining statements are left to the reader. □

Proposition 2.1.9. *If $\varphi : F \to G$ and $\psi : G \to H$ are homomorphisms, then the composition $\psi \circ \varphi : F \to H$ is a homomorphism.*

Proof. For all $a, b \in F$, we have $(\psi \circ \varphi)(a + b) = \psi(\varphi(ab)) = \psi(\varphi(a)\varphi(b)) = \psi(\varphi(a))g(\varphi(b))$. □

Definition 2.1.10. *A pair of homomorphisms $\varphi : F \to G$ and $\psi : G \to H$ is* exact *(at G) if the kernel of ψ equals the image of φ. A sequence of maps*

$$1 \to F \to G \to H \to 1 \tag{2.1}$$

is a short exact sequence *if it is exact at F, G, and H. Here 1 denotes the trivial group with a single element.*

The short exact sequence in (2.1) splits *if there is a homomorphism $\mu : H \to G$ so that $g \cdot h$ is the identity.*

Note that in equation (2.1), exactness at F is equivalent to φ being injective and exactness at H is equivalent to ψ being surjective.

Proposition 2.1.11. *If $1 \to F \to G \to H \to 1$ is a short exact sequence and all three groups are finite, then $|G| = |F| \cdot |H|$.*

Proof. Let φ denote the homomorphism from F to G, and let ψ denote the homomorphism from G to H. Since ψ is surjective, there is a subset U of G that maps one to one and onto H. If b is any element of G, then there is some $u \in U$ so that $\psi(u) = \psi(b)$. Then bu^{-1} maps to the identity in H, so $bu^{-1} = a \in \mathrm{Im}(\varphi)$. Thus we can write $b = au$ with $a \in \mathrm{Im}(\varphi)$. Suppose that $au = a'u'$ for some $a, a' \in \mathrm{Im}(\varphi)$ and $u, u' \in U$. Then $u'u^{-1} = (a')^{-1}a \in \mathrm{Im}(\varphi)$. It follows that $\psi(u'u^{-1}) = e_H$, so $\psi(u') = \psi(u)$. By the choice of U, we have $u = u'$. Then also $a = a'$. It follows that for each b there is a unique representation in the form $b = au$. The proposition is follows from this. □

Proposition 2.1.12. *Suppose* $1 \to F \to G \to H \to 1$ *is a short exact sequence with* $\varphi : F \to G$ *and* $\psi : G \to H$*, all three groups are Abelian, and the short exact sequence splits via a homomorphism* $h : H \to G$*, then there is an isomorphism between* $F \times H$ *and* G *given by* $(a, b) \mapsto \varphi(a)\mu(b)$*. Conversely, if* $G = F \times H$*, then there is a short exact sequence as in (2.1), where* ψ *is the projection map and* φ *maps* a *to* $(a, 1)$*.*

Proof. In Proposition 2.1.11 we can take U to be the image of μ to prove the first statement. The converse is trivial. □

2.1.d Quotients

If m is any positive integer, then the set of multiples of m, $m\mathbb{Z}$, is a subgroup of the (additive) group $\mathbb{Z}$. In Sect. 2.1.a the *quotient group* $\mathbb{Z}/m\mathbb{Z}$ is defined as the set of equivalence classes of $\mathbb{Z}$ under the following equivalence relation: $a \equiv b \pmod{m}$ if $a - b \in m\mathbb{Z}$.

More generally, suppose G is any group and H is a subgroup of G. Define an equivalence relation by saying $a \sim b$ if there is an $h \in H$ such that $b = ah$ (The proof that this is an equivalence relation is left as an exercise). The equivalence class of a is $aH = \{ah : \ h \in H\}$ and is called the *left coset of* a. The set of left cosets is denoted G/H. It is not always possible to form a group out of these cosets (but see Sect. 2.1.e).

In fact, we could have started by defining $a \sim' b$ if there is an $h \in H$ such that $b = ha$. This is also an equivalence relation. The equivalence class Ha of a with respect to this relation is called the *right coset of* a, the set of which is denoted $H\backslash G$. If G is Abelian, then $Ha = aH$ for all $a \in G$. More generally:

Definition 2.1.13. *If* H *is a subgroup of* G*, then* H *is* normal in G *if for every* $a \in G$*, we have* $aH = Ha$ *or equivalently, if* $aha^{-1} \in H$ *for every* $a \in G$ *and every* $h \in H$*.*

Theorem 2.1.14. *If* H *is normal in* G*, then* $G/H = H\backslash G$ *is a group under the operation* $(aH)(bH) = abH$*.* □

In this case, G/H is called the *quotient group of* G *modulo* H. The natural mapping $G \to G/H$ (given by $a \mapsto aH$) is a homomorphism. If the set of left cosets is finite, then we say H has *finite index* in G. The number of left cosets (which equals the number of right cosets) is called the *index of* H *in* G. Thus if H is normal in G and of finite index, then G/H is finite and $|G/H|$ equals the index of H in G. If G is finite, so is G/H, and we have $|G/H| = |G|/|H|$.

Theorem 2.1.15. *If* $\varphi : G \to G'$ *is a homomorphism then the following statements hold.*

1. $\mathrm{Ker}(\varphi)$ *is normal in* G*.*
2. *The quotient* $G/\mathrm{Ker}(\varphi)$ *is isomorphic to* $\mathrm{Im}(\varphi)$*.*

3. *Conversely, if H is a normal subgroup of G, then the natural mapping $a \mapsto aH$ is a surjection from G to G/H with kernel equal to H.*

□

Thus if H is normal in G, then we have a short exact sequence

$$1 \to H \to G \to G/H \to 1.$$

2.1.e Conjugacy and groups acting on sets

Two elements a and b in a group G are *conjugate* if there is an element $g \in G$ such that $b = gag^{-1}$. This is an equivalence relation on G whose equivalence classes are called *conjugacy classes*. If G is Abelian, then every conjugacy class has a single element, but if $ab \neq ba$, then both a and bab^{-1} are distinct elements in the same conjugacy class. Thus the number of conjugacy classes gives some measure of how far G is from being Abelian. If H and H' are subgroups of G, we say they are conjugate if there is an element $g \in G$ such that $H' = hHg^{-1}$.

An *action* of a group G on a set S is a mapping $G \times S \to S$, written $(g, s) \mapsto g \cdot s$, such that $(gh) \cdot s = g \cdot (h \cdot s)$ for all $g, h \in G$ and all $s \in S$. It follows that the identity $e \in G$ acts trivially ($e \cdot s = s$) and that each $g \in G$ acts by permutations. The *orbit* of an element $s \in S$ is the set

$$G \cdot s = \{g \cdot s : \ g \in G\}.$$

If $H \subset G$ is a subgroup then G acts on G/H by $g \cdot xH = (gx) \cdot H$.

If G acts on S, and if $s \in S$, define the *stabilizer* or *isotropy subgroup*,

$$\mathrm{Stab}_G(s) = \{g \in G : \ g \cdot s = s\}.$$

It is a subgroup of G. If $s, s' \in S$ are in a single orbit then their stabilizers are conjugate. In fact if $s' = gs$ then $\mathrm{Stab}_G(s') = g\mathrm{Stab}_G(s)g^{-1}$. The action of G on S is *transitive* if there is a single orbit, i.e., for every $s, s' \in S$ there exists $g \in G$ such that $s' = g \cdot s$. Suppose this to be the case, choose a "base point" $s_0 \in S$, and let $H = \mathrm{Stab}_G(s_0)$. This choice determines a one to one correspondence $\varphi : G/H \to S$ with $\varphi(gH) = g \cdot s_0$. The mapping φ is then compatible with the actions of G on G/H and on S, that is, $g \cdot \varphi(xH) = \varphi(g \cdot xH)$ for all $g \in G$ and all $xH \in G/H$. If $|G| < \infty$ it follows that $|S| = |G|/|H|$ divides $|G|$.

The group G acts on itself by translation ($g \cdot x = gx$) and by conjugation ($g \cdot x = gxg^{-1}$). The first action is transitive; the second is not, and its orbits are the conjugacy classes of G.

2.1.f Finitely generated Abelian groups

An Abelian group G is *finitely generated* if there is a finite set $V \subseteq G$ such that every element of G is equal to a finite product of elements of V. We state without proof the fundamental theorem of finite Abelian groups (See, for example, Lang [119, p. 46]):

Theorem 2.1.16. *Let G be a finitely generated Abelian group. Then G is isomorphic to a direct product of cyclic groups.*

Corollary 2.1.17. *Let G be a finite Abelian group with nm elements, where n and m are relatively prime positive integers. Then there are groups H_1 and H_2 with n and m elements, respectively, so that G is isomorphic to $H_1 \times H_2$.*

An element g in an Abelian group G is a *torsion* element if $g \neq 0$ and if some finite sum $g + g + \cdots + g = 0$ vanishes. That is, if it has finite order. The group G is *torsion-free* if it contains no torsion elements.

Corollary 2.1.18. *Let G be a finitely generated torsion-free Abelian group. Then G is isomorphic to a direct product of finitely many copies of $\mathbb{Z}$.*

2.2 Rings

Many important algebraic structures come with two interrelated operations. For example, addition and multiplication of integers, rational numbers, real numbers, and complex numbers; AND and XOR of Boolean valued functions; and addition and multiplication of $n \times n$ matrices of integers, etc.

Definition 2.2.1. *A ring R is a set with two binary operations $+$ and $\cdot$ and two distinguished elements $0, 1$ which satisfy the following properties for all $a, b, c \in R$:*

1. *R is an Abelian group with operation $+$ and identity 0;*
2. *$a \cdot (b \cdot c) = (a \cdot b) \cdot c$ and $1 \cdot a = a \cdot 1 = a$; and*
3. *$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ and $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$ (the distributive law).*

It follows that $a \cdot 0 = 0$ for all a, since $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. If $0 = 1$ then $R = \{0\}$ is the zero ring. It is common to denote by R^+ the Abelian group that is obtained from R by forgetting the multiplication operation.

A ring R is *commutative* if $a \cdot b = b \cdot a$ for all $a, b \in R$. *Throughout this book, all rings are commutative unless otherwise stated.* We generally write ab for the product $a \cdot b$.

2.2.a Units and zero divisors

Let R be a commutative ring. An element $a \in R$ is a *unit* if it is invertible, that is, if there exists $b \in R$ so that $ab = 1$. In this case b is unique. The collection of all units in R is denoted $R^\times$. It forms an Abelian group (under multiplication). An element $a \in R$ is a *zero divisor* if there exists a nonzero element $b \in R$ such that $ab = 0$. The ring of integers $\mathbb{Z}$ has no zero divisors, but only

1 and -1 are units. However if the ring R is finite then a given element is either a unit or a zero divisor. Indeed, let $\varphi_a : R \to R$ be the mapping which is given by multiplication by a. If φ_a is one to one, then it is also onto, hence a is invertible. If φ_a is not one to one, then there exist $b \neq c$ so that $ab = ac$ or $a(b-c) = 0$, so a is a zero divisor. If $a, b \in R$ and $ab = 0$, then b is said to *annihilate* a, and the set of such b is called the *annihilator* of a, denoted Z_a.

Let $b \in R$ be a unit. The smallest integer $m > 0$ such that $b^m = 1$ is called the multiplicative *order* of b, if such an $m < \infty$ exists; otherwise b is said to have infinite order. If $b \in R$ has order $m < \infty$, if $u \in R$ and if $t > 0$ is the smallest integer such that $(b^t - 1)u = 0$ then t divides m. (For, the group $\mathbb{Z}/(m)$ acts transitively on the set $\{u, bu, \cdots, b^{t-1}u\}$ with $k \in \mathbb{Z}/(m)$ acting by multiplication by b^k.) In particular, if $s > 0$ is relatively prime to m then $b^s - 1$ is not a zero divisor in R

Definition 2.2.2. *An* integral domain *(also called an* entire ring*) is a commutative ring with no zero divisors. A field is a commutative ring in which every nonzero element is invertible.*

In particular, a finite integral domain is necessarily a field. Every commutative ring R embeds in a ring $S^{-1}R$ which has the property that every element is either a zero divisor or is invertible, cf. Section 2.2.e.

2.2.b Ideals and quotients

Definition 2.2.3. *A* subring *S of a ring R is a subset of R, which is a ring under the same operations as R, and with the same zero and identity.*

If I is an additive subgroup of R (meaning that if $a, b \in I$ then $a + b \in I$ and $-a \in I$) then the quotient R/I is the set of equivalence classes under the equivalence relation $a \sim b$ if $a - b \in I$. The equivalence class containing $a \in R$ is the coset $a + I$. Then R/I is an Abelian group under addition: $(a + I) + (b + I) = a + b + I$. However, the multiplication operation on R does not necessarily induce a well defined multiplication on R/I. For if $a' \sim a$, say, $a' = a + c$ and if $b' \sim b$, say, $b' = b + d$ (where $c, d \in I$) then $a'b' = ab + ad + bc + cd$ which is not equivalent to ab unless $ad + bc + cd \in I$. The following definition is necessary and sufficient to ensure this holds for all $a, b \in R$ and $c, d \in I$.

Definition 2.2.4. *An ideal is an additive subgroup $I \subset R$ such that for any $a \in I$ and for any $b \in R$ we have $ab \in I$.*

It follows that the set of equivalence classes R/I inherits a ring structure from R if and only if I is an ideal. Two elements $a, b \in R$ are said to be *congruent modulo I* if they are in the same equivalence class. That is, if $a - b \in I$. Each equivalence class is called a *residue class modulo I*.

An ideal I is *proper* if $I \neq R$, in which case it does not contain any units. An ideal I is *principal* if there exists an element $a \in R$ such that $I = \{ar : r \in R\}$, in which case we write

$I = (a)$. If I, J are ideals then the sum $I + J$ is the set of all sums $a + b$ where $a \in I$ and $b \in J$. It is an ideal and is the smallest ideal containing both I and J. The intersection $I \cap J$ is also an ideal. The *product ideal* IJ is the set of all finite sums $\sum a_i b_i$ where $a_i \in I$ and $b_i \in J$. An ideal $I \subset R$ is *maximal* if I is proper and is not a proper subset of any other proper ideal. An ideal I is *prime* if $ab \in I$ implies $a \in I$ or $b \in I$. An ideal $I \subset R$ is *primary* if $I \neq R$ and whenever $ab \in I$, either $a \in I$ or $b^n \in I$ for some $n \geq 1$.

A field contains only the ideals (0) and (1).

Theorem 2.2.5. *Let R be a commutative ring. Then the following statements hold.*

1. *An ideal $P \subset R$ is maximal if and only if R/P is a field (called the* residue field with respect to P*).*
2. *An ideal $P \subset R$ is prime if and only if R/P is an integral domain. (See Definition 2.2.13.)*
3. *Every maximal ideal is prime.*

Proof. (1) Let P be maximal and $a \in R - P$. Then $J = \{ab + c : b \in R, c \in P\}$ is closed under addition and under multiplication by elements of R. It contains P (take $b = 0$) and a (take $b = 1$ and $c = 0$) so it properly contains P. By maximality it is not a proper ideal, so it must not be a proper subset of R. That is, $J = R$. In particular, $1 \in J$, so $1 = ab + c$ for some $b \in R$ and $c \in P$. Therefore $(a+P)(b+P) = ab+P = 1-c+P = 1+P$ so $a+P$ is invertible in R/P. Thus R/P is a field. On the other hand, suppose R/P is a field and J is an ideal containing P. Let $a \in J - P$. Then $a + P$ is invertible in R/P, so there is a $b \in R$ such that $(a + P)(b + P) = 1 + P$. That is, such that $ab = 1 + c$ for some $c \in P$. But then $1 = ab - c \in J$. By closure under multiplication by R, we have $R \subseteq J$. But this contradicts the fact that J is an ideal. Therefore P is maximal.

(2) Let $a, b \in R$. Then $(a + P)(b + P) = 0$ in R/P if and only if $ab \in P$. If P is prime, this says $(a + P)(b + P) = 0$ implies $a \in P$ or $b \in P$, which implies $a + P = 0$ or $b + P = 0$ in R/P, so R/P is an integral domain. Conversely, if R/P is an integral domain, then $ab \in P$ implies $(a + P)(b + P) = 0$ which implies $a + P = 0$ or $b + P = 0$. That is, $a \in P$ or $b \in P$, so P is a prime ideal.

(3) This follows from (1) and (2). □

For example, consider the ring of ordinary integers $\mathbb{Z}$. Let I be an ideal containing a nonzero element. Multiplication by -1 preserves membership in I, so I contains a positive element. Let m be the least positive element of I. Suppose that $a \in I$ is any other element of I. Then $\gcd(m, a) = um + va$ for some integers u and v, so $\gcd(m, a) \in I$. We have $\gcd(m, a) \leq m$, so by the minimality of m, $\gcd(m, a) = m$. That is, m divides a. Since every multiple of m is in I, it follows that I consists exactly of the multiples of m. In particular, $I = (m)$ is principal.

The ideal (m) is contained in the ideal (n) if and only if m is a multiple of n. The ideal (m) is prime if and only if m is prime. In this case it is also maximal. It is primary if and only if m is a power of a prime.

Definition 2.2.6. *A function $\varphi : R \to S$ from a ring R to a ring S is a* ring homomorphism *if $\varphi(a+b) = \varphi(a)+\varphi(b)$ and $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in R$. The homomorphism φ is a* surjection *(or* epimorphism*) if it is onto. It is an* injection *(or* monomorphism*) if it is one to one. It is an* isomorphism *if it is both an injection and a surjection. It is an* endomorphism *if $R = S$. It is an* automorphism *if it is an endomorphism and an isomorphism.*

The set of automorphisms of a ring S forms a group under composition, denoted by $\mathrm{Aut}(S)$. More generally, if R is a subring of S (we also say that S is an *extension of* R), then the set of automorphisms of S whose restrictions to R are the identity forms a subgroup $\mathrm{Aut}_R(S)$. The proof of the following theorem is left as an exercise.

Theorem 2.2.7. *If $\varphi : R \to S$ is a ring homomorphism, then*

$$\mathrm{Ker}(\varphi) = \{\mathrm{r} \in \mathrm{R} : \ \varphi(\mathrm{r}) = 0\}$$

is an ideal of R, the image of φ is a subring of S, and φ induces an isomorphism between $R/\mathrm{Ker}(\varphi)$ and $\mathrm{Im}(\varphi)$. Conversely, if I is an ideal of R then the map $a \mapsto a+I$ is a surjective homomorphism from $R \to R/I$ with kernel I.

If $f : R \to S$ is a surjective ring homomorphism with kernel $I \subset R$, then

$$0 \to I \to R \to S$$

is a short exact sequence of additive groups. Sometimes there is also an injection $g : S \to R$ such that $f \circ g$ is the identity function (a *right inverse* of f). In this case it makes sense to think of S as a subring of R so that R is an algebra over S. We say that g is a *splitting* of f.

If R is a ring and $a \in R$, then the annihilator Z_a of a is an ideal. It is proper because $1 \notin Z_a$.

2.2.c Characteristic

Let R be a commutative ring. If m is a nonnegative integer, we write $m \in R$ for the sum $1+1\cdots+1$ (m times). This defines a homomorphism from $\mathbb{Z}$ into R. That this function is a homomorphism can be shown by a series of induction arguments. In fact this is the unique homomorphism from $\mathbb{Z}$ into R, since any such homomorphism is completely determined by the facts that $1_{\mathbb{Z}}$ maps to 1_R, and the ring operations are preserved. The kernel of this homomorphism is an ideal in $\mathbb{Z}$, hence by the example in Section 2.2.b is of the form (m) for some nonnegative integer m. This integer is called the *characteristic* of R. For any $a \in R$, we have $ma = a+a+\cdots+a$ (m times). Hence if the characteristic is nonzero, it is the smallest positive integer m such that $ma = 0$ for all $a \in R$. If the characteristic is zero, then no such m exists and $\mathbb{Z}$ is isomorphic to a subring of R. Otherwise $\mathbb{Z}/(m)$ is isomorphic to a subring of R. If R is finite then its characteristic is positive since the sequence of elements $1, 2, 3, \cdots \in R$ must eventually lead to a repetition.

Theorem 2.2.8. *If R is an integral domain then its characteristic is either 0 or is a prime number. In particular, the characteristic of any finite field is prime.*

Proof. Suppose R is an integral domain. Let $k > 0$ be the characteristic and suppose $k = mn$ (in $\mathbb{Z}$), with $m, n > 0$. Then $mn = 0$ in R, so $m = 0$ or $n = 0$ in R. Suppose $m = 0$. For any $c \in R$, the element $c + \cdots + c$ (m times) is $mc = 0$. By the minimality of k, we must have $m = k$ and $n = 1$. A similar argument holds when $n = 0$ in R. It follows that k is prime. □

Lemma 2.2.9. *Let R be a commutative ring. If the characteristic k of R is a prime number, and if q is any positive power of k then*

$$(a + b)^q = a^q + b^q \in R \tag{2.2}$$

for every $a, b \in R$.

Proof. Suppose k is prime and $0 < m < k$. The binomial coefficient

$$\binom{k}{m} = \frac{k!}{m!(k-m)!}$$

is divisible by k since k appears as a factor in the numerator but not in the denominator. Consequently $(a + b)^k = a^k + b^k$ and equation (2.2) follows by induction. □

If k is not prime, then equation (2.2) is generally false.

2.2.d The Ring $\mathbb{Z}/(N)$ and primitive roots

In this section we continue the example of the modular integers introduced in Section 2.1.a. Fix a nonzero integer N. The ring $\mathbb{Z}/(N)$ is the (cyclic) group of order N, $\mathbb{Z}/(N)$, together with the operation of multiplication. The same symbol is used for both structures, which often causes some confusion. The group $\mathbb{Z}/(N)$ is sometimes referred to as the *additive group* of $\mathbb{Z}/(N)$. The characteristic of the ring $\mathbb{Z}/(N)$ is $|N|$.

As in Section 2.2.c, the mapping $\pmod N : \mathbb{Z} \to \mathbb{Z}/(N)$ is a ring homomorphism. If $x \in \mathbb{Z}$ we sometimes write $\overline{x} \in \mathbb{Z}/(N)$ for its reduction modulo N. Conversely, it is customary to represent each element $y \in \mathbb{Z}/(N)$ by the corresponding integer $\widehat{y} \in \mathbb{Z}$ with $0 \le \widehat{x} \le N - 1$, but note that this association $\mathbb{Z}/(N) \to \mathbb{Z}$ is neither a group nor a ring homomorphism. It is also common to omit the "bar" and the "hat", thereby confusing the integers between 0 and $N - 1$ with $\mathbb{Z}/(N)$.

If m divides N then the mapping $\pmod m : \mathbb{Z}/(N) \to \mathbb{Z}/(m)$ is a ring homomorphism. If a, b are nonzero, relatively prime integers, then the mapping

$$\mathbb{Z}/(ab) \to \mathbb{Z}/(a) \times \mathbb{Z}/(b) \tag{2.3}$$

given by $x \mapsto (x \pmod a, x \pmod b)$ is a ring isomorphism. (Since both sides have the same number of elements it suffices to check that the kernel is zero. But if x is divisible by a and by b and if a, b are relatively prime, then x is divisible by ab. This is a special case of the *Chinese remainder theorem*, Theorem 2.2.18).

Let $x \in \mathbb{Z}$. The following statements are equivalent:

1. The element $\overline{x} = x \pmod N \in \mathbb{Z}/(N)$ is invertible in $\mathbb{Z}/(N)$.

2. The element $\overline{N} = N \pmod x \in \mathbb{Z}/(x)$ is invertible in $\mathbb{Z}/(x.)$

3. The integers x and N are relatively prime.

4. There exists $n > 0$ so that $x|(N^n - 1)$

5. There exists $m > 0$ so that $N|(x^m - 1)$

6. The element $\overline{x} \in \mathbb{Z}/(N)$ generates the *additive* group of $\mathbb{Z}/(N)$. That is, the elements $\{0, \overline{x}, \overline{x} + \overline{x}, \overline{x} + \overline{x} + \overline{x}, \cdots\}$ account for all the elements in $\mathbb{Z}/(N)$.

As we saw in Section 2.2.c, the units in $\mathbb{Z}/(N)$ form an Abelian group under multiplication, the *multiplicative group* $\mathbb{Z}/(N)^\times$. Euler's totient function, $\phi(N) = |\mathbb{Z}/(N)^\times|$ is defined to be the number of units in $\mathbb{Z}/(N)$. For any $y \in \mathbb{Z}/(N)^\times$ there is a least power d, called the *order* of y and denoted $d = \mathrm{ord}_N(y)$, such that $y^d = 1 \in \mathbb{Z}/(N)$. It follows from Theorem 2.1.3 that $d|\phi(N)$, so if y is relatively prime to N then $y^{\phi(N)} \equiv 1 \pmod N$, which is called *Fermat's congruence* or Fermat's little theorem. The least n, m in (4), (5) is $n = \mathrm{ord}_x(\overline{N})$ and $m = \mathrm{ord}_N(\overline{x})$ respectively.

Lemma 2.2.10. *Let N be a positive integer. Then the following statements hold.*

1. *If $N = \prod_{i=1}^k p_i^{m_i}$ is the prime factorization of N then $\phi(N) = \prod_{i=1}^k \phi(p_i^{m_i})$.*

2. *$\phi(p^m) = p^{m-1}(p-1)$ if p is prime.*

3. *$N = \sum_{d|N} \phi(d)$.*

Proof. The first statement follows from equation (2.3). In the second statement, since p is prime, the only integers that are not relatively prime to p^m are the multiples of p. There are p^{m-1} of these in $\mathbb{Z}/(p^m)$, which leaves $p^{m-1}(p-1)$ integers that are relatively prime to p, proving the second statement. (In fact, the group $\mathbb{Z}/(p^m)^\times$ is described in Section 4.2.a: it is cyclic of order $p^{n-1}(p-1)$ if $p > 2$. If $p = 2$ and $m \geq 3$ then it is a product of two cyclic groups, one of order 2

(generated by -1) and one of order p^{m-1}, generated by 5.) For the third statement, consider the set of fractions $\{1/N, 2/N, \cdots, N/N\}$. They are distinct, positive, and ≤ 1. Reduce each of these to its lowest terms. Then the denominator of each fraction will be a divisor d of N. For a given denominator d the possible numerators will be any integer relatively prime to d and $\leq d$, so there are $\phi(d)$ of them. Therefore, adding $\phi(d)$ over all $d|N$ gives N. □

From the comments in the preceding paragraph, it follows that the multiplicative group $\mathbb{Z}/(N)^\times$ is cyclic if and only if $N = p^m, 2p^m, 2$, or 4, where $p \geq 3$ is an odd prime. In this case a generator $a \in \mathbb{Z}/(N)^\times$ is called a *primitive root* modulo N. The number of primitive roots modulo N is therefore $\phi(\phi(N))$. The *Artin conjecture* states in part that each prime number $p \in \mathbb{Z}$ is a primitive root modulo q for infinitely many primes q. The following proposition helps enormously in verifying primitivity modulo a prime power p^t. (cf. Section 16.1.)

Lemma 2.2.11. *Let p be prime and let $s \geq 1, t \geq 1, b \in \mathbb{Z}$. Then b is a unit modulo p^s if and only if it is a unit modulo p^t.*

Proof. We may assume that $s = 1$. If a is a unit modulo p^t, then $p^t|bc - 1$ for some b, so $p|bc - 1$ as well, and b is a unit modulo p.

Conversely, suppose b is a unit modulo p, so $p|bc - 1$ for some c. We claim by induction that for all i there is a c_i so that $p^i|bc_i - 1$. Indeed, for $i \geq 2$ by induction let

$$bc_{i-1} = 1 + p^{i-1}d_{i-1}.$$

Then by the binomial theorem,

$$(bc_{i-1})^p = (1 + p^{i-1}d_{i-1})^p = 1 + p^i d_{i-1} + (p^{i-1})^2 z$$

for some integer z. But $2(i-1) = 2i - 2 \geq i$, so

$$b(b^{p-1}c_{i-1}^p) \equiv 1 \pmod{p^i}$$

as claimed. In particular, b is a unit modulo p^t. □

Proposition 2.2.12. *Suppose $N = p^t$ with $p \geq 3$ an odd prime and $t \geq 2$. Let $2 \leq a \leq N - 1$. Then a is primitive modulo N if and only if a is primitive modulo p^2. This holds if and only if a is primitive modulo p, and p^2 does not divide $a^{p-1} - 1$.*

Proof. If a is primitive modulo p^t, then by Lemma 2.2.11 every unit b modulo p or p^2 is a unit modulo p^t. Thus b is congruent to a power of a modulo t, and hence also modulo p and p^2. Thus a it is primitive modulo p and p^2. We prove the converse by induction on t, following [83, Section 4.1 Theorem 2]. Fix $t \geq 3$ and suppose that a is primitive in $\mathbb{Z}/(p^s)$ for all $s < t$. The order of

a is a divisor of $\phi(p^t) = p^{t-1}(p-1)$, the cardinality of the group of units in $\mathbb{Z}/(p^t)$. We want to show that the order of a is not a divisor of $p^{t-1}(p-1)/r$, for any prime divisor r of $p^{t-1}(p-1)$. First we take $r = p$. Since

$$a^{p^{t-3}(p-1)} = a^{\phi(p^{t-2})} \equiv 1 \pmod{p^{t-2}}$$

we have

$$a^{p^{t-3}(p-1)} = 1 + cp^{t-2}$$

for some $c \neq 0$, and c is relatively prime to p since a is primitive for $\mathbb{Z}/(p^{t-1})$. Then

$$a^{p^{t-2}(p-1)} = (1 + cp^{t-2})^p \equiv 1 + cp^{t-1} \not\equiv 1 \pmod{p^t}$$

since $\binom{p}{i}$ is a multiple of p. This shows that the order of a modulo p^t is not $\phi(p^t)/r$ with $r = p$.

Now suppose that r is a prime divisor of $p-1$ and

$$a^{p^{t-1}(p-1)/r} \equiv 1 \pmod{p^t}.$$

Let b be a primitive element modulo p^t and $a \equiv b^k \pmod{p^t}$. Then

$$b^{kp^{t-1}(p-1)/r} \equiv 1 \pmod{p^t}$$

so $p^{t-1}(p-1)$ divides $kp^{t-1}(p-1)/r$. Equivalently, r divides k. But then

$$a^{p^{t-2}(p-1)/r} \equiv b^{kp^{t-2}(p-1)/r} \equiv 1 \pmod{p^{t-1}}$$

as well and this is a contradiction. This proves the first statement.

We have shown that if a is primitive modulo p^t then a is primitive modulo p, and p^2 does not divide $a^{p-1} - 1$. Now we prove the converse. If p^2 does not divide $a^{p-1} - 1$, then the only way that a can fail to be primitive modulo p^2 is if a has order modulo p^2 dividing $p(p-1)/r$ for some prime divisor of $p-1$. But as we saw in the previous paragraph, this implies that a has order modulo p dividing $(p-1)/r$, which would contradict a's primitivity modulo p. □

2.2.e Divisibility in rings

Let R be a commutative ring. If $a, b \in R$ then a is a *divisor* of b if there exists $c \in R$ such that $ac = b$, in which case we write $a|b$. The element a is a *unit* if it is invertible, or equivalently, if it is a divisor of 1. Elements $a, b \in R$ are *associates* if $a = \epsilon b$ for some unit ϵ. A nonzero element $c \in R$ is a *common divisor* of a and b if $c|a$ and $c|b$. It is a *greatest common divisor* of a and b (written $c = \gcd(a, b)$) if it is a common divisor and if every other common divisor of a and b divides c. An element $c \neq 0$ is a *common multiple* of a and b if $a|c$ and $b|c$. It is a *least common multiple*

(written $c = \operatorname{lcm}(a, b)$) if it is a common multiple and if it divides every other common multiple of a and b.

A nonzero element $r \in R$ is *prime* if (r) is a proper prime ideal, meaning that if $ab \in (r)$ then $a \in (r)$ or $b \in (r)$. It is *primary* if (r) is primary, meaning that $ab \in (r)$ implies $a \in (r)$ or $b^n \in (r)$ for some $n > 0$. It is *irreducible* if it is not a unit and if $r = ab$ implies that a or b is a unit. Two nonzero non-units $r, s \in R$ are *coprime* or *relatively prime* if $(r) + (s) = R$ or equivalently if there exist $a, b \in R$ so that $1 = ar + bs$. See also Theorem 2.2.15.

Definition 2.2.13. *Let R be a commutative ring.*

1. *R is an* integral domain *(or simply a* domain*, or* entire*) if it has no zero divisors.*
2. *R is* principal *if every ideal in R is principal. It is a* principal ideal domain *or* PID *if it is principal and is an integral domain.*
3. *R is a* GCD ring *if every pair of elements has a greatest common divisor.*
4. *R is a* local ring *if it contains a unique maximal ideal (which therefore consists of the set of all non-units).*
5. *R is a* unique factorization domain *(or* UFD*, or* factorial*) if it is an integral domain and every nonunit $a \in R$ has a factorization into a product*

$$a = \prod_{i=1}^{m} p_i \tag{2.4}$$

   of irreducible elements (not necessarily distinct), which is unique up to reordering of the p_is and multiplication of the p_is by units. That is, if $a = \prod_{i=1}^{n} q_i$, then $m = n$ and there is a permutation σ of $\{1, \cdots, m\}$ so that p_i and $q_{\sigma(i)}$ are associates.
6. *R is a* factorization ring *if every nonunit $a \in R$ has a factorization into a product of irreducible elements, not necessarily distinct, and not necessarily in a unique way. An entire factorization ring is a* factorization domain*.*
7. *R is* Noetherian *if every increasing sequence of ideals $I_1 \subset I_2 \subset \cdots$ stabilizes at some finite point, or equivalently, if every ideal is finitely generated.*
8. *R is* Euclidean *if there is a function $\delta : R \to \{0, 1, 2, \cdots\} \cup \{-\infty\}$ such that (1) for every $a, b \in R$ with a and b both nonzero, we have $\delta(ab) \geq \delta(a)$, and (2) for every $a, b \in R$ with $b \neq 0$ there exist $q \in R$ (the quotient) and $r \in R$ (the remainder)*
   so that

$$a = qb + r \quad \text{and} \quad \delta(r) < \delta(b). \tag{2.5}$$

Theorem 2.2.14 summarizes the various inclusions among the special types of rings that we have discussed. We have included the polynomial ring $R[x]$ for ease of reference although it will not be considered until Section 2.4.

Theorem 2.2.14. *Let R be a commutative ring and let $R[x]$ be the ring of polynomials with coefficients in R (see Section 2.4). Then we have the following diagram of implications between various possible properties of R.*

$$\begin{array}{ccccccccccc} \textit{field} & \Longrightarrow & \textit{Euclidean} & \Longrightarrow & \textit{PID} & \Longrightarrow & \textit{UFD} & \Longrightarrow & \textit{entire} & \Longrightarrow & R[x]\textit{entire} \\ \Downarrow & & & & & & \Downarrow & & & & \\ R[x]\textit{Euclidean} & & & & & & \textit{GCD} & & & & \end{array}$$

If R is finite and entire then it is a field. If R is an order in an algebraic number field (see Section 3.4.c) then it is entire and Noetherian. The following additional implications hold.

$$\begin{array}{ccccccc} \textit{PID} & \Longrightarrow & \textit{Noetherian} & \Longrightarrow & \textit{factorization} & & \\ & & \Downarrow & & & & \textit{factorization} \\ & & R[x]\textit{Noetherian} & \Longrightarrow & R[x]\textit{factorization} & \Longleftarrow & \textit{domain} + \textit{GCD} \end{array}$$

Proof. The properties of the polynomial ring $R[x]$ are proved in Lemma 2.4.1 and Theorem 2.4.2. If R is a field then it is Euclidean with $\delta(0) = -\infty$ and $\delta(r) = 0$ for all nonzero elements $r \in R$.

To show that every Euclidean ring is a PID, let R be Euclidean. Suppose $a \in R$ is nonzero. We can write $0 = qa + r$ with $\delta(r) < \delta(a)$. Suppose that q is nonzero. Then $\delta(r) = \delta(-qa) \geq \delta(a)$, which is a contradiction. Thus $q = 0$ so $r = 0$. But then we must have $\delta(0) < \delta(a)$ for every $a \neq 0$. In particular, $\delta(a) \geq 0$ if a is nonzero. Now let I be a nonzero ideal in R. Let $a \in I - \{0\}$ be an element such that $\delta(a)$ is minimal. There is at least one such element since $\delta(I - \{0\}) \subset \mathbb{N}$ has a least element (by the well ordering principal). We claim that $I = (a)$. Let b be any other element in I. Then $b = qa + r$ for some $q, r \in R$ such that $\delta(r) < \delta(a)$. But $r = b - qa \in I$, so $r = 0$. That is, $b = qa$, as claimed. Moreover, if $0 = ab$ for some nonzero a, then the argument above shows that $b = 0$, so R is an integral domain.

Now assume that R is a PID. If a and b are two elements of R, then the ideal (a, b) has a principal generator, $(a, b) = (c)$. Thus c divides both a and b, and $c = ua + vb$ for some $u, v \in R$. Therefore any common divisor of a and b divides c as well. That is, c is a GCD of a and b. It follows that R is a GCD ring. It also follows that the GCD c can be written in the form $c = ua + vb$.

To see that R is Noetherian, let $I_1 \subset I_2 \subset \cdots$ be an increasing chain of ideals in R. The union of the I_ns is an ideal, so it is principal,

$$\cup_{n=1}^{\infty} I_n = (a)$$

for some a. But there is a natural number n with $a \in I_n$, so the chain stabilizes at I_n.

Suppose that R is Noetherian. We prove that R is a factorization ring. That is, that every element $a \in R$ has a prime factorization. Let S be the set of nonzero elements of R that do not have prime factorizations, and suppose S is nonempty. Let $a \in S$. Then a is not prime, so we

can write $a = bc$ with neither b nor c in (a). Since a is in S, either b or c is in S. Repeating this infinitely gives a chain $(a_1) \subset (a_2) \subset \cdots$ with $a_i \in S$ and $(a_i) \neq (a_{i+1})$ for every $i \geq 1$. This contradicts the fact that R is Noetherian.

Now we return to the case when R is a PID (and hence a GCD ring and Noetherian and so a factorization ring) and prove uniqueness of factorizations. Suppose $a \in R$ is irreducible and $a|bc$. If $a \not| b$, then 1 is a gcd of a and b, so we have

$$1 = ua + vb,$$

for some $u, v \in R$. Thus $c = uac + vbc$, so $a|c$. That is, if $a|bc$, then $a|b$ or $a|c$. In other words, a is prime if a is irreducible. Suppose some nonunit $b \in R$ can be factored in two ways,

$$b = \prod_{i=1}^{k} p_i = \prod_{i=1}^{\ell} q_i.$$

Since b is not a unit, we have $k > 0$ and $\ell > 0$. We use induction on k. Since $p_k | \prod_{i=1}^{\ell} q_i$, we have $p_k | q_n$ for some n by the primality of p_k, say $q_n = dp_1$. By the irreducibility of p_k and q_n, d is a unit. Then $\prod_{i=1}^{k-1} p_i = d(\prod_{i=1}^{\ell} q_i)/q_n$, and the result follows by induction. This completes the proof that a PID R is a UFD.

The implication UFD $\implies$ GCD is straightforward. Every UFD is an integral domain by definition. This completes the first diagram of implications.

The implication (finite + entire $\implies$ field) was proven in Section 2.2.a. An order R in a number field is a free $\mathbb{Z}$ module of finite rank, so the same is true of any ideal in R, hence such an ideal is finitely generated (as an Abelian group). Thus R is Noetherian.

The proof that (Noetherian $\implies$ $R[x]$ Noetherian) is fairly long and will be omitted; it is called Hilbert's basis theorem, see any book on commutative algebra, for example [3]. The remaining results involving polynomials are proved in Section 2.4.a. □

Theorem 2.2.15. *Let R be a commutative ring and let $a, b \in R$. Then*

1. *The element a is prime if and only if it has the following property: if $a|cd$ then $a|c$ or $a|d$.*
2. *If a is prime and is not a zero divisor, then a is irreducible.*
3. *If R is a UFD, then a is prime if and only if a is irreducible.*
4. *The elements a and b are coprime if and only if (the image of) a is invertible in $R/(b)$ (if and only if the image of b is invertible in $R/(a)$).*
5. *If a and b are coprime, then every common divisor of a and b is a unit.*
6. *If R is a PID and if every common divisor of a and b is a unit, then a and b are coprime.*
7. *If R is a PID and $a \in R$, then a is prime if and only if (a) is maximal (if and only if $R/(a)$ is a field).*

Proof. Part (1) is just a restatement of the definition that (a) is a prime ideal.

Now suppose a is prime and is not a zero divisor, and suppose $a = cd$. Then either $c \in (a)$ or $d \in (a)$; we may assume the former holds. Then $c = ea$ for some $e \in R$, so $a = cd = ead$ or $a(1 - ed) = 0$. Since a is not a zero divisor, we have $ed = 1$ hence d is a unit. This proves (2).

For part (3), first suppose that $a \in R$ is irreducible and let $cd \in (a)$. Then $cd = ae$ for some element $e \in R$. The right side of this equation is part of the unique factorization of the left side, so a must divide either c or d. Therefore either $c \in (a)$ or $d \in (a)$. The converse was already proven in part (2). (Note that a UFD contains no zero divisors, due to the unique factorization of 0.)

For part (4), if a is invertible in $R/(b)$ then there exists $c \in R$ so that $ac \equiv 1 \pmod{b}$, meaning that there exists $d \in R$ so that $ac = 1 + db$. Hence $(a) + (b) = R$. The converse is similar.

For part (5), supposing a and b are coprime, we may write $1 = ac + bd$ for some $c, d \in R$. If $e|a$ and $e|b$ then $a = fe$ and $b = ge$ for some $f, g \in R$. This gives $1 = (fc + gd)e$ so e is invertible.

For part (6), Suppose R is a PID. Given a, b the ideal $(a) + (b)$ is principal, so it equals (c) for some $c \in R$, which implies that $c|a$ and $c|b$. Therefore c is a unit, so $(a) + (b) = (c) = R$.

For part (7), we have already shown, in Theorem 2.2.5 that (a) maximal implies that a is prime. For the converse, suppose that (a) is prime and that $(a) \subset (b) \neq R$. Then b is not a unit, and $a = cb$ for some $c \in R$. Since the ring R is also a UFD, the element a is irreducible, so c is a unit. Therefore $(a) = (b)$ hence (a) must be maximal. □

2.2.f Examples

Here are a few standard examples of rings.

1. The integers $\mathbb{Z}$ is a Euclidean ring with $\delta(a) = |a|$.

2. The rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$ are fields.

3. If $k = mn$ is a composite integer (with $m, n \geq 2$) then $\mathbb{Z}/k\mathbb{Z}$ is not an integral domain since $m \cdot n = 0$.

4. If R is a ring and S is a nonempty set, then the set of functions from S to R is a ring with the operations $(f + g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$. The zero is the function $z(x) = 0$ for all x, and the identity is the function $i(x) = 1$ for all x.

5. If R is a ring then the collection $R[x]$ of polynomials with coefficients in R (see Section 2.4) is a ring.

6. Let G be an Abelian group with operation $*$ and identity e. The set E of endomorphisms of G is a ring with the operations $+_E$ = "product" and $\cdot_E$ = "composition". The zero is the function $z(a) = e$ for all a, and the identity is the function $i(a) = a$ for all a.

2.2.g The Euclidean algorithm

If R is Euclidean (and hence a GCD domain) via function $\delta : R \to \{0, 1, 2, \cdots\} \cup \{-\infty\}$, then the Euclidean algorithm, given in Figure 2.1, computes the gcd of any two elements. We assume that addition and multiplication of elements in R are atomic operations and that given $a, b \in R$, we can compute $q, r \in R$ as in equation (2.5). The algorithm assumes that a and b are nonnegative.

```
EUCLID(a, b)
    begin
    while (b ≠ 0) do
        Let a = qb + r
        (a, b) = (b, r)
     od
    return(a)
    end
```

Figure 2.1: The Euclidean Algorithm.

The proof of correctness of the Euclidean algorithm is essentially the same as in the integer case, which can be found in most general texts on algorithms. The time complexity depends on the ring, and in particular on the maximum time $M(d)$ it takes to compute q and r as in equation (2.5) when $\max(\delta(a), \delta(b)) \leq d$.

If $d = \max(\delta(a), \delta(b))$ decreases by an additive constant ϵ at each stage, then the complexity is at most $O(dM(d))$. This is the case when $R = \mathbb{F}[x]$ for a finite field F and $\delta(a) = \deg(a)$. In this case M is the time required to multiply polynomials, say $M(d) \in O(d \log(d))$ using fast Fourier transforms. The resulting complexity of the Euclidean algorithm is $O(\deg(a)^2 \log(\deg(a)))$. However a better bound can be found in this case by taking into account the degrees of the intermediate quotients. Two degree d polynomials can be divided in time $O(d(e + 1))$, where e is the degree of the quotient. Suppose that the sequence of polynomials produced by the algorithm is $r_0 = a, r_1 = b, r_2, \cdots, r_n$. Then $n \leq d$. If r_i has degree d_i, then the ith quotient has degree at most $d_{i-1} - d_i$. Thus the complexity is in

$$O(\sum_{i=0}^{d} d_{i-1}(d_{i-1} - d_i + 1)) \in O(d \sum_{i=1}^{d} (d_{i-1} - d_i + 1)) = O(d(d_0 - d_n + d)) = O(d^2).$$

If for some constant $\epsilon < 1$, $\delta(a)$ is decreased by a factor of ϵ after a constant number c of steps, then a simple bound on the complexity is $O(\log(d)M(d))$. This is the case when $R = \mathbb{Z}$

and $\delta(a) = |a|$. However a better bound can be found in this case by taking into account the actual numbers involved. Two k-bit numbers can be divided in time $O(k(\ell+1))$, where ℓ is the number of bits in the quotient. Suppose that the sequence of numbers produced by the algorithm is $r_0 = a, r_1 = b, r_2, \cdots, r_n$. Then $n \leq k$. If r_i has k_i bits, then the ith quotient has at most $k_{i-1} - k_i + 1$ bits. Thus the complexity is in

$$O(\sum_{i=1}^{k} k_{i-1}(k_{i-1} - k_i + 2)) \in O(k \sum_{i=1}^{k}(k_{i-1} - k_i + 2)) = O(k(k_0 - k_n + 2k)) = O(k^2) = O(\log(d)^2).$$

Theorem 2.2.16. *If $R = \mathbb{F}[x]$ for a finite field F, then the complexity of the Euclidean algorithm on inputs of degree $\leq d$ is in $O(d^2)$. If $R = \mathbb{Z}$, the the complexity of the Euclidean algorithm on inputs of size $\leq d$ is in $O(\log(d)^2)$.*

If R is a Euclidean domain, then (by Theorem 2.2.14) it is also a PID. If $a, b \in R$, then the ideal generated by a and b has a generator c. As in the proof of Theorem 2.2.14, $c = \gcd(a, b)$ and there are elements $u, v \in R$ so that $c = ua + vb$. The elements u and v are sometimes called *Bézout coefficients* It turns out that with a simple modification, the Euclidean algorithm can be used to compute the Bézout coefficients. This can be described by keeping track of all the intermediate information in the computation of the Euclidean algorithm:

$$r_0 = a, u_0 = 1, v_0 = 0;$$

$$r_1 = b, u_1 = 0, v_1 = 1;$$

and for $i \geq 1$

$$r_{i+1} = r_{i-1} - q_i r_i, u_{i+1} = u_{i-1} - q_i u_i, v_{i+1} = v_{i-1} - q_i v_i.$$

The sequence halts with $i = n$ so that $r_n = 0$. The sequence $(u_i, v_i, r_i), i = 0, 1, \cdots, n$ is called the *Bézout sequence of a and b*. We have the following facts which we will make use of in Section 19.3.b on rational approximation of N-adic numbers using the Euclidean algorithm.

Lemma 2.2.17.

1. $r_1 > r_2 > \cdots > r_{n-1} = \gcd(a, b) \geq 0$.

2. *For $0 \leq i \leq n$,*
$$u_i a + v_i b = r_i. \tag{2.6}$$

3. *For $0 \leq i \leq n-1$,*
$$u_i v_{i+1} - u_{i+1} v_i = (-1)^i. \tag{2.7}$$

4. *If i is even then $u_i \geq 0$ and $v_i \geq 0$. If is odd then $u_i \geq 0$ and $v_i \geq 0$.*

5. *$|u_1| < |u_2| < \cdots |u_n|$ and $|v_0| < |v_1| < \cdots |v_n|$.*

6. *For $0 \leq i \leq n-1$, $|u_{i+1} r_i| \leq b$, $|v_{i+1} r_i| \leq a$, $|u_i r_{i+1}| \leq b$, and $|v_i r_{i+1}| \leq a$.*

2.2.h Fractions

The field of rational numbers $\mathbb{Q}$ is constructed from the ring of integers $\mathbb{Z}$ as the set of all fractions a/b, where we identify a/b with $(ax)/(bx)$ for any nonzero integer x. A similar construction can be made in great generality. Let R be a commutative ring. A subset S of R is *multiplicative* if it is closed under multiplication. If S is any multiplicative subset of R, define the ring $S^{-1}R$ to be the collection of all formal symbols a/b (where $a \in R$ and $b \in S$), under the following equivalence relation: $a/b \sim a'/b'$ if there exists $s \in S$ such that

$$s(ab' - ba') = 0. \tag{2.8}$$

Addition and multiplication of fractions are defined by the usual formulas:

$$\frac{a}{b} + \frac{a'}{b'} = \frac{ab' + a'b}{bb'} \quad \text{and} \quad \frac{a}{b}\frac{a'}{b'} = \frac{aa'}{bb'}.$$

The ring $S^{-1}R$ consists of a single element if $0 \in S$, so sometimes this case is excluded.

Now suppose that S does not contain any zero divisors (which will always be the case in our applications). Then equation (2.8) may be replaced by the more familiar equivalence relation: $ab' = a'b$. The natural mapping $R \to S^{-1}R$ (which takes a to $a/1$) is an injection, so $S^{-1}R$ "contains" R. Every element of S has become invertible in $S^{-1}R$. If the set S consists of all the elements that are not zero divisors, then an element of $S^{-1}R$ is either a zero divisor or else it is invertible. In this case, the ring $S^{-1}R$ is called the (full) *ring of fractions* of R. If R is an integral domain then its ring of fractions is a field, which is called the *fraction field* of R. See for example, Section 3.4.7, Section 5.2 and Section 5.4.

2.2.i Chinese remainder theorem

If R_1 and R_2 are rings then their Cartesian product $R_1 \times R_2$ is a ring under the coordinate-wise operations of addition and multiplication. The Chinese remainder theorem gives a sufficient condition under which a ring may be decomposed as a product.

Theorem 2.2.18. *Let R be a ring and let $I_1, \cdots, I_k$ be ideals such that $I_i + I_j = R$ for every $i \neq j$. Let*

$$I = \cap_{j=1}^{k} I_j.$$

Then for every $a_i, \cdots, a_k \in R$ there is an element $a \in R$ such that for every i, $a \equiv a_i \pmod{I_i}$. The element a is unique modulo I. Furthermore,

$$R/I \cong \prod_{j=0}^{k} R/I_j.$$

Proof. For $k = 1$ the statement is trivial. If $k = 2$, then there are elements $b_1 \in I_1$ and $b_2 \in I_2$ so that $1 = b_1 + b_2$. Let $a = a_1 b_2 + a_2 b_1$.

Now suppose $k > 2$. For every i let

$$J_i = \prod_{j \neq i} I_j.$$

For every $i \geq 2$ there are elements $c_i \in I_1$ and $b_i \in I_i$ such that $1 = c_i + b_i$. In particular,

$$\prod_{i=2}^{k} (c_i + b_i) = 1.$$

This product is in $I_1 + J_1$, so $R = I_1 + J_1$. Similarly, $R = I_j + J_j$ for every j. By the theorem in the case of two ideals, there is an element $d_j \in R$ such that $d_j \equiv 1 \pmod{I_j}$ and $d_j \equiv 0 \pmod{J_j}$. Then $a = a_1 d_1 + \cdots + a_k d_k$ satisfies our requirements

For each i there is a reduction homomorphism φ_i from R/I to R/I_i. This induces a homomorphism φ from R/I to

$$\prod_{j=1}^{k} R/I_j$$

whose kernel is

$$I = \bigcap_{j=1}^{k} I_j.$$

Thus φ is injective. By the first part it is surjective, hence an isomorphism. This also proves the uniqueness of a. □

Corollary 2.2.19. *Suppose R is a PID and $b_1, \cdots, b_k \in R$ are pairwise relatively prime. If $a_1, \cdots, a_k \in R$, then there exists an element $a \in R$ such that for every i, $a \equiv a_i \pmod{b_i}$.*

Proof. By Theorem 2.2.18 it suffices to show that for each $i \neq j$ we have $(b_i) + (b_j) = R$. The set $(b_i) + (b_j)$ is an ideal. Since R is a PID, there is some $b \in R$ so that $(b_i) + (b_j) = (b)$. This says that b is a common divisor of b_i and b_j, so b is a unit by assumption. Thus $(b_i) + (b_j) = R$. □

By Theorem 2.2.14, Corollary 2.2.19 applies in particular when R is a Euclidean domain. The case when $R = \mathbb{Z}$ is the classical Chinese Remainder Theorem.

2.2.j Vector spaces

In many settings we have a notion of one algebraic object "acting on" another by multiplication. For example, a real number r acts on the set of points in the plane by $(x, y) \mapsto (rx, ry)$.

Definition 2.2.20. *A* vector space *over a field F is a set V such that V is an Abelian group with an operation $+$, and there is a function $\cdot$ from $F \times V$ to V such that for all $a, b \in F$ and $u, v \in V$*

1. $a \cdot (u + v) = (a \cdot u) + (a \cdot v)$;
2. $(ab) \cdot u = a \cdot (b \cdot u)$;
3. $(a + b) \cdot u = (a \cdot u) + (b \cdot u)$; *and*
4. $1 \cdot u = u$.

It follows from these axioms that for every $u \in V$, $0 \cdot u = 0$.

For example, the set of points in the real plane is a vector space over the real numbers. If F is a field which is a subring of a ring R, then R is a vector space over F (just use the multiplication in R for the action of F on R). If F is a field and S is a nonempty set, then the set of functions from S to F is a vector space over F with the operations $(f+g)(x) = f(x)+g(x)$ and $(a \cdot f)(x) = af(x)$ for $a \in F$, $x \in S$, and $f, g : S \to F$. Various restrictions can be put on the functions to produce interesting vector spaces (e.g., continuity if $S = F = \mathbb{R}$).

Let V be a vector space over a field F. The elements of V are called *vectors*. A *linear combination* of vectors $v_1, v_2, \cdots, v_k \in V$ is a vector $a_1v_1 + a_2v_2 + \cdots + a_kv_k$ with $a_1, a_2, \cdots, a_k \in F$. A set of vectors $S \subseteq V$ is *linearly independent* if the only linear combination of elements of S that is zero is the one with all the coefficients a_i equal to zero. S *spans* V if every vector can be written as a linear combination of elements of S. S is a *basis* for V if it spans V and is linearly independent. The proof of the following is an exercise.

Theorem 2.2.21. *Let V be a vector space over a field F. If V has more than one element then it has a nonempty basis. If S is a basis, then every vector can be written uniquely as a linear combination of elements of S.*

If V has a basis S with a finite number of elements, then we say V is *finite dimensional with dimension* $= |S|$. In this case it can be shown that every basis has the same number of elements. In the important case when F is a subfield of a field E, E is called an *extension field*. If E is finite dimensional as a vector space over F, then its dimension is called *the degree of the extension* and is denoted $[E : F]$.

Theorem 2.2.22. *If F is a finite field and V is a finite dimensional vector space over F with dimension d, then $|V| = |F|^d$.*

Proof. Let S be a basis for V. Thus $|S| = d$. That is $S = \{v_1, v_2, \cdots, v_d\}$ for some $v_1, v_2, \cdots, v_d$. By the previous theorem, the elements of V are in one-to-one correspondence with the linear combinations $\sum_{i=1}^{d} a_i v_i$, $a_i \in F$. There are exactly $|F|^d$ such linear combinations. □

Definition 2.2.23. *If F is a field and V and W are vector spaces over F, then a function $L : V \to W$ is a* homomorphism *or is F-*linear *if it is a group homomorphism and for all $a \in F$ and $v \in V$ we have $L(av) = aL(v)$.*

If $S = \{v_1, v_2, \cdots, v_d\}$ is a basis for V, then an F-linear function L is completely determined by its values on the elements of S since

$$L\left(\sum_{i=1}^{d} a_i v_i\right) = \sum_{i=1}^{d} a_i L(v_i).$$

On the other hand, any choice of values for the $L(v_i)$ determines an F-linear function L. Furthermore, if $T = \{w_1, w_2, \cdots, w_e\}$ is a basis for W, then each value $L(v_i)$ can be expressed as a linear combination

$$L(v_i) = \sum_{j=1}^{e} b_{ij} w_j$$

with $b_{ij} \in F$.

Theorem 2.2.24. *If F is finite and V and W are finite dimensional with dimensions d and e, respectively, then there are $|F|^{de}$ F-linear functions from V to W.*

The image and kernel of L are Abelian groups, and it is straightforward to check that they are also vector spaces over F. Their dimensions are called the *rank* and *co-rank* of L, respectively. We leave it as an exercise to show that the rank plus the co-rank equals the dimension of V.

We can identify an element $\sum_i a_i v_i \in V$ with the column vector $(a_1, \cdots, a_d)^t$ (where the superscript t denotes the transpose of a matrix), and similarly for an element of W. Then the linear function L is identified with ordinary matrix multiplication by the matrix $B = [b_{ij}]$. The rank of L is the size of a maximal set of independent columns or independent rows of B.

If B is a square matrix, then the determinant of B is defined as usual in linear algebra. In this case the kernel is nonempty if and only if the determinant is zero.

If V and W are vector spaces over a field F, then the set of F-linear homomorphisms from V to W is denoted $\mathrm{Hom_F(V, W)}$. It is again a vector space over F with F acting by $(a \cdot L)(v) = L(av)$. By Theorem 2.2.24, if V and W are finite dimensional then the dimension of $\mathrm{Hom_F(V, W)}$ is the product of the dimensions of V and W.

In the special case when $W = F$, the dimension of $\mathrm{Hom_F(V, F)}$ is the same as that of V, so $\mathrm{Hom_F(V, F)}$ and are isomorphic as vector spaces over F (but not canonically – an isomorphism depends on a choice of bases). $\mathrm{Hom_F(V, F)}$ is called the *dual space* of V.

2.2.k Modules and lattices

The notion of a vector space over a field can be generalized to rings.

Definition 2.2.25. *Let $(R, +, \cdot, 0, 1)$ be a commutative ring. A* module over R *is an Abelian group $(M, +, 0_M)$ with an operation $\cdot$ from $R \times M$ to M such that for all $a, b \in R$ and $u, v \in M$*

1. $a \cdot (u + v) = (a \cdot u) + (a \cdot v)$;
2. $(ab) \cdot u = a \cdot (b \cdot u)$;
3. $(a + b) \cdot u = (a \cdot u) + (b \cdot u)$; *and*
4. $1 \cdot u = u$.

Again, it follows from these axioms that for every $u \in V$, $0 \cdot u = 0$.

For example, every Abelian group is a module over the integers (if $n \in \mathbb{Z}^+$, then $n \cdot a$ equals the sum of n copies of a). If f is a homomorphism from a ring R to a ring S, then S is a module over R with the operation $a \cdot u = f(a)u$.

It is apparent that the notion of basis does not make sense for modules in general – even a single element of a module may not be linearly independent. However, if there is a finite set of elements $m_1, \cdots, m_k \in M$ such that every element of M can be written (perhaps not uniquely) as a linear combination $a_1 m_1 + \cdots + a_k m_k$ with $a_1, \cdots, a_k \in R$, then we say that M is *finitely generated over* R. If M is finitely generated, then the size of the smallest set of generators for M over R is called the *R-rank* or simply the rank of M.

A module M over a ring R is *free* if M is isomorphic to the Cartesian product of a finite number of copies of R. That is, M is free if there are elements $m_1, \cdots, m_k \in M$ such that every element $m \in M$ can be represented uniquely in the form

$$m = \sum_{i=1}^{k} c_i m_i, \quad c_i \in R.$$

In this case the set $m_1, \cdots, m_k$ is called a basis of M over R.

Definition 2.2.26. *A* lattice L *is the set of integer linear combinations of a collection $U = \{u_1, \cdots, u_k\}$ of $\mathbb{R}$-linearly independent vectors in $\mathbb{R}^n$. The set U is called a* basis *for L. The lattice L is* full *if $k = n$, which we now assume. Then M_U is defined to be the matrix whose rows are $u_1, u_2, \cdots, u_n$. The volume of the parallelepiped spanned by these vectors is denoted $D_U = |\det(M_U)|$, and it is referred to as the* volume *of the lattice L, or the* determinant *of L.*

It is immediate that a lattice is a free $\mathbb{Z}$-module. A basis for a full lattice L is also a basis for $\mathbb{R}^n$. A full 2-dimensional lattice with basis $(5, 1)$, $(3, 4)$ is shown in Figure 2.2.k. The following theorem says that $\mathrm{vol}(L)$ is well-defined and it gives a way to tell when a collection of vectors forms a basis of a given lattice, cf. Chapter 19.

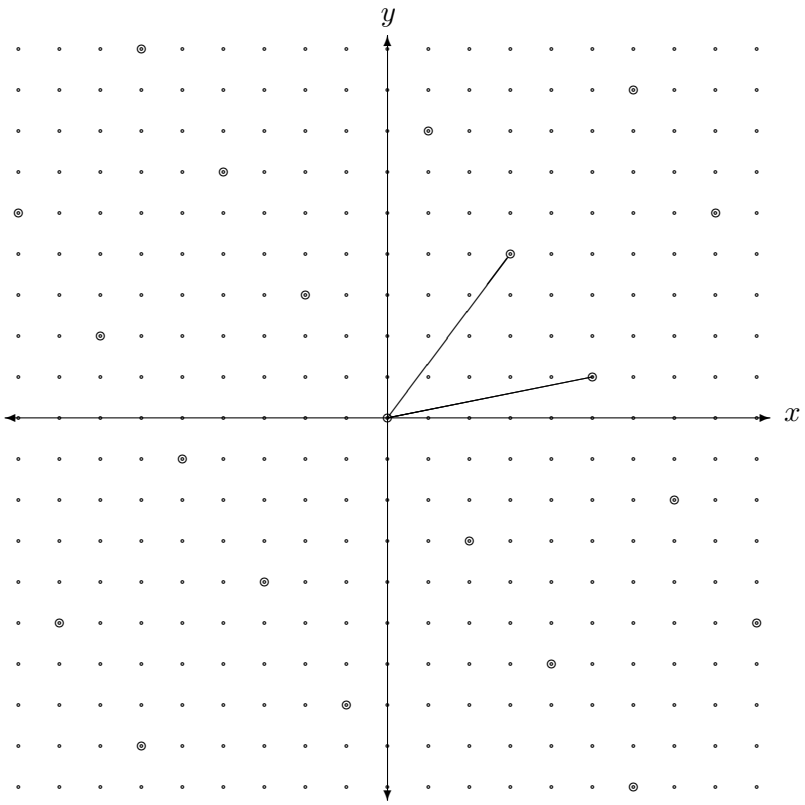


Figure 2.2: A lattice with basis $(5, 1)$, $(3, 4)$

Theorem 2.2.27. *Let L be a full lattice in $\mathbb{R}^n$. Then $\mathrm{vol}(L)$ is independent of the choice of basis. If $V = \{v_1, \cdots, v_n\} \subset L$ is any linearly independent set of vectors in L, then $D_V \neq 0$ and it is an integral multiple of $\mathrm{vol}(L)$. Moreover, $D_V = \mathrm{vol}(L)$ if and only if V is a basis of L.*

Proof. Let $U = \{u_1, \cdots, u_n\}$ be a basis of L. Then each v_i can be written as an integer linear combination of the u_j. This gives a matrix S with integer entries such that $M_V = SM_U$. The determinant of S is an integer so $D_V = |\det(S)|D_U$ is an integral multiple of D_U. If V is also a basis of L then we similarly obtain a matrix T with integer entries such that $M_U = TM_V$. This implies $TS = I$ so the determinants of S and T are integers with integer inverses, hence $|\det(S)| = 1$ and $D_U = D_V$. This proves the first two statements.

For the last statement, suppose $U \subset L$ and $V \subset L$ are collections of n linearly independent vectors, suppose U is a basis of L, and suppose $D_V = D_U$. We claim that V is also a basis of L. As above, write $V = SU$ where S is a matrix of integers. Then $\det(V) = \det(S)\det(U)$ so

$\det(S) = \pm 1$. By Cramer's rule, the inverse of S consists of rational numbers whose denominators are $\det(S)$, so S^{-1} also has integer entries. Hence, the equation $U = S^{-1}V$ expresses the u_i as integer linear combinations of the v_j, so V is also a basis for L. □

The proof of the following fact about lattice bases may be found in [17] Lemma 1, Section 2.6.

Theorem 2.2.28. *Let $L \subset K \subset \mathbb{R}^n$ be full lattices. Then K/L is a finite Abelian group. Let $u_1, \cdots, u_n$ and $v_1, \cdots, v_n$ be bases of L and K respectively. Each u_i is an integer linear combination of the vectors v_i, say, $u_i = A_{i1}v_1 + \cdots + A_{in}v_n$. Then the matrix $A = (A_{ij})$ has integer entries and $|K/L| = |\det(A)|$.* □

In a lattice, linear dependence over $\mathbb{R}$ implies linear dependence over $\mathbb{Z}$.

Lemma 2.2.29. *Let $u_1, \cdots, u_k$ be a set of vectors in $\mathbb{R}^n$ that is linearly independent over $\mathbb{R}$. Let v be a vector in the $\mathbb{Z}$-span of $u_1, \cdots, u_k$ and suppose that v is in the $\mathbb{R}$-span of $u_1, \cdots, u_\ell$ with $\ell \leq k$. Then v is in the $\mathbb{Z}$-span of $u_1, \cdots, u_\ell$*

Proof. Write $v = a_1u_1 + \cdots + a_ku_k$ with each $a_i \in \mathbb{Z}$, and $v = b_1u_1 + \cdots + b_\ell u_\ell$ with each $b_i \in \mathbb{R}$. By the uniqueness of the representation of a vector as a linear combination of a set of linearly independent vectors over a field, we have $a_1 = b_1, a_2 = b_2, \cdots, a_\ell = b_\ell$ and $a_{\ell+1} = \cdots = a_k = 0$. □

For any positive real number r, let $B_r(x) \subset \mathbb{R}^n$ denote the (closed) ball of radius r, centered at $x \in \mathbb{R}^n$. A subset $L \subset \mathbb{R}^n$ is *discrete* if for every x, r the set $B_r(x) \cap L$ is finite.

Theorem 2.2.30. *Every lattice $L \in \mathbb{R}^n$ is discrete.*

Proof. Any lattice is contained in a full lattice, so we may assume L is full, say with basis $u_1, \cdots, u_n$. Define $f : \mathbb{R}^n \to \mathbb{R}^n$ by $f(a_1u_1 + \cdots + a_nu_n) = (a_1, a_2, \cdots, a_n)$. Then f maps the lattice L isomorphically to the standard lattice $\mathbb{Z}^n \subset \mathbb{R}^n$ consisting of vectors with integer coordinates. For any x, r the image $f(B_r(x))$ is compact, hence closed and bounded, so it is contained in some n-cube with integer vertices and with edges of some (possibly very large) integer length, D. Such a cube contains $(D+1)^n$ integer vertices. Therefore $B_r(x)$ contains no more than $(D+1)^n$ lattice points in L. □

We leave as an exercise the proof that every discrete $\mathbb{Z}$-module in $\mathbb{R}^n$ is a lattice. This provides an alternate characterization of lattices that is sometimes used in the literature as definition. Although we will not need to use it, we state for completeness the following theorem of Minkowski,

Theorem 2.2.31. *Let $L \subset \mathbb{R}^n$ be a full lattice and let $X \subset \mathbb{R}^n$ be a bounded convex subset that is centrally symmetric. If $\mathrm{vol}(X) > 2^n \mathrm{vol}(L)$ then X contains a nonzero element of L.*

Sometimes a module M over a ring R has the structure of a commutative ring. If the function $a \mapsto a \cdot 1_M$ is a ring homomorphism, then we say that M is a (commutative) *S-algebra*. For example, every commutative ring is a $\mathbb{Z}$-algebra. If R is a subring of a ring R', then R' is an R-algebra. If R is commutative ring and S is a multiplicative set in R, then $S^{-1}R$ is an R-algebra. More generally, if I is an ideal of R and R/I is a subring of a ring R', then R' is an R-algebra.

2.2.l Inverse limits

The notions of directed system and inverse limit provide a powerful mechanism for studying infinite sequences.

Definition 2.2.32. *Let R be a ring and let $(P, \prec)$ be a partially ordered set. A* directed system of modules over R indexed by P *is a set of modules $\{M_r : r \in P\}$ and, for each pair $p, q \in P$ with $p \prec q$, a homomorphism $\mu_{q,p} : M_q \to M_p$. If $p \prec q \prec r$, then we must have $\mu_{r,p} = \mu_{q,p} \circ \mu_{r,q}$.*

If $\{M_r : r \in P\}$ is a directed system of modules over R indexed by P, then let the inverse limit of the system be

$$\varprojlim M_r = \varprojlim \{M_r : r \in P\} = \left\{ z \in \prod_{r \in P} M_r : \text{ if } p \prec q, \text{ then } \mu_{q,p}(z_q) = z_p \right\}.$$

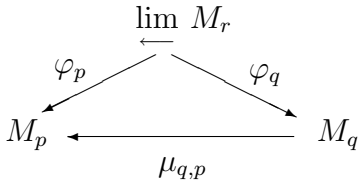
Here z_p denotes the pth component of $z \in \prod_{r \in P} M_r$.

Theorem 2.2.33. *The set $\varprojlim M_r$ is a module. For each $q \in P$ there is a homomorphism*

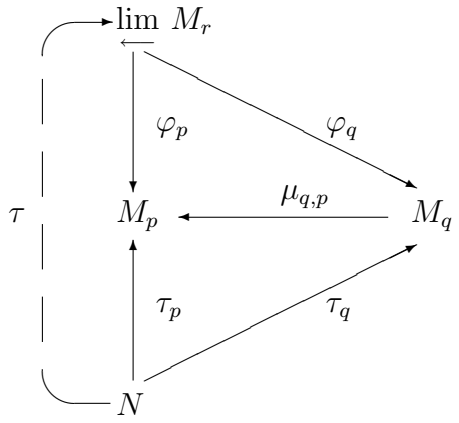
$$\varphi_q : \varprojlim M_r \to M_p$$

so that $\varphi_p = \mu_{q,p} \circ \varphi_q$ whenever $p \prec q$.

That is, the following diagram commutes.



If N is any R-module and $\{\tau_p : p \in P\}$ is a set of homomorphisms such that $\tau_p = \mu_{q,p} \circ \tau_q$ whenever $p \prec q$, then there is a unique homomorphism $\tau : N \to \varprojlim M_r$ so that $\tau_p = \varphi_p \circ \tau$. That is, the following diagram can be completed to a commutative diagram.



If also the M_p's are R-algebras and the homomorphisms $\mu_{q,p}$ are R-algebra homomorphisms, then $\varprojlim M_r$ is an R-algebra

Proof. Any Cartesian product $\prod_{r\in P} M_r$ of R-modules is an R-module, and $\varprojlim M_r$ is a subset that is closed under addition and scalar multiplication, so is also an R-module. The function φ_p is simply the restriction of the projection on M_p to $\varprojlim M_r$. The commutativity of the first diagram follows from the constraint on the elements of $\varprojlim M_r$.

If N and $\{\tau_p\}$ are as in the second condition and $a \in N$, then we define $\tau(a)$ to be the element of $\prod_{r\in P} M_r$ whose rth component is $\tau_r(a)$. That $\tau(a) \in \varprojlim M_r$ follows from the commutativity of the τ_r and the $\mu_{q,p}$. It is immediate that the second diagram commutes and that τ is unique.

The extension to R-algebras is straightforward. □

In the language of category theory, $\varprojlim M_r$ is a universal object for the directed system $\{M_r : r \in P\}$. This theorem often allows simple proofs that certain rings defined by different infinite constructions are isomorphic.

2.3 Characters and Fourier transforms

The Fourier transform can be defined in tremendous generality. In this section we describe the main properties of the Fourier transform for finite Abelian groups.

2.3.a Basic properties of characters

Definition 2.3.1. *A (complex)* character *of an Abelian group G is a group homomorphism from G to the multiplicative group $\mathbb{C}^\times = \mathbb{C} - \{0\}$ of the complex numbers. That is, it is a function $\chi : G \to \mathbb{C}$ such that $\chi(a+b) = \chi(a)\chi(b)$ for all $a, b \in G$. Such a character is* nontrivial *if $\chi(a) \neq 1$ for some a. The trivial character is denoted 1, and the collection of all characters of G is denoted $\widehat{G}$.*

The group operation in an Abelian group is usually denoted "+", and this can lead to some confusion since a character takes values in a multiplicative group. In particular, if χ is a character of G then $\chi(mg) = \chi(g)^m$ (for any integer m), and $\chi(0) = 1$. For example, if $G = \mathbb{Z}/(2)$ then there is a unique nontrivial character χ and it converts $\{0,1\}$ sequences into $\{\pm 1\}$ sequences. If G is a finite Abelian group then $|\chi(g)| = 1$ for all $g \in G$ (since $\chi(g)^{|G|} = 1$) so χ takes values in the set $\mu_{|G|}$ of roots of unity. It follows that $\chi(-g) = \overline{\chi}(g)$ (complex conjugate) for all $g \in G$.

The set of characters $\widehat{G}$ of a group G is itself a group with group operation defined by

$$(\chi_1 \cdot \chi_2)(a) = \chi_1(a)\chi_2(a)$$

and with the trivial character as identity. If $G = \mathbb{Z}/(N)$ is the additive group of integers modulo N then the group $\widehat{G}$ of characters is also cyclic and is generated by the primitive character $\chi(k) = e^{2\pi i k/N}$ for $k \in \mathbb{Z}/(N)$. If $G = G_1 \times G_2$ is a product of two groups then $\widehat{G} = \widehat{G}_1 \times \widehat{G}_2$. Specifically, if χ is a character of G then there are unique characters χ_1, χ_2 of G_1, G_2 (respectively) such that $\chi(g_1, g_2) = \chi_1(g_1)\chi_2(g_2)$, namely $\chi_1(g_1) = \chi(g_1, 1)$ and $\chi_2(g_2) = \chi(1, g_2)$ (for any $g_1 \in G_1$ and $g_2 \in G_2$). From this, together with the fundamental theorem for finite Abelian groups 2.1.16, it follows that the collection $\widehat{G}$ of characters of a finite Abelian group G is itself a finite Abelian group which is isomorphic to G. (The corresponding statement for infinite Abelian groups is false: for example, any nonzero $x \in \mathbb{C}$ defines a character of the integers $\mathbb{Z}$ by setting $\chi(m) = x^m$.)

Proposition 2.3.2. *Let G be a finite Abelian group, let $\chi : G \to \mathbb{C}^\times$ be a character, and let $g \in G$. Then*

$$\sum_{h \in G} \chi(h) = \begin{cases} 0 & \text{if } \chi \neq 1 \\ |G| & \text{if } \chi = 1 \end{cases} \tag{2.9}$$

and

$$\sum_{\psi \in \widehat{G}} \psi(g) = \begin{cases} 0 & \text{if } g \neq 0 \\ |G| & \text{if } g = 0. \end{cases} \tag{2.10}$$

Proof. If χ is nontrivial, there exists $a \in G$ with $\chi(a) \neq 1$. Then

$$\chi(a) \sum_{h \in G} \chi(h) = \sum_{h \in G} \chi(ah) = \sum_{h' \in G} \chi(h')$$

so

$$(1-\chi(a))\sum_{h\in G}\chi(g)=0.$$

For the second statement, note that g determines a character ψ_g of $\widehat{G}$ by the equation $\psi_g(\chi)=\chi(g)$. This character is nontrivial precisely when $g\neq 0$. In this case, the sum is $\sum_{\chi\in\widehat{G}}\psi_g(\chi)$, which is zero by the first part of the lemma. □

Corollary 2.3.3. *If G is a finite Abelian group and if $g,h\in G$ with $g\neq h$, then there exists a character χ such that $\chi(g)\neq\chi(h)$.*

Proof. If $\chi(g-h)=1$ for every $\chi\in\widehat{G}$, then summing over all characters gives $|G|$. By equation (2.10) we conclude that $g-h=0$. □

Corollary 2.3.4. *(Orthogonality relations) If G is a finite Abelian group and if $\psi,\chi\in\widehat{G}$ are distinct characters then*

$$\sum_{g\in G}\psi(g)\overline{\chi}(g)=0. \tag{2.11}$$

If $g,h\in G$ are distinct elements then

$$\sum_{\chi\in\widehat{G}}\chi(g)\overline{\chi}(h)=0. \tag{2.12}$$

Proof. The first equation follows by applying Proposition 2.3.2 to the character $\psi\chi^{-1}$. The second equation is $\sum_\chi\chi(g-h)=0$, also by Proposition 2.3.2. □

2.3.b Fourier transform

Let G be a finite Abelian group and $f:G\to\mathbb{C}$ be a function. Its *Fourier transform* $\widehat{f}:\widehat{G}\to\mathbb{C}$ is defined by

$$\widehat{f}(\chi)=\sum_{g\in G}\chi(g)f(g).$$

There are three standard properties of the Fourier tranform. First, the *inversion formula*

$$f(g)=\frac{1}{|G|}\sum_{\chi\in\widehat{G}}\widehat{f}(\chi)\overline{\chi}(g) \tag{2.13}$$

expresses an arbitrary function f as a linear combination of characters, so in particular, the characters span the group $\mathbb{C}[G]$ of complex-valued functions on G. Equation (2.13) follows immediately

from the orthogonality relation for characters, for the sum on the right hand side is

$$\frac{1}{|G|}\sum_{\chi\in\widehat{G}}\sum_{h\in G} f(h)\chi(h)\overline{\chi}(g) = \frac{1}{|G|}\sum_{h\in G} f(h)\sum_{\chi\in\widehat{G}}\chi(h-g) = f(g)$$

by equation (2.10). Second, the *convolution formula*

$$\widehat{f}\cdot\widehat{h} = \widehat{f*h} \tag{2.14}$$

expresses the product of $\widehat{f},\widehat{h}$ as the Fourier transform of the *convolution*

$$(f*h)(y) = \sum_{g\in G} f(g)h(y-g).$$

Finally, *Parseval's formula* says that for any function $f: G\to\mathbb{C}$,

$$|G|\sum_{g\in G}|f(g)|^2 = \sum_{\chi\in\widehat{G}}|\widehat{f}(\chi)|^2. \tag{2.15}$$

To see this, multiply $\widehat{f}(\chi)=\sum_g \chi(g)f(g)$ by its complex conjugate, $\sum_h \overline{\chi}(h)\overline{f}(h)$ to get

$$\sum_{\chi}|\widehat{f}(\chi)|^2 = \sum_{\chi}\sum_{g}\sum_{h} f(g)\overline{f}(h)\chi(g)\overline{\chi}(h) = \sum_{g,h} f(g)\overline{f}(h)\sum_{\chi}\chi(g)\overline{\chi}(h).$$

The inner sum vanishes unless $g=h$, which leaves $|G|\sum_g f(g)\overline{f}(g)$ as claimed.

If $G\cong\mathbb{Z}/(N)$ is a cyclic group then a choice $\zeta\in\mathbb{C}$ of primitive N-th root of unity determines an isomorphism $G\cong\widehat{G}$ which takes 1 to the character χ_1 with $\chi_1(k)=\zeta^k$. The other nontrivial characters χ_m are powers of this: $\chi_m(k)=\zeta^{mk}$. If $f: G\to\mathbb{C}$ is a function, its Fourier transform $\widehat{f}$ may be considered as a function $\widehat{f}: G\to\mathbb{C}$ by writing $\widehat{f}(m)$ rather than $\widehat{f}(\chi_m)$. Thus

$$\widehat{f}(m) = \sum_{k=0}^{N-1}\zeta^{mk}f(k). \tag{2.16}$$

Finally we remark that throughout this section, it is possible to replace the complex numbers $\mathbb{C}$ with any field K, provided K contains $|G|$ distinct solutions to the equation $x^{|G|}=1$. No changes to any of the proofs are needed; see Section 3.2.h. The resulting function $\widehat{f}$ is defined on all K-valued characters $\chi: G\to K^\times$. If K is a finite field then $\widehat{f}$ is called the *discrete Fourier transform*.

A generalized discrete Fourier transform, applicable when $x^{|G|}-1$ has repeated roots in K, is described in Section 18.4.b. Applications of the Fourier transform appear in Sections 13.4 and 18.3.

2.4 Polynomials

In this section we describe some of the basic properties of the ring of polynomials. The polynomial ring is among the most fundamental algebraic constructions. It is needed for much of the analysis of shift register sequences.

2.4.a Polynomials over a ring

Throughout this section R denotes a commutative ring. A *polynomial over* R is an expression

$$f = f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d = \sum_{i=0}^{d} a_i x^i$$

where $a_0, a_1, \cdots, a_d \in R$ and x is an indeterminate. The a_i are called the *coefficients of* R. When writing polynomials we may omit terms whose coefficients equal zero. We may also write the terms in a different order. If $a_d \neq 0$, then we say that $f(x)$ has degree $d = \deg(f(x))$. In this case a_d is called the *leading coefficient* of $f(x)$. We say $\deg(0) = -\infty$. If $\deg(f(x)) = 0$ then $f(x)$ is a *constant polynomial*. If $a_d = 1$ then $f(x)$ is *monic*. The term a_0 is called the *constant term*. The value of $f(x)$ at an element $b \in R$ is $f(a) = \sum_{i=0}^{d} a_i b^i$. An element $a \in R$ is a *root of* $f(x)$ if $f(a) = 0$. If $g(x) = \sum_{i=0}^{e} b_i x^i$ is a second polynomial over R, then we define

$$(f+g)(x) = f(x) + g(x) = \sum_{i=0}^{\max(d,e)} (a_i + b_i) x^i$$

(where we may have to extend one of the polynomials with zero coefficients so that this makes sense) and

$$(fg)(x) = f(x)g(x) = \sum_{i=0}^{d+e} \left(\sum_{j=\max(0,i-e)}^{\min(d,i)} a_j b_{i-j} \right) x^i.$$

The set of polynomials over R is denoted $R[x]$. The operations of addition and multiplication make $R[x]$ into a ring whose zero is the polynomial with every $a_i = 0$, and whose identity is the polynomial with $a_0 = 1$ and $a_i = 0$ for $i \geq 1$. The proof of the following lemma is straightforward.

Lemma 2.4.1. *If* $f(x), g(x) \in R[x]$, *then* $\deg(f+g) \leq \max(\deg(f), \deg(g))$ *with equality if* $\deg(f) \neq \deg(g)$. *Also,* $\deg(fg) \leq \deg(f) + \deg(g)$, *and equality can fail only when the product of the leading coefficients of* f *and* g *equals zero. In particular, if* R *is an integral domain then so is* $R[x]$.

If R is an integral domain, then the units in $R[x]$ are exactly the polynomials with degree zero and whose constant terms are units of R. This is false in general. For example, if $R = \mathbb{Z}/(4)$, then $(1+2x)^2 = 1$, so $1+2x$ is a unit with degree one. The following result says that sometimes we can perform division with remainder in $R[x]$.

Theorem 2.4.2. *(Division Theorem for polynomials)*
Let $f(x), g(x) \in R[x]$. Suppose the leading coefficient of g is a unit in R. Then there exist unique polynomials $q, r \in R[x]$ such that $\deg(r) < \deg(g)$ and

$$f(x) = q(x)g(x) + r(x).$$

Proof. By induction on the degree d of f. If $\deg(f) < \deg(g)$, take $q = 0$ and $r = f$. Otherwise, suppose f has leading coefficient a_d. Suppose g has degree $e \leq d$ and leading coefficient b_e. Then we have $f(x) = a_d b_e^{-1} x^{d-e} g(x) + f'(x)$ for some polynomial f'. The degree of f' is less than the degree of f, so by induction we have $f' = q'g + r$. It follows that $f = (a_d b_e^{-1} + q')x^{d-e}g + r$. For uniqueness, suppose $f = q_1 g + r_1 = q_2 g + r_2$ with $\deg(r_i) < \deg(g)$. Then $0 = (q_1 - q_2)g + (r_1 - r_2)$. The leading coefficient of g is invertible, and $\deg(r_1 - r_2) < \deg(g)$. It follows that the leading coefficient of $q_1 - q_2$ is zero, that is, $q_1 - q_2 = 0$. Therefore $r_1 - r_2 = 0$. □

Corollary 2.4.3. *If R is a field then $R[x]$ is a Euclidean domain with $\delta(f) = \deg(f)$.*

Theorem 2.4.4. *If a is a root of $f(x) \in R[x]$, then there exists a polynomial $q(x) \in R[x]$ such that*

$$f(x) = (x - a)q(x).$$

If R is an integral domain, then the number of distinct roots of f is no more than the degree of f (but see exercise 16).

Proof. Use the division theorem (Theorem 2.4.2) with $g = x - a$. The remainder r has degree zero but has a as a root. Thus r is zero. If R is an integral domain and if $b \neq a$ is another root of $f(x)$ then b is necessarily a root of $q(x)$. So the second statement follows by induction. □

The following theorem completes the proof of Theorem 2.2.14.

Theorem 2.4.5. *Suppose R is a GCD ring and a factorization domain. Then $R[x]$ is a factorization domain.*

Proof. We claim that every $f \in R[x]$ can be factored into a product of irreducibles. First we show that every $f \in R[x]$ has an irreducible divisor. Suppose not, and let d be the smallest degree of an element $f \in R[x]$ that has no irreducible divisor. Since R is a factorization domain, $d > 0$. Moreover, f is reducible. That is, $f = gh$ with neither g nor h a unit. The elements g and h

have no irreducible divisors since such a divisor would be a divisor of f as well. In particular, $\deg(h) > 0$. But then $\deg(g) < \deg(f)$ since R is an integral domain and this contradicts the minimality of $\deg(f)$.

Now let f be any element in $R[x]$. We already know that if f has degree zero, then it has an irreducible factorization, so let f have positive degree. Let a be the greatest common divisor of the coefficients of f and let $g = f/a$. If g has an irreducible factorization, then we obtain an irreducible factorization of f by multiplying those of g and a. Thus we may assume that the greatest common divisor of the coefficients of f is 1.

Now we use induction on the degree of f. If f has degree 1, then it is irreducible since no non-unit of R divides f other than an associate of f. If f has degree greater than 1, then by the first paragraph of this proof f has an irreducible divisor h. But h has positive degree so f/h has degree less than $\deg(f)$. By induction f/h has an irreducible factorization. Multiplying this by h gives an irreducible factorization of f. □

A root a of polynomial f is said to be *simple* if a is not a root of $f(x)/(x-a)$.

Lemma 2.4.6. *Let $q = \sum_{i=0}^{m} q_i x^i \in R[x]$ be a polynomial with coefficients in R. Consider the following statements*

1. *q_0 is invertible in R.*
2. *The polynomial x is invertible in the quotient ring $R[x]/(q)$.*
3. *The polynomials $q(x)$ and x are relatively prime in the ring $R[x]$.*
4. *There exists an integer $T > 0$ such that $q(x)$ is a factor of $x^T - 1$.*
5. *There exists an integer $T > 0$ such that $x^T = 1$ in the ring $R[x]/(q)$.*

Then statements (1), (2), and (3) are equivalent and if they hold, then

$$x^{-1} = -q_0^{-1}(q_1 + q_2 x + \cdots + q_m x^{m-1})$$

in $R[x]/(q)$. Statements (4) and (5) are equivalent (and the same T works for both) and $x^{-1} = x^{T-1}$ in $R[x]/(q)$. Statement (4) (or (5)) implies (1), (2), and (3). If R is finite then (1) (or (2) or (3)) implies (4),(5).

Proof. The statements are all straightforward except (possibly) the last one. Suppose that R is finite. Then the quotient ring $R[x]/(q)$ also contains finitely many elements so the powers $\{x^n\}$ of x in this ring cannot all be different. Hence there exists T such that $x^{n+T} \equiv x^n \pmod{q}$ for all sufficiently large n. Under assumption (2) this implies that $x^T \equiv 1 \pmod{q}$. In other words, q divides the polynomial $x^T - 1$, as claimed. □

When condition (4) (or (5)) in Lemma 2.4.6 holds, the smallest T such that $q(x)|(x^T - 1)$ is called the *order* of the polynomial q. This is admittedly confusing terminology since, in the

language of group theory, the order of the polynomial q is the order of x in the group $(R[x]/(q))^\times$. If condition (4) does not hold, then one may say that q does not have an order, or that its order is infinite. (For example, if $R = \mathbb{Q}$ the polynomial $q(x) = x - 2$ has infinite order.)

The following theorem will be useful when we discuss roots of unity.

Theorem 2.4.7. *Let a and b be positive integers. Then over any ring R the polynomial $x^a - 1$ divides $x^b - 1$ if and only if a divides b.*

Proof. By the Division Theorem for integers, we can write $b = qa + r$ with $0 \leq r < a$. Then

$$x^b - 1 = (x^{b-a} + x^{b-2a} + \cdots + x^r)(x^a - 1) + x^r - 1.$$

Sine $\deg(x^r - 1) < \deg(x^a - 1)$, it follows that $x^a - 1$ divides $x^b - 1$ if and only if $x^r - 1 = 0$. This holds if and only if $r = 0$, hence if and only if a divides b. □

2.4.b Polynomials over a field

Theorem 2.4.8. *If F is a field, then $F[x]$ is Euclidean with $\delta(f) = \deg(f)$. Every ideal in $F[x]$ has a unique monic principal generator. Any $f(x) \in F[x]$ can be written in the form*

$$f(x) = a p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$$

where $a \in F$, the p_i are distinct monic irreducible elements of $F[x]$, and the e_i are positive integers. This representation is unique apart from changing the order of the p_i.

Proof. It follows from Theorem 2.4.2 that $F[x]$ is Euclidean. It is also principal and is a UFD by Theorem 2.2.14. Each irreducible polynomial has a unique monic associate (divide by the leading coefficient). This accounts uniquely for a. □

It also follows from Theorem 2.2.14 that $F[x]$ is a GCD ring, but to be precise we have the following theorem.

Theorem 2.4.9. *Let F be a field and $f_1, \cdots, f_k \in F[x]$, not all zero. There is a unique monic $g \in F[x]$ such that (1) g divides every f_i and (2) if h divides every f_i then h also divides g. Moreover, g can be written in the form*

$$g = h_1 f_1 + h_2 f_2 + \cdots + h_k f_k \tag{2.17}$$

for some $h_1, h_2, \cdots, h_k \in F[x]$.

Proof. Let $I = \{h_1 f_1 + h_2 f_2 + \cdots + h_k f_k : h_1, h_2, \cdots, h_k \in F[x]\}$. Then I is an ideal in $F[x]$, so by Theorem 2.4.8, I has a unique monic generator g. Since $g \in I$, g can be written in the form in equation (2.17). It follows that any h that divides every f_i also divides g. Since $f_i \in I$, g divides f_i. □

We write $g = \gcd(f_1, \cdots, f_k)$. It can be found by the usual Euclidean algorithm by repeatedly using Theorem 2.4.2. There is also a notion of least common multiple in $F[x]$. The following theorem later allows us to construct finite fields of all possible sizes. The proof is omitted.

Theorem 2.4.10. *If F is a finite field and d is a positive integer, then there is at least one irreducible polynomial of degree d in $F[x]$.*

If $F \subseteq E$ are fields and if $a \in E$ is an element that is the root of some polynomial with coefficients in F, then we say a is *algebraic over F*. A polynomial $f \in F[x]$ is called a *minimal polynomial* of a (over F) if it is monic, if $f(a) = 0$ and if it is a polynomial of smallest degree with these properties.

Theorem 2.4.11. *Suppose a is algebraic over F. Then it has a unique minimal polynomial $f \in F[x]$. The minimal polynomial f is also the unique monic irreducible polynomial in $F[x]$ having a as a root. If $g \in F[x]$ is any other polynomial such that $g(a) = 0$ then f divides g in $F[x]$.*

Proof. If two monic polynomials $f, g \in F[x]$ have the same (minimal) degree and both have a as a root then $f - g$ has smaller degree, which is a contradiction. If f is a minimal polynomial of a and $f = gh$, then $0 = f(a) = g(a)h(a)$ so $g(a) = 0$ or $h(a) = 0$. By the minimality of f, whichever factor has a as a root must have the same degree as f, so f is irreducible.

Now suppose f is a monic irreducible polynomial such that $f(a) = 0$. The set

$$J = \{h \in F[x] : \ h(a) = 0\}$$

is an ideal, so it is principal. It contains f, but f is irreducible, so $J = (f)$ is the ideal generated by f, and f is the unique monic polynomial with this property. If $g(a) = 0$ then $g \in J$ so g is a multiple of f. In particular, f is the minimal polynomial of a. □

More generally, we can think consider the "operator" on rings that takes a ring R to the polynomial ring $R[x]$. Strictly speaking this is not a function since there is no set of all rings. Rather, it is a (covariant) functor on the category of rings. We shall not, however pursue these notions in this book.

2.5 Exercises

1. Prove that if G_1 and G_2 are groups, then the direct product $G_1 \times G_2$ is a group. Prove that $G_1 \times G_2$ is Abelian if G_1 and G_2 are Abelian.

2. Describe the set of all subgroups of the group $\mathbb{Z}/m\mathbb{Z}$.

3. Let $\varphi : G \to H$ be a group homomorphism. Prove that $\mathrm{Ker}(\varphi)$ is a subgroup of G and $\mathrm{Im}(\varphi)$ is a subgroup of H.

4. Let G be a group and let H be a subgroup of G. Prove that the relation defined by $a \sim b$ if there is an $h \in H$ such that $b = ah$ is an equivalence relation. Find an example where the definition $aHbH = abH$ does not make the set of equivalence classes into a group.

5. Prove that a subgroup H of a group G is normal if and only if for every $a \in G$ and $h \in H$, we have $aha^{-1} \in H$.

6. Theorem 2.1.15: Let $\varphi : G \to G'$ be a homomorphism.

   a. Prove that $\mathrm{Ker}(\varphi)$ is normal in G.
   b. Prove that the quotient $G/\mathrm{Ker}(\varphi)$ is isomorphic to $\mathrm{Im}(\varphi)$.
   c. Conversely, prove that if H is a normal subgroup of G, then the map $a \mapsto aH$ is a surjection from G to G/H with kernel equal to H.

7. Show that the set of endomorphisms of an Abelian group is a ring.

8. Theorem 2.2.7:

   a. Suppose $\varphi : R \to S$ is a ring homomorphism. Prove that $\mathrm{Ker}(\varphi)$ is an ideal of R and φ induces an isomorphism between $R/\mathrm{Ker}(\varphi)$ and the image of f.
   b. Prove that if I is an ideal of R, then the map $a \mapsto a + I$ is a homomorphism from R onto R/I with kernel I.

9. Prove that a GCD ring with no infinite chain of proper ascending ideals is also a LCM (least common multiple) ring.

10. Let $\{R_s : s \in S\}$ be a family of rings. Prove that R_S is the unique (up to isomorphism) ring such that if T is any ring and $\psi_s : T \to R_s$ any set of homomorphisms, then there is a homomorphism $g : T \to R_S$ such that $\psi_s = \varphi_s \circ g$ for every $s \in S$.

11. Prove that if V is a vector space over a field F, then for every $u \in V$ we have $0 \cdot u = 0$.

12. Theorem 2.2.21:

   a. Prove that every vector space has a basis. (Hint: use Zorn's Lemma.)
   b. Prove that if S is a basis for a vector space V, then every vector can be written uniquely as a linear combination of elements of S.

13. Prove by induction on the dimension that every discrete $\mathbb{Z}$-module in $\mathbb{R}^n$ is a lattice.

14. Let $f(x) = a_0 + a_1 x + \cdots + a_d x^d \in \mathbb{C}[x]$ be a polynomial and let $F : \mathbb{Z} \to \mathbb{C}$ be the function $F(i) = a_i$ (and $F(i) = 0$ if $i < 0$ or $i > d$). Let $g(x) = b_0 + b_1 x + \cdots + b_e x^e$ and let $G : \mathbb{Z} \to \mathbb{C}$ be the corresponding function. Show that the product $f(x)g(x)$ polynomial corresponds to the convolution $F * G$.

15. Develop a theory of characters as functions with values in an arbitrary field F rather than $\mathbb{C}$. For certain parts you will need to assume that F contains the n-th roots of unity.

16. Let $R = \mathbb{Z} \times \mathbb{Z}$. Let $f(x) = (1, 0)x - (1, 0) \in R[x]$. Show that f has infinitely many roots in the ring R.

Chapter 3 Fields

Fields are rings where every nonzero element is a unit. Many sequence generators can be viewed as implementing multiplication by a fixed element in a ring. Since finite fields have cyclic groups of units, they provide a source of large period sequence generators. In Section 3.1 we describe the Galois theory of field extensions. In Sections 3.2 and 3.4 we study in some detail two important classes of fields – finite fields, which give us a way to make algebraic constructions with finite alphabets, and algebraic number fields, which generalize the field of rational numbers. In Section 3.5 we describe local fields. These are fields that are complete with respect to a notion of convergent sequences. Elements of these fields can sometimes be viewed as infinite sequences over some alphabet. We also study quadratic forms, which are the source of several important constructions of sequences with good correlation properties (see Section 3.3).

3.1 Field extensions

In this section we summarize (without proofs) some standard facts about field extensions.

3.1.a Galois group

If F is a field and E is a ring, then the kernel of any nonzero homomorphism $F \to E$ is the zero ideal (the only proper ideal), so every homomorphism is an injection. We say that E is an *extension* of F. Elements of E can then be added and multiplied by elements of F so E becomes a vector space over F. The dimension of E as a vector space over F is called the *degree* or the *dimension* of the extension.

A field R is *algebraically closed* if every polynomial $p(x) \in F[x]$ factors completely, $p(x) = k(x-a_1)(x-a_2)\cdots(x-a_n)$ where $\deg(p) = n$ and where $k, a_i \in F$. Every field F is contained in an algebraically closed field $\overline{F}$ of finite degree over F, called an *algebraic closure* of F.

If G is a subgroup of the group of automorphisms of a field E, then the set of elements in E that are fixed by every automorphism in G (that is, $\sigma(a) = a$ for every $a \in E$ and every $\sigma \in G$) is denoted E^G. It is necessarily a field since it is closed under addition, multiplication, and inverse. If $F \subset E$ are fields then the group $\mathrm{Aut}_F(E)$ of automorphisms of E which fix each element of F is the *Galois group* of E over F and it is denoted by $\mathrm{Gal}(E/F)$. If $G = \mathrm{Gal}(E/F)$, then $F \subseteq E^G$. If in fact $F = E^G$, then we say that E *is a Galois extension of* F . The discovery of Galois extensions by Evariste Galois was a turning point in the understanding of the nature of algebraic equations and triggered a great transformation in the way mathematics was done.

If $F \subset E$ is a finite extension of fields and if $\overline{F}$ is an algebraic closure of F then there are finitely many embeddings $h_1, \cdots, h_n : E \to \overline{F}$. If E is a Galois extension of F then these embeddings all have the same image. In this case, a choice of one embedding (say, h_1) determines a one to one correspondence $h_i \leftrightarrow \sigma_i$ with elements of the Galois group $\mathrm{Gal}(E/F)$ by $h_i(x) = h_1(\sigma_i(x))$.

Proposition 3.1.1. *Let $F \subset E$ be a finite extension and let $T : E \to F$ be a nonzero F-linear map. Then for any F-linear map $f : E \to F$ there exists a unique element $a \in E$ such that $f(x) = T(ax)$ for all $x \in E$.*

Proof. The field E has the structure of a vector space over F, of some finite dimension, say, n. Then $\mathrm{Hom_F(E, F)}$ is the dual vector space and it also has dimension n. Each $a \in E$ gives an element $f_a \in \mathrm{Hom_F(E, F)}$ by $f_a(x) = T(ax)$ which is also nonzero unless $a = 0$. So the association $a \mapsto f_a$ gives a mapping $E \to \mathrm{Hom_F(E, F)}$ which is a homomorphism of n dimensional vector spaces, whose kernel is 0. Therefore it is an isomorphism. □

3.1.b Trace and norm

Let E be an extension of degree $n < \infty$ of a field F. Choose a basis $e_1, e_2, \cdots, e_n$ of E as a vector space over F. Each $a \in E$ defines a mapping $L_a : E \to E$ by $L_a(x) = ax$. This mapping is E-linear, hence also F-linear, so it can be expressed as an $n \times n$ matrix M_a with respect to the chosen basis. If $a \neq 0$ then the matrix M_a is invertible. The *trace*, $\mathsf{Tr}_F^E(a)$ and *norm* $\mathsf{N}_F^E(a)$ are defined to be the trace and determinant (respectively) of the matrix M_a. It is common to write $\mathsf{Tr}(a) = \mathsf{Tr}_F^E(a)$ and $\mathsf{N}(a) = \mathsf{N}_F^E(a)$ if the fields E and F are understood.

Theorem 3.1.2. *Let $F \subset E$ be a finite extension of fields.*

1. *For all $a, b \in E$ and $c \in F$ we have $\mathsf{Tr}(a + b) = \mathsf{Tr}(a) + \mathsf{Tr}(b)$ and $\mathsf{Tr}(ca) = c\mathsf{Tr}(a)$. That is, Tr is F-linear.*
2. *For all $a, b \in E$ we have $\mathsf{N}(ab) = \mathsf{N}(a)\mathsf{N}(b)$ so $\mathsf{N}_F^E : E^\times \to F^\times$ is a homomorphism of multiplicative groups. It is surjective.*
3. *If $F \subset L \subset E$ are finite extensions then for all $a \in E$,*
$$\mathsf{Tr}_F^L(\mathsf{Tr}_L^E(a)) = \mathsf{Tr}_F^E(a) \quad \text{and} \quad \mathsf{N}_F^L(\mathsf{N}_L^E(a)) = \mathsf{N}_F^E(a).$$
4. *If E is a Galois extension of F then $\mathsf{Tr}_F^E : E \to F$ is nonzero and*
$$\mathsf{Tr}_F^E(a) = \sum_{\sigma \in \mathrm{Gal}(E/F)} \sigma(a) \quad \text{and} \quad \mathsf{N}_F^E = \prod_{\sigma \in \mathrm{Gal}(E/F)} \sigma(a).$$

Parts (1) and (2) are straightforward. We omit the proofs (see [119], [17]) of parts (3) and (4) but we will return to the trace and norm in Section 3.2 and Section 3.4. There are situations in which the trace $\mathsf{Tr}_F^E : E \to F$ is the zero map, but if E, F are finite fields or if $\mathrm{char}(E) = 0$ or if E is a Galois extension of F then the trace is not zero.

3.2 Finite fields

In this section we analyze the structure of finite fields, or *Galois fields*. For a more complete treatment see the excellent reference by Lidl and Niedereitter [123]. Our first task is to identify all finite fields and all inclusion relations among them.

3.2.a Basic properties

Theorem 3.2.1. *Let p be a prime number. For each $d > 0$ there is (up to isomorphism) a unique field $\mathbb{F}_{p^d}$ with p^d elements. These account for all finite fields. If $e > 0$ is another integer, then there is an inclusion $\mathbb{F}_{p^d} \subseteq \mathbb{F}_{p^e}$ if and only if d divides e. That is, the (combinatorial) lattice of finite fields with characteristic p under inclusion is isomorphic to the lattice of whole numbers under divisibility. The subfield $\mathbb{F}_{p^d}$ consists of those elements a of $\mathbb{F}_{p^e}$ satisfying $a^{p^d} = a$.*

The field $\mathbb{F}_{p^d}$ is sometimes denoted $GF(p^d)$ (for "Galois field"). The proof of Theorem 3.2.1 will occupy the rest of Section 3.2.a.

Suppose d is a positive integer and F is a finite field with r elements. Let $f(x)$ be an irreducible polynomial over F with degree d. Then by Theorem 2.2.5.4, $F[x]/(f(x))$ is a field. It has r^d elements. In particular, if p is a prime integer and we take $F = \mathbb{Z}/(p)$, then this together with Theorem 2.4.10 shows that there exists a finite field of order p^d for every prime p and positive integer d.

Next suppose F is a finite field with characteristic $p > 0$. Recall that we showed in Theorem 2.2.8 that p is prime. It follows that the mapping $\mathbb{Z}/(p) \to F$ which takes an element n to $1 + 1 + \cdots + 1$ (n times) is a ring homomorphism. So we can view $\mathbb{Z}/(p)$ as a subfield of F. Hence F has the structure of a finite dimensional vector space over $\mathbb{Z}/(p)$. By Theorem 2.2.22, F has p^d elements for some d.

Proposition 3.2.2. *If $F \subseteq E$ are two finite fields, then E and F have the same characteristic. If p is the characteristic, then $|F| = p^d$ and $|E| = p^e$ for some integers d and e such that d divides e.*

Proof. If F has characteristic p and E has characteristic r, then $|F| = p^d$ and $|E| = r^e$ for some d and e. But E is a vector space over F, so $r^e = (p^d)^k$ for some k. Thus $r = p$ and $e = dk$. □

To complete the picture of the set of finite fields we want to show that there is, up to isomorphism, a unique finite field of a given cardinality. First we need a lemma.

Lemma 3.2.3. *If F is a finite field, then every $a \in F$ is a root of the polynomial $x^{|F|} - x$ and we have*

$$x^{|F|} - x = \prod_{a \in F} (x - a).$$

No other element of any extension field of F is a root of this polynomial.

Proof. The multiplicative group of F has order $|F|-1$, so by Theorem 2.1.3 any nonzero element $a \in F$ satisfies $a^{|F|-1} = 1$. Therefore any element $a \in F$ satisfies $a^{|F|} = a$. That is, every a is a root of the polynomial $x^{|F|} - x$. It follows that $x - a$ divides $x^{|F|} - x$. Furthermore, the degree of $x^{|F|} - x$ equals $|F|$, so there are no other roots of this polynomial in E. The factorization follows from Theorem 2.4.4. □

Corollary 3.2.4. *Suppose E is a field, p is a prime number, and d is a positive integer. Then E contains at most one subfield of order p^d.*

Proof. Suppose F is a subfield of E of order p^d. By Lemma 3.2.3 every $a \in F$ is a root of $x^{p^d} - x$, and there are no other roots of this polynomial in E.

Now suppose F' is another subfield of E of order p^d. The same reasoning applies to F'. Thus $F = F'$. □

Proposition 3.2.5. *Let p be a prime number and let $d > 0$ be an integer. Any two finite fields with p^d elements are isomorphic.*

Proof. Let $E = (\mathbb{Z}/(p))[x]/(f(x))$, where $f(x)$ is an irreducible polynomial with degree d and coefficients in $\mathbb{Z}/(p)$. It is enough to show that any field F with p^d elements is isomorphic to E.

By Lemma 3.2.3, every $a \in E$ satisfies $a^{p^d} = a$. In particular, $x^{p^d} - x = 0$ in E, so $f(x)$ divides $x^{p^d} - x$ as polynomials. That is, $x^{p^d} - x = f(x)g(x)$ for some $g(x) \in (\mathbb{Z}/(p))[x]$.

On the other hand, we can think of $x^{p^d} - x$ as a polynomial over F. By the same reasoning, every element of F is a root of this polynomial, so

$$f(x)g(x) = x^{p^d} - x = \prod_{a \in F} (x - a).$$

In particular, $f(x)$ factors into linear factors over F. Let a be a root of $f(x)$ in F. If the elements $\{1, a, a^2, \cdots, a^{d-1}\}$ were linearly dependent over $(\mathbb{Z}/(p))[x]$, a would be a root of a lower degree polynomial, and this polynomial would divide $f(x)$. That would contradict the irreducibility of $f(x)$. Thus they are linearly independent and hence a basis (F has dimension d over $(\mathbb{Z}/(p))[x]$). That is, every b in F can be written

$$b = \sum_{i=0}^{d-1} c_i a^i,$$

with $c_i \in (\mathbb{Z}/(p))[x]$. We define a function

$$L\left(\sum_{i=0}^{d-1} c_i a^i\right) = \sum_{i=0}^{d-1} c_i x^i$$

from F to E. This function is one-to-one and it can be checked that it preserves multiplication and addition. Hence it is an isomorphism. □

Thus for each prime power $q = p^d$ there is a unique field $\mathbb{F}_q$ with q elements.

Proposition 3.2.6. *Let p be prime and let d, e be positive integers. Then the field $F = \mathbb{F}_{p^d}$ may be realized as a subfield of $E = \mathbb{F}_{p^e}$ if and only if d divides e. In this case it is the set*

$$F = \left\{ x \in E : \ x^{p^d} = x \right\}.$$

Proof. If F is a subfield of E then E is a vector space over F, of some dimension k. Consequently $|E| = |F|^k$ so $e = dk$. To prove the converse, assume $e = dk$ for some positive integer k. Let $q = p^d = |F|$. Recall from Lemma 3.2.3 that E consists of the distinct roots of the polynomial $x^{p^e} - x = x^{q^k} - x$. This polynomial is divisible by the polynomimal $x^q - x$, for the quotient is

$$x^{(q^k-1)-(q-1)} + x^{(q^k-1)-2(q-1)} + \cdots + x^{q-1} + 1.$$

Thus E contains a set S of q distinct roots of the polynomial $(x^q - x)$. By Lemma 2.2.9, both addition and multiplication commute with raising to the qth power, so the subset $S \subset E$ is a field. Therefore it is isomorphic to the field $F = \mathbb{F}_q$. □

Suppose $f \in F[x]$ is irreducible. Recall that in the terminology of Section 2.4.a, the order of f is the smallest T such that $f(x)|(x^T - 1)$. This is the order of x in the group of units of $F[x]/(f)$, a group that has $|F|^{\deg(f)} - 1$ elements. Thus by Theorem 2.1.3 the order of f divides $|F|^{\deg(f)} - 1$.This completes the proof of Theorem 3.2.1.

3.2.b Galois groups of finite fields

Some of the preceding notions can be understood in terms of Galois groups (see Section 3.1.a). Let $E = \mathbb{F}_{p^e}$ where p is prime. By Lemma 2.2.9 the mapping $\sigma : E \to E$, $\sigma(x) = x^p$ is a field automorphism, meaning that it is additive, multiplicative, and invertible. However $\sigma^e(x) = x^{p^e} = x$ so $\sigma^e = I$ is the identity. Thus the various powers of σ (including $\sigma^0 = I$) form a cyclic group of automorphisms, of order e, which fix each element of $\mathbb{F}_p$.

Proposition 3.2.7. *The group $\{\sigma^0 = I, \sigma, \cdots, \sigma^{e-1}\}$ is the Galois group* $\mathrm{Gal}(E/\mathbb{F}_p)$.

Proof. The Galois group $\mathrm{Gal}(E/F)$ is the set of automorphisms of E that fix each element of F. So it suffices to show that any automorphism $\tau : E \to E$ is some power of σ. Let f be an irreducible polynomial over $\mathbb{F}_p$ with degree e, and let a be a root of f. Then $\mathbb{F}_{p^d} = \mathbb{F}_p[a]$ and $1, a, a^2, \cdots, a^{e-1}$ is a basis for $\mathbb{F}_{p^e}$ over $\mathbb{F}_p$. Thus to show that two automorphisms are equal, it suffices to show that they are equal on a. We have that $\sigma^i(f) = f$ for every i, so $\sigma^i(a)$ is a root of f. Similarly, $\tau(a)$ is a root of f. The $\sigma^i(a)$ are distinct – otherwise a and hence $\mathbb{F}_{p^e}$ are in a proper subfield, which is a contradiction. Thus there are $e = \deg(f)$ of them, and they account for all the roots of f. In particular, $\tau(a) = \sigma^i(a)$ for some i. So $\tau = \sigma^i$, proving the proposition. □

Theorem 3.2.8. *Let $F = \mathbb{F}_{p^d} \subset E = \mathbb{F}_{p^e}$ be finite fields. Then the Galois group* $\mathrm{Gal}(E/F)$ *is a cyclic subgroup group of* $\mathrm{Gal}(E/\mathbb{F}_p)$*, of order e/d. It is generated by the automorphism $\sigma^d : x \mapsto x^{p^d}$. The field $F \subset E$ consits of those elements of E that are fixed by every element of* $\mathrm{Gal}(E/F)$ *(which is the same as being fixed by the generator σ^d). Consequently the field E is a* Galois extension *of the field F.*

Proof. It follows from Proposition 3.2.6 that F is the subfield of E that is fixed by σ^d. So the various powers of σ^d are contained in $\mathrm{Gal}(E/F)$. By Proposition 3.2.7, every automorphism of F is some power of σ. But d is the smallest power of σ that fixes F because the equation $x^{p^k} = x$ has at most p^k solutions. Consequently $\mathrm{Gal}(E/F)$ consists of all powers of σ^d. These elements form a cyclic subgroup of $\mathrm{Gal}(E/\mathbb{F}_p)$ of order e/d. □

Thus we have an inclusion reversing correspondence between the lattice of subfields of $\mathbb{F}_{p^d}$ and the lattice of subgroups of $\mathrm{Gal}(\mathbb{F}_{p^d}/\mathbb{F}_p)$. The main theorem of Galois theory describes the solutions of a polynomial equation in terms of the Galois group.

Theorem 3.2.9. *Let F be a finite field with q elements and let $f(x) \in F[x]$ be a polynomial of degree d with coefficients in F. Let E be an extension field of F and suppose $\alpha \in E$ is a root of f. Then for any $\sigma \in Gal(E/F)$, the element $\sigma(\alpha) \in E$ is also a root of f. If f is irreducible in $F[x]$ and if E is the degree d extension of F then all the roots of f are contained in E. They consist exactly of the Galois conjugates,*

$$\sigma_i(\alpha) = \alpha^{q^i},$$

where $0 \leq i \leq d-1$. That is, where σ_i ranges over all elements of $Gal(E/F)$.

Proof. Let $q = |F|$. The Galois group $Gal(E/F)$ is cyclic and it is generated by the mapping $\sigma : E \to E$ given by $\sigma(a) = a^q$. If $f(x) = \sum_{i=0}^{d} a_i x^i$ and if $\alpha \in E$ is a root of f, then

$$0 = \sigma(f(\alpha)) = \left(\sum_{i=0}^{d} a_i \alpha^i \right)^q = \sum_{i=0}^{d} a_i^q \alpha^{iq} = \sum_{i=0}^{d} a_i \sigma(\alpha) = f(\sigma(\alpha))$$

(by Lemma 2.2.9), so $\sigma(\alpha)$ is also a root of f.

Now suppose f is irreducible and, without loss of generality, monic. Then it is the minimal polynomial of α by Theorem 2.4.11. But the polynomial

$$g(x) = \prod_{\tau \in Gal(E/F)} (x - \tau(\alpha)) \in E[x]$$

has the same degree as f, and it is fixed under each element of $Gal(E/F)$. So $g \in F[x]$, and it has α as a root. Therefore $g = f$, so the roots of f are all the Galois conjugates of α. □

3.2.c Primitive elements

To work within a particular finite field F, it is useful to have some structural information. An element $a \in F$ is called *primitive* if every nonzero element of F can be written as a power of a. A polynomial $f \in \mathbb{F}_p[x]$ of degree d is primitive if it is irreducible and if one (and hence all) of its roots in $\mathbb{F}_{p^d}$ are primitive elements. The following lemma will be used in Section 12.2.

Lemma 3.2.10. *Let $F = \mathbb{F}_q$ be the field with q elements. Let $f \in F[x]$ be a polynomial. Then f is primitive if and only if its order is $q^{\deg(f)} - 1$.*

Proof. In the ring $F[x]/(f)$ the element x is a root of the polynomial $f(x)$. If x is primitive then the order of x is $T = |F| - 1 = q^{\deg(f)} - 1$. Thus T is the smallest integer such that $x^T = 1 \pmod f$, which is to say that T is the smallest integer such that f divides $x^T - 1$. Thus the order of f is T. The converse is similar. □

We next show that every finite field has primitive elements. This implies that the multiplicative group of a finite field is cyclic.

Proposition 3.2.11. *The finite field $\mathbb{F}_{p^d}$ has $\phi(p^d - 1)$ primitive elements.*

Proof. Suppose that $a \in \mathbb{F}_{p^d}$ has order e. That is, $a^e = 1$ and no smaller positive power of a equals 1. Then the elements $1, a, a^2, \cdots, a^{e-1}$ are distinct and are all roots of $x^e - 1$. That is,

$$x^e - 1 = (x-1)(x-a)(x-a^2)\cdots(x-a^{e-1}).$$

It follows that every element whose eth power equals 1 is a power of a, and an element $b = a^i$ has order e if and only if $\gcd(i, e) = 1$. Thus if there is at least one element of order e, then there are exactly $\phi(e)$. That is, for every e there are either 0 or $\phi(e)$ elements of order e.

Furthermore, by Lemma 3.2.3 every nonzero $a \in F$ is a root of the polynomial $x^{p^d-1} - 1$. Thus if there is an element in F with order e, then e divides $p^d - 1$. By Lemma 2.2.10, for any positive integer k

$$\sum_{e|k} \phi(e) = k.$$

Thus we have

$$\begin{aligned} p^d - 1 &= \sum_{e|p^d-1} |\{a \in F : \text{ the order of } a = e\}| \\ &\leq \sum_{e|p^d-1} \phi(e) = p^d - 1. \end{aligned}$$

Therefore the two sums are equal. Since each term in the first sum is less than or equal to the corresponding term in the second sum, each pair of corresponding terms must be equal.

In particular, the number elements with order $p^d - 1$ equals $\phi(p^d - 1) > 0$. □

In fact, it can be shown that every finite field $\mathbb{F}_{p^d}$ has a *primitive normal basis* over a subfield $\mathbb{F}_{p^c}$. This is a basis of the form $a, a^{p^c}, \cdots, a^{p^{d-c}}$ with a primitive. The interested reader can find the details in [123, Section 2.3].

3.2.d Roots of unity

Let $N \in \mathbb{Z}$ be a positive integer. Over the complex numbers the polynomial $x^N - 1$ factors completely into distinct linear factors

$$x^N - 1 = \prod_{j=0}^{N-1} (x - \zeta^j)$$

where $\zeta \in \mathbb{C}$ is a primitive N-th root of unity, for example, $\zeta = e^{2\pi i/N}$. These N-th roots of unity form an Abelian group under multiplication, denoted μ_N, that is isomorphic to $\mathbb{Z}/(N)$. The field $\mathbb{Q}(\zeta)$ is called a *cyclotomic field*. It is a Galois extension of $\mathbb{Q}$ of degree $[\mathbb{Q}(\zeta) : \mathbb{Q}] = \phi(n)$. The Galois group is Abelian and is isomorphic to $\mathbb{Z}/(N)^\times$. If $s \in \mathbb{Z}/(N)^\times$ is relatively prime to N then the corresponding element $\sigma_s \in \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$ acts on $\mathbb{Q}(\zeta)$ by $\sigma_s(\zeta^k) = \zeta^{ks}$. The following fact is used in Proposition 13.4.4.

Lemma 3.2.12. *Let $\zeta \in \mathbb{C}$ be a primitive Nth root of unity and let $p(x) \in \mathbb{Q}[x]$ be a polynomial with rational coefficients. Suppose $|p(\zeta)|^2 \in \mathbb{Q}$ is a rational number. Then $|p(\zeta^s)|^2 = |p(\zeta)|^2$ for any integer s relatively prime to N with $1 \leq s \leq N-1$.*

Proof. Let $p(x) = a_0 + a_1 x + \cdots + a_d x^d$ with $a_i \in \mathbb{Q}$. Then $|p(\zeta)|^2 = p(\zeta)\overline{p(\zeta)} \in \mathbb{Q}$ is fixed under the action of the element $\sigma_s \in \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$, where $\overline{\zeta} = \zeta^{-1}$ denotes complex conjugation. Therefore

$$\begin{aligned} p(\zeta)\overline{p(\zeta)} &= \sigma_s\left(p(\zeta)\overline{p(\zeta)}\right) \\ &= (a_0 + a_1\zeta^s + \cdots + a_k\zeta^{ks})(a_0 + a_1\zeta^{-s} + \cdots + a_k\zeta^{-ks}) \\ &= p(\zeta^s)\overline{p(\zeta^s)} \end{aligned}$$

□

The situation is more complicated over a finite field. Let $F = \mathbb{F}_q$ be a finite field of characteristic p. Let $N \in \mathbb{Z}$ be a positive integer. Define d as follows. Write $N = p^e n$ where p does not divide n. Let $d = \mathrm{ord}_n(q)$. (In the group theoretic sense: the image of q in $\mathbb{Z}/(n)$ is invertible, and d is the least integer such that $q^d \equiv 1 \pmod{n}$, cf. Section 2.2.d.) The following theorem says that $x^N - 1$ factors completely in the extension field $\mathbb{F}_{q^d}$ of $\mathbb{F}_q$, but the roots are not distinct if $p^e > 1$.

Theorem 3.2.13. *Given N, q as above, there exists $\beta \in \mathbb{F}_{q^d}$ such that*

$$x^N - 1 = \prod_{i=0}^{n-1} (x - \beta^i)^{p^e}. \tag{3.1}$$

Moreover, $\mathbb{F}_{q^d}$ is the smallest extension of $\mathbb{F}_q$ over which $x^N - 1$ splits into linear factors.

Proof. Let $\alpha \in \mathbb{F}_{q^d}$ be a primitive element and let $\beta = \alpha^{(q^d-1)/n}$. Since $q^d \equiv 1 \pmod{n}$ the exponent $(q^d - 1)/n$ is an integer. The powers $\beta^0, \beta^1, \cdots, \beta^{n-1} \in \mathbb{F}_{q^d}$ are distinct, and $\beta^n = 1$. Thus β is a primitive n-th root of unity, and $x^n - 1 = \prod_{i=1}^{n-1}(x - \beta^i)$. Equation (3.1) follows. The minimality of $\mathbb{F}_{q^d}$ is left as an exercise. □

The factors in equation (3.1) can be grouped together to give the factorization of $x^N - 1$ over the field $\mathbb{F}_q$. Let $\gamma = \beta^k \in \mathbb{F}_{q^d}$ be any root of $x^n - 1$. Then the remaining roots of the minimal polynomial of γ over $\mathbb{F}_q$ are $\{\gamma^{q^i} = \beta^{kq^i} : i \geq 0\}$. The set of exponents

$$C_k(q) = C_k = \{k, qk \pmod{n}, q^2k \pmod{n}, \cdots\}$$

is called the *kth cyclotomic coset modulo n relative to q* (the terms "modulo n" and "relative to q" may be omitted if n and/or q are understood). The minimal polynomial of γ is then the product

$$f_k(x) = \prod_{i \in C_k} (x - \beta^i).$$

If $C_{j_1}, \cdots, C_{j_m}$ are the distinct cyclotomic cosets in $\{0, 1, \cdots, n-1\}$, then they form a partition of $\{0, 1, \cdots, n-1\}$ and the desired factorizations are

$$x^n - 1 = \prod_{i=1}^{m} f_k(x) \quad \text{and} \quad x^N - 1 = \prod_{i=1}^{m} f_k(x)^{p^e}.$$

3.2.e Trace and norm on finite fields

Theorem 3.2.14. *Let $F = \mathbb{F}_q \subset E = \mathbb{F}_{q^n}$ be finite fields of characteristic p. Then the following statements holds.*

1. *The trace function $\mathsf{Tr}_F^E : E \to F$ is given by*

$$\mathsf{Tr}_E^F(a) = a + a^q + a^{q^2} + \cdots + a^{q^{n-1}} = \sum_{\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})} \sigma(a). \tag{3.2}$$

2. *The norm is given by*

$$\mathsf{N}_F^E(a) = \prod_{\sigma \in \mathrm{Gal}(E/F)} \sigma(a) = a^{(q^n-1)/(q-1)} \in F.$$

3. *The trace is nonzero and for all $c \in F$, we have $|\{a \in E : \mathsf{Tr}(a) = c\}| = p^{e-d}$.*

4. *For all* $0 \neq c \in F$ *we have* $|\{a \in E : \mathsf{N}(a) = c\}| = (|E| - 1)/(|F| - 1)$.
5. *For all* $a \in E$ *we have* $\mathsf{Tr}_F^E(a^p) = \mathsf{Tr}_F^E(a)^p$.
6. $\mathsf{Tr}_F^E(1) \in \mathbb{F}_p$ *and* $\mathsf{Tr}_F^E(1) \equiv n \pmod{p}$.
7. *If* $L : E \to F$ *is an* F*-linear function, then there is an element* $a \in E$ *such that* $L(b) = \mathsf{Tr}(ab)$ *for all* $b \in E$.

Proof. 1. Verification of the second equality in equation (3.2) is left as an exercise. It follows that the quantity on the right side of equation (3.2), which we denote by $T(a)$, is fixed by each $\sigma \in \mathrm{Gal}(E/F)$ so it is indeed an element of F. Both maps T and Tr are F-linear, hence are equal if and only if they are equal on a basis. Let $a \in E$ be a root of an irreducible polynomial of degree n over F. Then the set $\{1, \alpha, \alpha^2, \cdots, \alpha^{n-1}\}$ forms a basis for E (over F). If the minimal polynomial for a is

$$f(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + x^n$$

then the matrix

$$M_a = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ & & \vdots & & \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{m-2} & -a_{m-1} \end{pmatrix} \tag{3.3}$$

is the corresponding *companion matrix*: it has 1 in each entry of the superdiagonal, $-a_0, \cdots, -a_{d-1}$ in the last row, and 0s elsewhere. The characteristic polynomial of the matrix M_a is exactly the polynomial $f(x)$, so the eigenvalues of M_a (i.e. the roots of its characteristic polynomial) are the Galois conjugates of a. So the trace of M_a is $-a_{n-1}$. On the other hand,

$$f(x) = \prod_{\sigma \in \mathrm{Gal}(E/F)} (x - \sigma(a)),$$

so $-a_{n-1} = \sum_{\sigma \in \mathrm{Gal}(E/F)} \sigma(a)$.

2. The same argument applies to the determinant of the matrix M_a, which is

$$(-1)^n a_0 = (-1)^{2n} \prod_{\sigma \in \mathrm{Gal}(E/F)} \sigma(a).$$

3. Thinking of E as an n-dimensional vector space over F, the mapping $\mathsf{Tr} : E \to F$ is linear, so its rank is either zero or 1. If the rank is zero then $\mathsf{Tr}(x) = 0$ for all $x \in E$, however this equation is a polynomial of degree q^{n-1} so it has at most $q^{n-1} < q^n$ solutions. Therefore $\mathsf{Tr} : E \to F$ is surjective so its kernel $K = \mathsf{Tr}^{-1}(0) \subset E$ is a vector subspace of dimension $n - 1$ which therefore contains q^{n-1} elements. For any $0 \neq a \in F$ the set $\mathsf{Tr}^{-1}(a)$ is a translate of K, that is, an affine subspace of the same dimension, which therefore contains the same number of elements.

4. The norm is a homomorphism $\mathsf{N}_F^E : E^\times \to F^\times$. If α is a primitive element in E, then

$$\mathsf{N}_F^E(\alpha) = \prod_{i=0}^{n-1} \alpha^{q^i} = \alpha^{\sum_{i=0}^{n-1} q^i} = \alpha^{(q^n-1)/(q-1)},$$

which is primitive in F. Thus N_F^E surjective so its kernel is a subgroup of order $(|E|-1)/(|F|-1)$.

5. All the operations used to define Tr commute with raising to the pth power.

6. We have $\mathsf{Tr}(1) = 1 + 1^q + \cdots + 1^{q^{n-1}} = 1 + 1 + \cdots + 1$, with n terms.

7. This is a special case of Proposition 3.1.1. □

3.2.f Quadratic equations in characteristic 2

Let $F = \mathbb{F}_{2^r}$ be a finite field of characteristic 2 and let $a, b, c \in F$ with $a \neq 0$. The trace function is a necessary ingredient for determining when the quadratic equation

$$ax^2 + bx + c = 0 \tag{3.4}$$

has a solution $x \in F$.

Theorem 3.2.15. *If $a, b \neq 0$ then the quadratic equation (3.4) has a solution $x \in F$ if and only if $\mathsf{Tr}_{\mathbb{F}_2}^F(ac/b^2) = 0$, in which case it has two distinct solutions. If $b = 0$ (and $a \neq 0$) then it has a unique solution.*

Proof. First consider the case $a = b = 1$. To solve the equation $x^2 + x = c$, consider the following sequence, where $\phi : F \to F$ is the linear map, $\phi(x) = x^2 - x = x^2 + x$,

$$0 \longrightarrow \mathbb{F}_2 \longrightarrow F \xrightarrow{\phi} F \xrightarrow{\mathsf{Tr}_{\mathbb{F}_2}^F} \mathbb{F}_2 \longrightarrow 0$$

Then this sequence is exact (see Definition 2.1.10): exactness at the first term is immediate and exactness at the fourth terms follows from Theorem 3.2.14. The kernel $\mathrm{Ker}(\phi) = \{0, 1\} = \mathbb{F}_2$ is 1-dimensional. It follows that the sequence is exact at the second term, and in particular, the mapping ϕ is two-to-one. Therefore the mapping ϕ has rank $r - 1$. But $\mathsf{Tr} : F \to \mathbb{F}_2$ is surjective so its kernel also has dimension $r - 1$. Since $\mathrm{Im}(\phi) \subset \mathrm{Ker}(\mathsf{Tr})$, and they have the same dimension, they must coincide. In other words, $c \in \mathrm{Ker}(\mathsf{Tr})$ if and only if $c = x^2 - x$ for some x, and in this case there are two such values of x.

	$Q(x,y)$	conditions	$N_v,\ v=0$	$N_v,\ v\neq 0$
Type I	xy	$b\neq 0,\ \mathsf{Tr}^F_{\mathbb{F}_2}(ac/b^2)=0$	$2q-1$	$q-1$
Type II	x^2	$b=0$	q	q
Type III	$h(x^2+y^2)+xy$	$b\neq 0,\ \mathsf{Tr}^F_{\mathbb{F}_2}(ac/b^2)=1$	1	q

Table 3.1: Quadratic forms in characteristic 2.

For the general case, the transformation

$$x=\frac{b}{a}y$$

converts equation (3.4) into the equation

$$\frac{b^2}{a}(y^2+y)=c$$

which therefore has a solution if and only if

$$\mathsf{Tr}^F_{\mathbb{F}_2}(ac/b^2)=0,$$

in which case it has two solutions. Finally, if $b=0$ then the equation $x^2=c/a$ has one solution because 2 is relatively prime to 2^r-1 so the mapping $F\to F$ given by $x\to x^2$ is invertible (and in fact it is an isomorphism of $\mathbb{F}_2$-vector spaces). □

Corollary 3.2.16. *Let $a,b,c\in F=\mathbb{F}_{2^r}$. Fix $h\in F$ with $\mathsf{Tr}^F_{\mathbb{F}_2}(h)=1$. Then the quadratic form*

$$Q(x,y)=ax^2+bxy+cy^2$$

can be transformed, using a linear transformation of variables, into one of the three quadratic forms in Table 3.1. The number N_v of solutions to the equation $Q(x,y)=v$ is also given.

Proof. The transformation $x\to\alpha x+\beta y$ changes Q into the quadratic form

$$a\alpha^2x^2+b\alpha xy+(a\beta^2+b\beta+c)y^2.$$

By Theorem 3.2.15 if $\mathsf{Tr}^F_{\mathbb{F}_2}(ac/b^2)=0$ then $\beta\in F$ can be chosen so that the coefficient of y^2 vanishes. Taking $\alpha=1/b$ leaves

$$Q=\frac{a}{b^2}x^2+xy.$$

Now the transformation

$$y \to \frac{a}{b^2}x + y$$

changes Q to xy. This is Type I.

If $\mathsf{Tr}_{\mathbb{F}_2}^F(ac/b^2) \neq 0$ then choose β so that $a\beta^2 + b\beta + c = b^2h^2/a$ (which is possible, by Theorem 3.2.15) and choose $\alpha = bh/a$. Then the transformation $x \to \alpha x + \beta y$ converts the quadratic form Q into the form

$$\frac{b^2h}{a}(hx^2 + xy + hy^2).$$

The further transformation

$$x \to \frac{\sqrt{a}}{b\sqrt{h}}x \quad \text{and} \quad y \to \frac{\sqrt{a}}{b\sqrt{h}}y$$

transforms Q into $hx^2 + xy + hy^2$. This is Type II. Finally, if $b = 0$ (and $a, c \neq 0$), use

$$x \to \frac{x + \sqrt{c}y}{\sqrt{a}}$$

to convert $Q(x, y)$ into x^2. This is Type III.

Counting the number N_v of solutions to $Q(x, y) = v$ is trivial for Types I and II. For Type III, if $v = 0$ then for any nonzero choice of y we need to solve for x in the equation $hx^2 + xy + hy^2 = 0$. Since $\mathsf{Tr}_{\mathbb{F}_q}^F(h) = 1$ this has no solutions. Thus $(0, 0)$ is the unique solution. If $v \neq 0$, imagine choosing y and solving for x, which will be possible if and only if $\mathsf{Tr}_{\mathbb{F}_2}^F(h^2y^2 - hv) = 0$. As y varies in F the quantity inside the trace varies among all elements of F, and $q/2$ of these have trace zero. For each such choice of y there are two distinct choices for x, for a total of q solutions. □

3.2.g Characters and exponential sums

Let F be a finite field, say, $|F| = q = p^r$ where p is a prime number. Let $F^\times$ be the group of all nonzero elements of F under multiplication and let F^+ be the group of all elements of F under addition. A character $\chi : F^+ \to \mathbb{C}^\times$ is called an *additive* character. If χ is a nontrivial additive character then every additive character is of the form $\psi(x) = \chi(Ax)$ for some element $A \in F$. (Different values of A give distinct characters, and there are $|F|$ of them, which therefore account for all additive characters.)

A character $\psi : F^\times \to \mathbb{C}^\times$ is called a *multiplicative* character of F. It is common to extend each multiplicative character $\psi : F^\times \to \mathbb{C}$ to all of F by setting $\psi(0) = 0$. If q is odd then the *quadratic character*

$$\eta(x) = \begin{cases} 1 & \text{if } x \text{ is a square} \\ -1 & \text{otherwise} \end{cases} \tag{3.5}$$

is a multiplicative character. If p is an odd prime and $0 \neq x \in \mathbb{F}_p$ then the *Legendre symbol* is

$$\left(\frac{x}{q}\right) = \eta(x).$$

Since the prime field $\mathbb{F}_p = \mathbb{Z}/(p)$ is cyclic, the additive group F^+ is isomorphic to the additive group $(\mathbb{Z}/(p))^r$, so we obtain from Section 2.3.b the notion of a Fourier transform $\widehat{f}$ of any function $f : F \to \mathbb{C}$, with respect to this additive group structure. Since the multiplicative group $F^\times$ is cyclic, we obtain a second notion of Fourier transform of any function $f : F^\times \to \mathbb{C}$. Equation (2.16) gives explicit formulae for these Fourier transforms. In this case they are sometimes called the Hadamard and Walsh transforms (respectively).

If ψ is a multiplicative character one can take its Fourier transform $\widehat{\psi}$ with respect to the additive structure to obtain the *Gauss sum*

$$\widehat{\psi}(\chi) = G(\psi, \chi) = \sum_{g \in F} \chi(g)\psi(g) = \sum_{g \in F^\times} \chi(g)\psi(g) \tag{3.6}$$

for any additive character χ. Conversely, equation (3.6) may be interpreted as the Fourier transform $\widehat{\chi}$ of the additive character χ evaluated on the multiplicative character ψ. The results in Section 2.3.b therefore give a number of simple facts concerning Gauss sums. In particular, the Fourier expansion of a multiplicative character ψ in terms of additive characters as in equation (2.13) gives

$$\psi(g) = \frac{1}{|F|} \sum_{\chi} G(\psi, \chi)\overline{\chi}(g) = \frac{1}{|F|} \sum_{\chi} G(\psi, \overline{\chi})\chi(g).$$

We state without proof the following classical *Gauss bound* and *Weil bound* (cf. [123] Section 5.2, Section 5.4) and its improvement by Carlitz and Uchiyama [22].

Theorem 3.2.17. *If χ, ψ are nontrivial additive and multiplicative $\mathbb{C}$-valued characters (respectively) of a finite field F then*

$$|G(\psi, \chi)| = \sqrt{|F|}.$$

Theorem 3.2.18. *Let $F = \mathbb{F}_{p^r}$ with p prime. Let $f \in F[x]$ be a polynomial of degree $n \geq 1$. Let χ be a nontrivial additive character of F. Suppose either (a) $\gcd(n, p) = 1$ or (b) f is not of the form $g^p - g + b$ where $g \in F[x]$ and $b \in F$. Then*

$$\left| \sum_{x \in F} \chi(f(x)) \right| \leq (n-1)\sqrt{|F|}. \tag{3.7}$$

Let ψ be a nontrivial multiplicative character of F of order $m > 1$ and let d be the number of distinct roots of f in its splitting field over F. Instead of assumptions (a) or (b) above, assume

that the monic polynomial $a^{-1}f$ (where a is the leading coefficient of f) is not the m-th power of a polynomial $g(x) \in F[x]$. Then

$$\left|\sum_{x\in F} \psi(f(x))\right| \leq (d-1)\sqrt{|F|}. \tag{3.8}$$

The following theorem of A. Weil [189], [173] combines all of the above.

Theorem 3.2.19. *Let $F = \mathbb{F}_q$ be a finite field. Let ψ be a nontrivial multiplicative character of order d. Let χ be a nontrivial additive character. Let $f(x), g(x) \in F[x]$ be polynomials with $n = \deg(g)$. Assume that $f(x)$ has m distinct roots in F, and further assume that $\gcd(d, \deg(f)) = \gcd(q, \deg(g)) = 1$. Then*

$$\left|\sum_{x\in F} \psi(f(x))\chi(g(x))\right| \leq (m+n-1)\sqrt{|F|}.$$

For polynomials in several variables there is the following bound of Deligne [36] (Theorem 8.4):

Theorem 3.2.20. *Let F be a finite field and let χ be a nontrivial additive character. Let $f(x_1, x_2, \cdots, x_n)$ be a polynomial of degree m. Assume m is relatively prime to $|F|$. Assume also that the homogeneous part of f of maximal degree ($= m$) is nonsingular, when it is considered as a form over the algebraic closure of F. Then*

$$\left|\sum_{x\in F^n} \chi(f(x))\right| \leq \left((m-1)\sqrt{|F|}\right)^n.$$

3.2.h The Discrete Fourier transform

While the (usual) Fourier transform involves complex valued functions, the discrete Fourier transform involves functions with values in a finite field F. It is defined in a manner completely analogous to equation (2.16), provided the field F contains all the required roots of unity. (For cyclic groups $G = \mathbb{Z}/(N)$, this assumption may be relaxed, see Section 18.4.b.)

Let G be a finite Abelian group and let $N \in \mathbb{Z}$ be its *characteristic*, that is, the smallest integer such that $0 = x + x + \cdots + x$ (N times) for all $x \in G$. (Then N divides $|G|$ and it is the order of the largest cyclic subgroup of G.) Let $F = \mathbb{F}_q$ be a finite field and suppose that N and q are relatively prime. Let $d = \mathrm{ord}_N(q)$ and let $E = \mathbb{F}_{q^d}$. By Theorem 3.2.13 the field E is the smallest field extension of F that contains all the N-th roots of unity. (In what follows, the field E may be replaced by any larger field.)

Continue to assume that $N = \mathrm{char}(G)$ and $\mathrm{char}(E)$ are relatively prime. The group of *discrete characters* $\widehat{G}$ is the set of (group) homomorphisms $\psi : G \to E^\times$. It forms a group under multiplication of characters, $(\chi\psi)(g) = \chi(g)\psi(g)$. If $G = \mathbb{Z}/(N)$ then $\widehat{G}$ is also cyclic of order N and is generated by the primitive character $\chi(x) = b^x$ where $b \in E$ is a primitive N-th root of unity. If $G = G_1 \times G_2$ then $\widehat{G} = \widehat{G}_1 \times \widehat{G}_2$. It follows that the group of characters $\widehat{G}$ is a finite Abelian group that is abstractly isomorphic to G and in particular that $|\widehat{G}| = |G|$. The proof of Proposition 2.3.2 works here too and we obtain the following.

Lemma 3.2.21. *Let G be a finite Abelian group characteristic N, $F = \mathbb{F}_q$, and $E = \mathbf{q^d}$ where $d = ord_N(q)$. Then the following hold.*

1. *If $\chi : G \to E^\times$ is a nontrivial character then $\sum_{g\in G} \chi(g) = 0$.*
2. *If $0 \neq g \in G$ then $\sum_{\chi\in\widehat{G}} \chi(g) = 0$.*
3. *If $g \neq h \in G$ then there exists $\chi \in \widehat{G}$ such that $\chi(g) \neq \chi(h)$.*
4. *If $\psi \neq \chi \in \widehat{G}$ then $\sum_{g\in G} \psi(g)\chi^{-1}(g) = 0$.*
5. *If $g \neq h \in G$ then $\sum_{\chi\in\widehat{G}} \chi(g)\chi^{-1}(h) = 0$.* □

With $G, F, E, \widehat{G}$ as above, for any $f : G \to E$ define its *Fourier transform* $\widehat{f} : \widehat{G} \to E$ by

$$\widehat{f}(\chi) = \sum_{g\in G} \chi(g)f(g).$$

Then the *convolution formula* (2.14) and the *Fourier inversion formula* (2.13) hold:

$$\widehat{f}\cdot\widehat{g} = \widehat{f * G} \quad \text{and} \quad f(g) = \frac{1}{|G|}\sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\chi^{-1}(g)$$

with the same proof as in Section 2.3.b. The proof in Section 2.3.b gives a weak analog to Parseval's equation,

$$\sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\widehat{f}(\chi^{-1}) = \sum_{g\in G} f(g)^2 \quad \text{and} \quad \sum_{\chi\in\widehat{G}} \widehat{f}(\chi)^2 = \sum_{g\in G} f(g)f(-g).$$

Although the discrete Fourier transform and the complex Fourier transform are entirely parallel in their definition and properties, there is no apparent relation between them.

If $G = \mathbb{Z}/(N)$, then a choice of primitive N-th root of unity $b \in E$ determines an isomorphism $G \cong \widehat{G}$ which takes $1 \in \mathbb{Z}/(N)$ to the character χ_1 with $\chi_1(k) = b^k$. The other characters are powers of this one: $\chi_m(k) = b^{mk}$. If $f : \mathbb{Z}/(N) \to E$ is a function, then its discrete Fourier transform may be considered as a function $\widehat{f} : \mathbb{Z}/(N) \to E$ by writing $\widehat{f}(m)$ rather than $\widehat{f}(\chi_m)$. In other words,

$$\widehat{f}(m) = \sum_{k=0}^{N-1} b^{mk} f(k) \quad \text{and} \quad f(g) = \frac{1}{N}\sum_{m=1}^{N-1} \widehat{f}(m)b^{-mg}.$$

3.3 Quadratic forms over a finite field

3.3.a Quadratic forms and their classification

The standard reference for this section is Lidl and Niederreiter's book on finite fields [123]. Let $F = \mathbb{F}_q$ be a finite field with q elements. A quadratic form in n variables over F is a polynomial $Q(x_1, x_2, \cdots, x_n)$ in n variables such that each term has degree two, that is,

$$Q(x_1, \cdots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j.$$

It follows that $Q(cx) = c^2 Q(x)$ for any $c \in F$. If q is odd then the function $Q(x)$ may be expressed as $Q(x) = x^t A x$ where A is the symmetric matrix $A_{ij} = \frac{1}{2}(a_{ij} + a_{ji})$ for $i \neq j$ and $A_{ii} = a_{ii}$. Consequently there is an associated bilinear form $B(x, y) = x^t A$ (here we are thinking of the vectors x and y as column vectors).

If q is even then every quadratic form Q may be expressed as $Q(x) = x^t A x$ where the matrix A is not necessarily symmetric, and where the transpose matrix A^t gives the same quadratic form. If $M : F^n \to F^n$ is an invertible $n \times n$ matrix representing a linear change of coordinates, then the matrix of the quadratic form with respect to the new coordinates is $M^t A M$. In particular, the determinant of A changes by the factor $\det(M)^2$.

The *rank* of a quadratic form Q is the smallest integer m such that there exists a linear change of variables $(x_1, \cdots, x_n) \to (y_1, \cdots, y_n)$ so that the resulting quadratic form involves only the variables $y_1, y_2, \cdots, y_m$. A quadratic form in n variables is *nondegenerate* if its rank is n. If $Q : F^n \to F$ is a quadratic form then there exists a *maximal nondegenerate subspace* that is, a subspace $V \subset F^n$ such that $\dim(V) = \text{rank}(Q)$ and so that Q restricted to V is nondegenerate. A vector $w \in F^n$ is in the *kernel* of Q if $Q(w + x) = Q(x)$ for all $x \in F^n$. The kernel of Q is a vector subspace of F^n and it is complementary to any maximal nondegenerate subspace $V \subset F^n$, meaning that $V \cap W = \{0\}$ and $V + W = F^n$.

If F is a field of characteristic 2 then every element is a square. But if the characteristic of F is odd then half the elements are squares, and the *quadratic character* $\eta : F^\times \to \{0, 1\}$ is defined by $\eta(x) = 1$ if $x = a^2$ for some $a \in F$ and $\eta(x) = -1$ otherwise. (It is customary to define $\eta(0) = 0$ as this convention can often be used to simplify various formulae.) Denote by $\Delta(Q)$ the determinant of the restriction of Q to a maximal nondegenerate subspace; it is well defined up to multiplication by a square in $\mathbb{F}$. If $\text{rank}(Q) = m$ define

$$\Delta'(Q) = \begin{cases} (-1)^{m/2}\Delta(Q) & \text{if } m \text{ is even,} \\ (-1)^{(m-1)/2}\Delta(Q) & \text{if } m \text{ is odd.} \end{cases} \tag{3.9}$$

If the characteristic of F is odd, the properties of the quadratic form Q depend on whether or not the element $\Delta'(Q) \in F$ is a square, that is, whether $\eta(\Delta'(Q))$ is $+1$ or -1.

The following theorem gives the classification of quadratic forms over the finite field $F = \mathbb{F}_q$ of arbitrary characteristic. Although the proofs are not difficult, they are tedious and they can be found in [123]. (See Theorem 3.2.15 for the case $m = 2$ and characteristic 2.) In this classification, the symbol B_m denotes the quadratic form

$$B_m(x_1, \cdots, x_n) = x_1x_2 + x_3x_4 + \cdots + x_{m-1}x_m. \tag{3.10}$$

Theorem 3.3.1. ([123] Thm 6.30) *Suppose Q is a quadratic form of rank m in $n \geq m$ variables over a field $F = \mathbb{F}_q$. If q is even, fix an element $h \in F$ such that $\mathsf{Tr}_{\mathbb{F}_2}^F(h) = 1$. Then there is a linear change of variables so that Q is one of the quadratic forms listed in Table 3.2.*

		q **even**
Type I	(m even)	$Q(x) = B_m(x)$
Type II	(m odd)	$Q(x) = B_{m-1}(x) + x_m^2$
Type III	(m even)	$Q(x) = B_{m-2}(x) + x_{m-1}x_m + h(x_{m-1}^2 + x_m^2)$

q **odd**
$Q(x) = a_1x_1^2 + a_2x_2^2 + \cdots + a_mx_m^2$, $a_i \neq 0$

Table 3.2: Classification of quadratic forms over $\mathbb{F}_q$

(If q is odd and if $b \in F$ is a fixed non-square then $Q(x)$ can even be reduced to one of the two quadratic forms $Q(x) = x_1^2 + \cdots + x_{m-1}^2 + ax_m^2$ where $a = 1$ or $a = b$, but for most purposes the above diagonal form suffices.)

3.3.b Solutions to $Q(x) + L(x) = u$

Let F be a finite field. In this section we wish to count the number of solutions to the equation $Q(x) + L(x) = u$ where $Q : F^n \to F$ is a quadratic form and $L : F^n \to F$ is a linear mapping. This calculation is the central step in determining the cross-correlation of m-sequences, geometric sequences, GMW sequences, and Gold sequences. In order to simplify the presentation we define the following function $\nu : F \to \{-1, q-1\}$ by

$$\nu(x) = \begin{cases} -1 & \text{if } x \neq 0 \\ q-1 & \text{if } x = 0 \end{cases}$$

With $\Delta'(Q) = \pm\Delta(Q)$ as defined in equation (3.9), the following theorem counts the number of solutions to the equation $Q(x) = u$. The proof is not difficult but it is tedious and it will be omitted. We use the convention that $\eta(0) = 0$.

Theorem 3.3.2. ([123] Thm. 6.26, 6.27, 6.31) *Let Q be a quadratic form of rank m in n variables over a field $F = \mathbb{F}_q$. Let $u \in F$. Then the number N of solutions to the equation $Q(x_1, x_2, \cdots, x_n) = u$ is given in Table 3.3.*

	q **even**	
Type I	(m even)	$N = q^{n-1} + \nu(u)q^{n-1-m/2}$
Type II	(m odd)	$N = q^{n-1}$
Type III	(m even)	$N = q^{n-1} - \nu(u)q^{n-1-m/2}$

q **odd**	
(m odd)	$N = q^{n-1} + \eta(u)\eta(\Delta')q^{n-(m+1)/2}$
(m even)	$N = q^{n-1} + \nu(u)\eta(\Delta')q^{n-1-m/2}$

Table 3.3: Number of solutions to $Q(x) = u$

We now use this result to describe the number of solutions $x \in F^n$ of the equation

$$Q(x) + L(x) = u \tag{3.11}$$

where Q is a quadratic form and L is a linear function. We first show that the case $\text{rank}(Q) < n$ can be reduced to the case when Q has maximal rank, then we count the number of solutions to (3.11) assuming Q has maximal rank. The answer shows, in particular, that if $Q \neq 0$ then the function $Q(x) + L(x)$ cannot be identically zero.

Proposition 3.3.3. *Let $F = \mathbb{F}_q$ be a finite field, let $Q : F^n \to F$ be a quadratic form with $m = \text{rank}(Q) < n$ and let $L : F^n \to F$ be a nonzero (hence, surjective) linear mapping. Let $V \subset F^n$ be a maximal subspace on which Q is nondegenerate. Then the number N_u of solutions to the equation $Q(x) + L(x) = u$ is*

$$N_u = \begin{cases} q^{n-1} & \text{if } \mathrm{Ker(Q)} \not\subset \mathrm{Ker(L)} \\ q^{n-m}N'_u & \text{if } \mathrm{Ker(Q)} \subset \mathrm{Ker(L)} \end{cases}$$

where N'_u is the number of solutions x to equation (3.11) with $x \in V$.

Proof. This follows immediately from the direct sum decomposition $F^n \cong \mathrm{Ker(Q)} \oplus \mathrm{V}$ and the fact that L can be written as the sum of a linear function L_1 on Ker(Q) and a linear function L_2 on V. □

Theorem 3.3.4. *Let $u \in \mathbb{F}_q$. Let $Q : \mathbb{F}_q^m \to \mathbb{F}_q$ be one of the standard quadratic forms of (maximal) rank m listed in the classification, Theorem 3.3.1. Let $L(x) = \sum_{i=1}^m c_i x_i$ be a linear function. Let $N = N_u$ denote the number of elements $x \in \mathbb{F}_q^m$ so that $Q(x) + L(x) = u$. Then N_u is given in Table 3.4, where $\tau(c,u) = 0$ if $c_m = 0$, otherwise*

$$\tau(c,u) = (-1)^{\mathsf{Tr}_{\mathbb{F}_2}^F\left(\frac{u+B_{m-1}(c)}{c_m^2}\right)} = \pm 1; \tag{3.12}$$

and where $R = R(Q,c)$ is given in equation (3.13).

	q **even**	
Type I	(m even)	$N = q^{m-1} + \nu(u+Q(c))q^{m/2-1}$
Type II	(m odd)	$N = q^{m-1} + \tau(c,u)q^{(m-1)/2}$
Type III	(m even)	$N = q^{m-1} - \nu(u+Q(c))q^{m/2-1}$

q **odd**	
(m odd)	$N = q^{m-1} + \eta(u+R)\eta(\Delta')q^{(m-1)/2}$
(m even)	$N = q^{m-1} + \nu(u+R)\eta(\Delta')q^{m/2-1}$

Table 3.4: Number of solutions to $Q(x) + L(x) = u$

Proof. When q is even, the results for Type I and III follow from Theorem 3.3.2 after an affine change of coordinates which replaces x_1 by $x_1 + c_2$, and x_2 by $x_2 + c_1$, etc. This eliminates the linear terms and replaces u with $u + Q(c)$. In the case of Type II ($Q(x) = B_{m-1}(x) + x_m^2$), the same trick eliminates the first $m-1$ linear terms and replaces u with $u + B_{m-1}(c)$. If $c_m = 0$ then we are done: there are q^{n-1} solutions as in Theorem 3.3.2. But if $c_m \neq 0$ we are left with the equation

$$B_{m-1}(x) = x_m^2 + c_m x_m + u + B_{m-1}(c).$$

The number of solutions $(x_1, \cdots, x_{m-1})$ to this equation depends on whether or not the right side vanishes. By Theorem 3.2.15, this in turn depends on

$$t = \mathsf{Tr}_{\mathbb{F}_2}^F\left(\frac{u + B_{m-1}(c)}{c_m^2}\right).$$

If $t = 0$ then there are two values of x_m for which the right side vanishes, and $q-2$ values for which the right side is nonzero. This gives

$$N = 2\left(q^{m-2} + (q-1)q^{(m-1)/2-1}\right) + (q-2)\left(q^{m-2} - q^{(m-1)/2-1}\right) = q^{m-1} + q^{(m-1)/2}.$$

If $t \neq 0$ then the right side never vanishes, so choosing x_m arbitrarily and then choosing the other variables $x_1, \cdots, x_{m-1}$ gives

$$N = q\left(q^{m-2} - q^{(m-1)/2-1}\right) = q^{m-1} - q^{(m-1)/2}.$$

If q is odd we may assume $Q(x) = \sum_{i=1}^{m} a_i x_i^2$ with $a_i \neq 0$ for all i. The substitution $x_i \to y_i - b_i/2a_i$ converts the equation $Q(x) + L(x) = u$ into the equation $Q(y) = u + R$ where

$$R = \text{“}Q\left(\frac{c}{2a}\right)\text{”} = a_1\left(\frac{c_1}{2a_1}\right)^2 + a_2\left(\frac{c_2}{2a_2}\right)^2 + \cdots + a_m\left(\frac{c_m}{2a_m}\right)^2. \tag{3.13}$$

By Theorem 3.3.2 the number of solutions to this equation is

$$N = q^{m-1} + \begin{cases} \eta(u+R)\eta(\Delta')q^{(m-1)/2} & \text{if } m \text{ is odd} \\ \nu(u+R)\eta(\Delta')q^{m/2-1} & \text{if } m \text{ is even.} \end{cases}$$

This completes the proof of Theorem 3.3.4. □

3.3.c The Quadratic form $\mathsf{Tr}(cx^d)$ for $d = q^i + q^j$

One important source of quadratic forms is the following. Let $F = \mathbb{F}_q \subset L = \mathbb{F}_{q^n}$ be finite fields. Let $c \in L$. Let $d = 1 + q^i$. Then, as shown below, the function $Q : L \to F$ defined by

$$Q(x) = \mathsf{Tr}_F^L(cx^d) \tag{3.14}$$

is a quadratic form. We may assume that $i < n$ because $x^{q^n} = x$ for all $x \in L$. We remark that the function $\mathsf{Tr}_F^L(cx^{d'})$ where $d' = q^j + q^i$ is no more general than (3.14), because the change of variable $y = x^{q^j}$ converts this form into $\mathsf{Tr}_F^L(cx^e)$ where $e = 1 + q^{i-j}$.

Theorem 3.3.5. *If $d = 1 + q^i$ then the function $Q(x) = \mathsf{Tr}_F^L(cx^d)$ is a quadratic form over F. Let $g = \gcd(i, n)$. The rank of this quadratic form is given in Table 3.5, where $g = \gcd(n, i)$, $e = 1+q^g$, $\eta = \eta(\Delta')$ as in equation (3.9), and $c = s^d$ means that c is a d-th power of some element $s \in L$.*

The following lemma is used in the proof of Theorem 3.3.5.

Lemma 3.3.6. *Let $n, j \geq 1$ and $b \geq 2$. The greatest common divisor $g = \gcd(n, j)$ is given in Table 3.6.*

Proof. The first statement follows from a simple calculation. It is possible to use the identity $b^{2k} - 1 = (b^k - 1)(b^k + 1)$ (with $k = j$ and $k = n$ respectively) to deduce the second and third statements from the first statement. □

q **even**				
Conditions			Type	Rank
n/g even	$n/2g$ odd	$c = s^d$	I	$n - 2g$
		$c \neq s^d$	III	n
	$n/2g$ even	$c = s^d$	III	$n - 2g$
		$c \neq s^d$	I	n
n/g odd			II	$n - g + 1$

q **odd**				
Conditions			Type	Rank
n/g even	$n/2g$ odd	$c^2 = s^e$ $c \neq s^e$	$\eta = -1$	$n - 2g$
		otherwise	$\eta = 1$	n
	$n/2g$ even	$c = s^e$	$\eta = 1$	$n - 2g$
		$c \neq s^e$	$\eta = -1$	n
n/g odd	n odd			n
	n even			n

Table 3.5: The quadratic form $\mathsf{Tr}(cx^d)$, $d = 1 + q^i$

$$\gcd(b^n - 1, b^j - 1) = b^g - 1$$

$$\gcd(b^n - 1, b^j + 1) = \begin{cases} 1 + b^g & \text{if } n/g \text{ is even} \\ 2 & \text{if } n/g \text{ is odd and } b \text{ is odd} \\ 1 & \text{if } n/g \text{ is odd and } b \text{ is even} \end{cases}$$

$$\gcd(b^n + 1, b^j + 1) = \begin{cases} 1 + b^g & \text{if } n/g \text{ is odd and } j/g \text{ is odd} \\ 2 & \text{if } n/g \text{ is even or } j/g \text{ is even, and } b \text{ is odd} \\ 1 & \text{if } n/g \text{ is even or } j/g \text{ is even, and } b \text{ is even} \end{cases}$$

Table 3.6: $\gcd(b^n \pm 1, b^j \pm 1)$

Proof of Theorem 3.3.5 Let $e_1, e_2, \cdots, e_r$ be a basis for L as a vector space over F. Let

$x = a_1e_1 + \cdots + a_re_r \in L$. Then

$$\begin{aligned}
\mathsf{Tr}_F^L(cx^{1+q^i}) &= \mathsf{Tr}_F^L\left[c\left(\sum_{h=1}^r a_he_h\right)\left(\sum_{h=1}^r a_he_h\right)^{q^i}\right] \\
&= \mathsf{Tr}_F^L\left[c\left(\sum_{h=1}^r a_he_h\right)\left(\sum_{h=1}^r a_he_h^{q_i}\right)\right] \\
&= \sum_{h=1}^r\sum_{k=1}^r b_{hk}a_ha_k
\end{aligned}$$

where

$$b_{hk} = \mathsf{Tr}_F^L\left(ce_he_k^{q^i}\right).$$

This is a quadratic form. In order to determine its rank we start by determining its kernel $W = \mathrm{Ker(Q)}$. Equating

$$\mathsf{Tr}_F^L(c(y+w)^{1+q^i}) = \mathsf{Tr}_F^L(cy^{1+q^i})$$

gives

$$\mathsf{Tr}_F^L(cw^{1+q^i} + cyw^{q^i} + cy^{q^i}w) = 0.$$

Hence $w \in W$ if and only if

$$\mathsf{Tr}_F^L(cw^{1+q^i}) = 0 \tag{3.15}$$

and

$$\mathsf{Tr}_F^L(cwy^{q^i}) = -\mathsf{Tr}_F^L(cw^{q^i}y) \text{ for every } y \in L.$$

Since $\mathsf{Tr}_F^L(x^q) = \mathsf{Tr}_F^L(x)$ the right side of this equation is unchanged if we raise its argument to the power q^i, which gives

$$\mathsf{Tr}_F^L((cw + c^{q^i}w^{q^{2i}})y^{q^i}) = 0$$

for all $y \in L$, so

$$cw = -c^{q^i}w^{q^2i}$$

or, assuming $w \neq 0$,

$$c^{q^i-1}w^{q^{2i}-1} = -1. \tag{3.16}$$

Let

$$z = cw^{1+q^i}.$$

Then equation (3.16) is equivalent to:

$$z^{q^i-1} = -1. \tag{3.17}$$

At this point we must separate the cases, when q is even or odd. From here on we let $g = \gcd(n, i)$.

q **even:** Here,

$$z^{q^i-1} = -1 = 1$$

so an element $w \in L$ is in Ker(Q) if and only if:

$$z = cw^{1+q^i} \in \mathbb{F}_{q^i} \cap \mathbb{F}_{q^n} = \mathbb{F}_{q^g} \quad \text{and} \quad \mathsf{Tr}_F^L(z) = 0. \tag{3.18}$$

The set

$$\mathbb{F}_{q^g} \cap \mathrm{Ker}(\mathsf{Tr}_{\mathrm{F}}^{\mathrm{L}})$$

is either all of $\mathbb{F}_{q^g}$, which is a vector space of dimension g, or it is a hyperplane in $\mathbb{F}_{q^g}$, which therefore has dimension $g-1$. It remains to determine the number of elements $w \in L$ that satisfy equation (3.17) with

$$z \in \mathbb{F}_{q^g} \cap \mathrm{Ker}(\mathsf{Tr}_{\mathrm{F}}^{\mathrm{L}}).$$

We claim that $W = \mathrm{Ker}(\mathrm{Q})$ is nonzero if and only if there exists $s \in L$ so that $s^d = c$. That is, so that

$$s^{1+q^i} = c. \tag{3.19}$$

First, suppose such an s exists. Then every element

$$z' \in \mathbb{F}_{q^g} \cap \mathrm{Ker}\left(\mathsf{Tr}_{\mathrm{K}}^{\mathrm{L}}\right)$$

gives rise to an element $w' \in \mathrm{Ker}(\mathrm{Q})$ as follows. Since $1+q^i$ is relatively prime to q^i-1 and hence also to q^g-1 there exists $h' \in \mathbb{F}_{q^g}$ such that

$$(h')^{1+q^i} = z'.$$

Set $w' = h'/s \in L$. It follows that

$$z' = c(w')^{1+q^i}$$

so w' satisfies equations (3.15) and (3.16).

We remark that if n/g is odd then $1+q^i$ is relatively prime to q^n-1 so such an element s satisfying equation (3.19) exists, hence the dimension of $W = \mathrm{Ker}(\mathrm{Q})$ is $g-1$. If n/g is even and if such an element s exists, then

$$\mathbb{F}_{q^g} \cap \mathrm{Ker}(\mathsf{Tr}_{\mathrm{K}}^{\mathrm{L}}) = \mathbb{F}_{\mathrm{q^g}};$$

for if $z' \in \mathbb{F}_{q^g}$ and n/g is even then (writing $E = \mathbb{F}_{q^g}$ to ease notation),

$$\mathsf{Tr}_F^L(z') = \mathsf{Tr}_F^E \mathsf{Tr}_E^L(z') = (n/g)\mathsf{Tr}_F^E(z') = 0. \tag{3.20}$$

Thus, in this case, the dimension of $W = \mathrm{Ker}(\mathrm{Q})$ is g.

Next, we prove the converse: suppose there exists $w \neq 0 \in W$; we claim there exists $s \in L$ satisfying equation (3.19). We may suppose that n/g is even (since the odd case was handled above). We find an element $u \neq 0 \in L$ so that u^{q^i+1} is a primitive element of $\mathbb{F}_{q^g}$. This will suffice because $z = cw^{q^i+1} \in \mathbb{F}_{q^g}$ so there exists m with $z = u^{(q^i+1)m}$, hence $c = (u^m/w)^{q^i+1}$.

The element $u \in L = \mathbb{F}_{q^n}$ should be taken to be a primitive element of the sub-field $\mathbb{F}_{q^{2g}}$ (which is contained in L since n/g is even). Then

$$u^{q^{2g}-1} = u^{(q^g-1)(q^g+1)} = 1.$$

Moreover, $u^{q^i+1} \in \mathbb{F}_{q^g}$ for the following reason: i/g is odd so by Lemma 3.3.6, q^g+1 divides q^i+1. Therefore

$$u^{(q^g-1)(q^i+1)} = 1.$$

Finally, u^{q^i+1} is primitive in $\mathbb{F}_{q^g}$ because q^i+1 and q^g-1 are relatively prime. We remark that as u varies within the field $\mathbb{F}_{q^{2g}}$ the element uw varies within Ker(Q) because

$$c(uw)^{1+q^i} \in \mathbb{F}_{q^i} \cap \mathbb{F}_{q^n}$$

and the trace condition is satisfied by equation (3.20). So in this case, $\dim(\mathrm{Ker}(\mathrm{Q})) = 2\mathrm{g}$.

q odd: In this case

$$z^{q^i-1} = -1$$

so $z^2 \in \mathbb{F}_{q^i}$. Hence

$$z \in \mathbb{F}_{q^{2i}} \cap \mathbb{F}_{q^n} = \mathbb{F}_{q^{\gcd(n,2i)}}.$$

If n/g is odd then $\gcd(n, 2i) = \gcd(n, i) = g$, so $z \in \mathbb{F}_{q^g} \subset \mathbb{F}_{q^i}$. Hence

$$z^{q^i-1} = 1, \tag{3.21}$$

which is a contradiction. Therefore $w = 0$, so the quadratic form Q has maximal rank, n.

Now suppose n/g is even, so $\gcd(n, 2i) = 2g$. Equation (3.21) holds, so

$$z^2 \in \mathbb{F}_{q^i} \cap \mathbb{F}_{q^n} = \mathbb{F}_{q^g},$$

so

$$z^{q^g-1} = \pm 1.$$

The $+1$ is not possible, so

$$z^{q^g-1} = -1. \tag{3.22}$$

We claim that

$$\text{if } x \in L = \mathbb{F}_{q^n} \text{ and } x^{q^g-1} = -1 \text{ then } \mathsf{Tr}_F^L(x) = 0.$$

For $x^{q^g} = -x$, let $T = x + x^q + \cdots + x^{q^{g-1}}$. Then

$$\mathsf{Tr}_F^L(x) = \mathsf{Tr}_F^E \mathsf{Tr}_E^L(x) = T - T + T - T \cdots \pm T$$

(where $E = \mathbb{F}_{q^g}$) and there are n/g terms (an even number), so this last sum vanishes.

It follows that (when n/g is even), $w \in \mathrm{Ker(Q)}$ if and only if

$$z = cw^{1+q^i}$$

satisfies equation (3.22). Thus, if $a \in \mathbb{F}_{q^{2g}}$ and $w \in \mathrm{Ker(Q)}$ then $aw \in \mathrm{Ker(Q)}$. If $v, w \neq 0 \in \mathrm{Ker(Q)}$ then

$$(v/w)^{(1+q^i)(q^i-1)} = 1$$

so

$$v/w \in \mathbb{F}_{q^{2i}} \cap \mathbb{F}_{q^n} = \mathbb{F}_{q^{2g}}.$$

In summary, either $\mathrm{Ker(Q)} = \{0\}$ or $\mathrm{Ker(Q)}$ has dimension $2g$.

It remains to determine when $\mathrm{Ker(Q)}$ has a nonzero element. Let $\alpha \in \mathbb{F}_{q^n}$ be a primitive element and set $c = \alpha^\ell$. Equation (3.22) becomes

$$\alpha^\ell w^{(1+q^i)} = \alpha^{(q^n-1)/2(q^g-1)}.$$

So we search for $w = \alpha^r$ such that $\alpha^{\ell + r(1+q^i)} = \alpha^{(q^n-1)/2(q^g-1)}$ or

$$\ell \equiv \frac{q^n - 1}{2(q^g - 1)} \pmod{q^g + 1}$$

since $\gcd(q^i + 1, q^n - 1) = q^g + 1$ in this case. If $n/(2g)$ is even then $q^g + 1$ divides $(q^n - 1)/2(q^g - 1)$ hence $w \neq 0$ exists if and only if $\ell \equiv 0 \pmod{q^g + 1}$, which is to say that $c = s^e$ for some $s \in L$, where $e = q^g + 1$. If $n/(2g)$ is odd then

$$\frac{q^n - 1}{2(q^g - 1)} \equiv \frac{q^g + 1}{2} \pmod{q^g + 1}$$

so $w \neq 0$ exists if and only if $\ell \equiv (q^g + 1)/2 \pmod{q^g + 1}$.

Determining $\eta(\Delta')$ We do not know $\eta(\Delta')$ when q is odd and n/g is odd. But if q is odd and n/g is even, it is possible to determine when $\Delta'(Q)$ is a square because this is detected (according to Theorem 3.3.2) by the number $|Z|$ of nonzero solutions Z to the equation $Q(x) = 0$. If $x \in Z$ and $0 \neq a \in L$ and if $a^{1+q^i} \in F$ then $a^{1+q^i} x \in Z$. Hence the group

$$G = \{a \in L : \ a^{1+q^i} \in F\} = \{a \in L : \ a^{(1+q^i)(q-1)} = 1\}$$

acts freely on Z so $|G| = (1+q^g)(q-1)$ divides $|Z|$. If n/g is even this says,

$$\frac{|G|}{q-1} = (q^g+1) \text{ divides } \frac{|Z|}{q-1} = \frac{q^n-1}{q-1} - \frac{q^{n-1}}{q^{\mathrm{rank}(Q)/2}}\left(q^{\mathrm{rank}(Q)/2} - \eta(\Delta'(Q))\right)$$

with $\eta(\Delta'(Q)) = \pm 1$. Together with the rank of Q determined above and the divisibility properties in Lemma 3.3.6, this gives the values of $\eta(\Delta'(Q))$ listed in Theorem 3.3.5. □

Corollary 3.3.7. *Let $F = \mathbb{F}_q \subset L = \mathbb{F}_{q^n}$ be finite fields, let $d = 1+q^i$, let $A, B \in L$ with $B \neq 0$, and let $F : L \to F$ be the function*

$$F(x) = \mathsf{Tr}_F^L(Ax + Bx^d).$$

Then F is identically zero if and only if the following conditions hold.

1. *$A = 0$.*
2. *n is even and $\gcd(i,n) = n/2$.*
3. *$\mathsf{Tr}_E^L(B) = 0$ where $E = \mathbb{F}_{q^{n/2}}$ (so that $F \subset E \subset L$.)*

Proof. We may assume that $i < n$ (since $x^{q^n} = x$ for all $x \in L$), so the second condition is equivalent to $i = n/2$. If the three conditions hold then $d = (|L|-1)/(|E|-1)$ so $x^d \in E$ for any $x \in L$. Therefore

$$\mathsf{Tr}_F^E\left(\mathsf{Tr}_E^L(Bx^d)\right) = \mathsf{Tr}_F^E\left(x^d \mathsf{Tr}_E^L(B)\right) = 0.$$

To prove the converse, let $L(x) = \mathsf{Tr}_F^L(Ax)$ and $Q(x) = \mathsf{Tr}_F^L(Bx^d)$. Suppose $F(x) = L(x) + Q(x)$ is identically zero. If $A \neq 0$ then Proposition 3.3.3 implies that Ker(Q) $\subset$ Ker(L) so Q is a non-vanishing quadratic form of some rank $m > 0$. So there exists a subspace $V \subset E$ whose F-dimension is m, such that the restriction of Q to V is non-degenerate. But $N_0 = q^n$ since F is identically zero, so in the notation of Proposition 3.3.3, $N_0' = q^m$, which is to say that the restriction of Q to V is zero. This is a contradiction. Therefore $A = 0$ which proves (1). Therefore the quadratic form Q has rank zero. By theorem 3.3.5 (and Table 3.5) it follows that $n = 2g$ where $g = \gcd(i,n)$, which proves (2). Thus $g = i = n/2$. To prove (3) we must consider two cases, depending on the parity of q. Let $E = \mathbb{F}_{q^g}$.

First suppose q is even. From Table 3.5 we see that $B = s^d$ for some $s \in E$. Therefore

$$\mathsf{Tr}_E^L(B) = s^d + s^{dq^g} = s^{1+q^g} + s^{(1+q^g)q^g} = s^{1+q^g} + s^{q^g+1} = 0.$$

Now suppose q is odd. From Table 3.5 we see (since $n/2g = 1$ is odd) that

$$B^2 = s^{1+q^g}$$

(for some $s \in L$) but B cannot be so expressed. Therefore

$$B^{q^g-1} = -1$$

since its square is

$$s^{(1+q^g)(q^g-1)} = 1.$$

In summary,

$$\mathsf{Tr}_E^L(B) = B + B^{q^g-1+1} = B - B = 0. \qquad \square$$

3.4 Algebraic number fields

3.4.a Basic properties

So far our examples of fields have consisted of finite fields and the familiar fields $\mathbb{Q}$, the rational numbers, $\mathbb{R}$, the real numbers, and $\mathbb{C}$, the complex numbers. Recall that we we obtain the various finite fields of characteristic $p > 0$ from the prime field $\mathbb{F}_p$ by constructing the quotient $\mathbb{F}_p[x]/(f(x))$ where $f(x)$ is an irreducible polynomial. We can think of this construction as adjoining a root (the variable x) of $f(x)$ to the field $\mathbb{F}_p$. Similarly, we obtain the complex numbers from the real numbers by adjoining a root of the polynomial $x^2 + 1$.

In this section we study a class of fields, called *algebraic number fields* that are obtained in the same way from the rational numbers. For the most part we omit proofs and leave the interested reader to find them in other references.

Definition 3.4.1. *An algebraic number field E is a finite extension of the rational numbers $\mathbb{Q}$.*

This means that E is a field that contains $\mathbb{Q}$ and that as a vector space over $\mathbb{Q}$ it is finite dimensional.

A complex number $a \in \mathbb{C}$ is *algebraic over* $\mathbb{Q}$, or simply *algebraic*, if it is a root of some polynomial $f(x) \in \mathbb{Q}[x]$ with coefficients in $\mathbb{Q}$. As in Theorem 2.4.11, there exists a unique monic minimal polynomial $f(x) \in \mathbb{Q}[x]$, irreducible in $\mathbb{Q}[x]$, such that $f(a) = 0$. If $\mathbb{Q}(a) \subset \mathbb{C}$ denotes the smallest field that contains both $\mathbb{Q}$ and a, then the mapping $\mathbb{Q}[x] \to \mathbb{Q}(a)$ which takes x to a induces an isomorphism

$$\mathbb{Q}[x]/(f) \to \mathbb{Q}(a),$$

where f is the minimal polynomial of a. The proof is left as an exercise. An important result is the following:

Theorem 3.4.2. *Suppose that E and F are algebraic number fields with $F \subseteq E$. Then there is an element $a \in E$ such that $E = F(a)$. In particular, every algebraic number field is of the form $\mathbb{Q}(a)$ for some algebraic number a.*

A field F is *algebraically closed* if every polynomial with coefficients in F splits as a product of linear factors. Every field is contained in an algebraically closed field, and any two minimal algebraically closed fields containing a given field F are isomorphic. Thus in general we may speak of *the* algebraic closure of a field F.

For example, $\mathbb{C}$ is algebraically closed. The set $\overline{\mathbb{Q}}$ of all algebraic numbers over $\mathbb{Q}$ is an algebraically closed subfield of $\mathbb{C}$, and we shall refer to this particular field as the algebraic closure of $\mathbb{Q}$. It is not a finite extension of $\mathbb{Q}$, so it is not a number field. However, this observation allows us to embed any algebraic number field in the complex numbers. For any prime number p, the set

$$\mathbb{F}_{p^\infty} = \cup_d \mathbb{F}_{p^d}$$

is a field. It is the algebraic closure of every $\mathbb{F}_{p^d}$.

Theorem 3.4.3. *Let F be a number field. Then there are exactly $[F:\mathbb{Q}]$ embeddings of F in $\mathbb{C}$. If $K \subset F$ is a subfield then every embedding $\tau : K \to \mathbb{C}$ extends to $[F:K]$ distinct embeddings $\sigma : F \to \mathbb{C}$ such that $\sigma(b) = \tau(b)$ for all $b \in K$.*

Proof. Let $F = \mathbb{Q}(a)$. An embedding σ of F in $\mathbb{C}$ is completely determined by its value on a. The image $\sigma(a)$ is a root of the minimal polynomial $f \in \mathbb{Q}[x]$ of a over $\mathbb{Q}$ (thinking of f as a polynomial over $\mathbb{C}$). It is straightforward to check that every root of f determines an embedding. The number of roots of f is exactly its degree, since $\mathbb{C}$ is algebraically closed. Thus the number of embeddings of F in $\mathbb{C}$ is exactly the degree of f, which equals $[F:\mathbb{Q}]$. The proof of the second statement is similar, and is left as an exercise. □

Theorem 3.4.4. *Let F be a number field and let $\sigma_1, \cdots, \sigma_d$ be the distinct embeddings of F in $\mathbb{C}$. Let $b \in F$ and let $e = [\mathbb{Q}(b):\mathbb{Q}]$. Then*

1. $\mathsf{Tr}_{\mathbb{Q}}^{F}(b) = \sigma_1(b) + \sigma_2(b) + \cdots + \sigma_d(b) = e\mathsf{Tr}_{\mathbb{Q}}^{\mathbb{Q}(b)}(b)$.
2. $\mathsf{N}_{\mathbb{Q}}^{F}(b) = \sigma_1(b)\sigma_2(b)\cdots\sigma_d(b) = \left(\mathsf{N}_{\mathbb{Q}}^{\mathbb{Q}(b)}(b)\right)^e$.
3. *If F is a Galois extension of $\mathbb{Q}$ then*

$$\mathsf{Tr}_{\mathbb{Q}}^{F}(b) = \sum_{\sigma \in \mathrm{Gal}(F/\mathbb{Q})} \sigma(b) \quad \text{and} \quad \mathsf{N}_{\mathbb{Q}}^{F}(b) = \prod_{\sigma \in \mathrm{Gal}(F/\mathbb{Q})} \sigma(b).$$

4. *The trace map $\mathsf{Tr}_{\mathbb{Q}}^{F} : F \to \mathbb{Q}$ is surjective.*

Proof. (See also Theorem 3.2.14.) Let $f(x) = a_0 + a_1 x + \cdots + a_e x^e$ be the minimum polynomial of b. By definition, the roots of f (in $\mathbb{C}$) are distinct (although they are not necessarily all contained in $\mathbb{Q}(b)$ or even in F). The set $\{1, b, b^2, \cdots, b^{e-1}\}$ forms a basis for $\mathbb{Q}(b)$ as a vector space over $\mathbb{Q}$. With respect to this basis, the matrix M_b for the mapping $\ell_b : \mathbb{Q}(b) \to \mathbb{Q}(b)$ ($\ell_b(a) = ba$) is the

companion matrix of $f(x)$,that is, it has ones on the superdiagonal, $-a_0, -a_1, \cdots, -a_{e-1}$ in the last rown, and zeroes elsewhere. See equation (3.3). The characteristic polynomial of M_b is exactly the polynomial $f(x)$ and the eigenvalues of M_b are the distinct roots of $f(x)$. So the trace and norm of b are the sum and product of the roots of $f(x)$ which are

$$\mathsf{Tr}_{\mathbb{Q}}^{\mathbb{Q}(b)}(b) = -a_{e-1} \quad \text{and} \quad \mathsf{N}_{\mathbb{Q}}^{\mathbb{Q}(b)}(b) = (-1)^e a_e.$$

Let $\tau_1, \cdots, \tau_e$ denote the distinct embeddings of $\mathbb{Q}(b)$ into $\mathbb{C}$. The elements $\tau_1(b), \tau_2(b), \cdots, \tau_e(b)$ are exactly the roots of the polynomial f. Consequently

$$\mathsf{Tr}_{\mathbb{Q}}^{\mathbb{Q}(b)}(b) = \tau_1(b) + \cdots + \tau_e(b) \quad \text{and } \mathsf{N}_{\mathbb{Q}}^{\mathbb{Q}(b)}(b) = \tau_1(b)\tau_2(b)\cdots\tau_e(b).$$

Let $u_1, \cdots, u_t$ be a basis for F over $\mathbb{Q}(b)$ (hence $te = d$). Then the set $\{u_i b^j\}$ $(1 \le i \le t,$ $0 \le j \le e-1)$ forms a basis for F over $\mathbb{Q}$. With respect to this basis the mapping $L_b : F \to F$ $(L_b(a) = ba)$ is a "block matrix" with t diagonal blocks, each of which is a copy of the matrix M_b. It follows that the characteristic polynomial of L_b is $(f(x))^t$, that $\mathsf{Tr}(L_b) = t\mathsf{Tr}(M_b)$ and that $\det(L_b) = (\det(M_b))^t$. By Theorem 3.4.3, each embedding $\tau_i : \mathbb{Q}(b) \to \mathbb{C}$ extends to t distinct embeddings $F \to \mathbb{C}$ but these embeddings all take $b \in F$ to the same element, $\tau_i(b)$. Therefore

$$\sum_{i=1}^{d} \sigma_i(b) = t \sum_{i=1}^{e} \tau_i(b) = t\mathsf{Tr}(M_b) = \mathsf{Tr}(L_b) = \mathsf{Tr}_{\mathbb{Q}}^{F}(b)$$

and

$$\prod_{i=1}^{d} \sigma_i(b) = \prod_{i=1}^{e} \tau_i(b)^t = (\det(M_b))^t = \det(L_b) = \mathsf{N}_{\mathbb{Q}}^{F}(b).$$

If F is a Galois extension of $\mathbb{Q}$ then the embeddings $\sigma_i : F \to \mathbb{C}$ have the same image, and the Galois group $\mathrm{Gal}(F/Q)$ permutes these embeddings. Consequently,

$$\mathsf{Tr}_{\mathbb{Q}}^{F}(b) = \sum_{\sigma \in \mathrm{Gal}(F/\mathbb{Q})} \sigma(b) \quad \text{and} \quad \mathsf{N}_{\mathbb{Q}}^{F}(b) = \prod_{\sigma \in \mathrm{Gal}(F/\mathbb{Q})} \sigma(b).$$

Finally, the trace $\mathsf{Tr}_{\mathbb{Q}}^{F}$ is surjective if and only if it is nonzero, but $\mathsf{Tr}_{\mathbb{Q}}^{F}(1) = d \neq 0$. □

3.4.b Algebraic integers

Just as algebraic number fields are generalizations of the rational numbers, there is a generalization of the rational integers $\mathbb{Z}$.

Definition 3.4.5. *An algebraic number a is an* algebraic integer *or is* integral *if its minimal polynomial $f \in \mathbb{Q}[x]$ over $\mathbb{Q}$ has all its coefficients in $\mathbb{Z}$.*

Theorem 3.4.6. *The following are equivalent*

1. *a is an algebraic integer.*
2. *$\mathbb{Z}[a]$ is a finitely generated $\mathbb{Z}$-module.*
3. *$a \in R$ for some ring $R \subseteq \mathbb{C}$ that is a finitely generated $\mathbb{Z}$-module.*
4. *$aM \subseteq M$ for some finitely generated $\mathbb{Z}$-module $M \subseteq \mathbb{C}$.*

Proof. If a is an algebraic integer, then a^d is a linear combination of $1, a, \cdots, a^{d-1}$ with integer coefficients, and it follows that $\mathbb{Z}[a]$ is generated as a $\mathbb{Z}$-module by $1, a, \cdots, a^{d-1}$. The implications (2) $\implies$ (3) $\implies$ (4) are straightforward.

To prove that (4) implies (1), suppose that M is generated by $m_1, \cdots, m_k$. Thus for $j = 1, \cdots, k$, we have

$$am_j = \sum_{i=1}^{k} b_{i,j} m_j \tag{3.23}$$

with $b_{i,j} \in \mathbb{Z}$. Let $c_{i,j} = b_{i,j}$ if $i \neq j$, and $c_{i,i} = b_{i,i} - x$. It follows from equation (3.23) that the determinant of the matrix $[c_{ij}]$ is zero at $x = a$. But the determinant of this matrix is a monic polynomial with integer coefficients, so a is algebraic. □

3.4.c Orders

Let F be an algebraic number field and let $m = [F : \mathbb{Q}]$. If $R \subset F$ is a subring, then it is automatically an integral domain. An *order* $R \subset F$ is a subring of F that is finitely generated as a $\mathbb{Z}$-module with rank m. In this case, Corollary 2.1.18 implies that R^+ is isomorphic to $\mathbb{Z}^m$. A standard result is the following.

Theorem 3.4.7. *A subring R in a number field F is an order in F if and only if it satisfies the following three conditions,*

1. *$R \cap \mathbb{Q} = \mathbb{Z}$*
2. *The fraction field (Section 2.2.h) of R is F.*
3. *The Abelian group $(R, +)$ is finitely generated.*

Except when $F = \mathbb{Q}$, there are infinitely many orders in F. Every order $R \subset F$ consists entirely of algebraic integers and in fact the intersection $\mathbb{Z}_F = F \cap \mathbb{F}A$ (where $\mathbb{F}A$ denotes the set of all algebraic integers) is an order which contains all the other orders in F. This maximal order $\mathbb{Z}_F$ is called the *ring of integers* of F. It is *integrally closed* in F, meaning that if $\alpha \in F$ is a root of a *monic* polynomial with coefficients in $\mathbb{Z}_F$ then $\alpha \in \mathbb{Z}_F$ also. In fact, $\mathbb{Z}_F$ is the integral closure of $\mathbb{Z}$ in F, that is, it consists of all elements $\alpha \in F$ which are roots of monic polynomials with coefficients in $\mathbb{Z}$. So a subset $R \subset F$ is an order if and only if it is a subring and it is contained in $\mathbb{Z}_F$ as a subgroup of finite index.

The ring of integers of $\mathbb{Q}$ is $\mathbb{Z}$; the ring of integers of $\mathbb{Q}[i]$ is $\mathbb{Z}[i]$. However the ring of integers of $\mathbb{Q}[\sqrt{5}]$ is larger than $\mathbb{Z}[\sqrt{5}]$ (which is an order). Rather, the ring of integers consists of all integer linear combinations of $(1+\sqrt{5})/2$ and $(1-\sqrt{5})/2$. For any number field F the maximal order $\mathbb{Z}_F$ has several particularly nice properties (it is a Dedekind ring, for example). However in Section 8.3.a we consider algebraic shift registers whose entries come from an arbitrary order in an arbitrary number field.

Lemma 3.4.8. *Let R be an order in a number field F. Let $a \in R$. Then the number of elements in the quotient ring is $|R/(a)| = |\mathsf{N}(a)|$.*

Proof. (See also Exercise 5.) Let $\{u_1, \cdots, u_n\}$ be an integer basis for the $\mathbb{Z}$-module R. Then the set $\{au_1, \cdots, au_n\}$ is an integer basis for the $\mathbb{Z}$-module (a). Each au_i is some integer linear combination, say, $au_i = A_{i1}u_1 + \cdots + A_{in}u_n$. On the one hand, the matrix $A = (A_{ij})$ describes the action of multiplication by a on M so $\mathsf{N}(a) = \det(A)$. On the other hand, according to Theorem 2.2.28, $|R/(a)| = |\det(A)|$. □

In particular, the theorem says that the norm of any algebraic integer is an ordinary integer. The absolute value of the norm gives a candidate function for defining division with remainder, see Definition 2.2.13. If $F = \mathbb{Q}(\sqrt{d})$ is a quadratic number field then its ring of integers is a Euclidean domain with respect to this function if and only if d =2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57, 73, -1, -2, -3, -7, -11, see [17] Section 3.2. A number field has *class number* one if and only if its ring of integers is a unique factorization domain. The *Stark-Heegner Theorem* states that the only quadratic imaginary number fields with class number one are $\mathbb{Q}(\sqrt{d})$ for $d = -1$, -2, -3, -7, -11, -19, -43, -67, -163. However the ring of integers of any number field is a *Dedekind domain*, and as such it admits unique prime decomposition of ideals, i.e., for any ideal I there are uniquely determined prime ideals $P_1, P_2, \cdots, P_k$ and integers $m_1, m_2, \cdots, m_k$ so that $I = P_1^{m_1} P_2^{m_2} \cdots P_k^{m_k}$. Moreover, we next show that every element in an order can be written in at least one way as a product of irreducible elements, so it is a *factorization ring*.

Theorem 3.4.9. *Let R be an order in a number field F and let $a \in R$. Then there is a unit u and irreducible elements $f_1, \cdots, f_k \in R$ so that $a = uf_1 f_2 \cdots f_k$. (If a is not a unit then the element u can be absorbed into one of the irreducible factors.)*

Proof. Use induction on $|\mathsf{N}(a)|$. If $|\mathsf{N}(a)| = 1$, then a is a unit so we are done. Likewise, if a is irreducible we are done. Otherwise we can write $a = bc$ where neither b nor c is a unit. Then $|\mathsf{N}(a)| = |\mathsf{N}(b)||\mathsf{N}(c)|$ and neither $|\mathsf{N}(b)|$ nor $|\mathsf{N}(c)|$ is equal to 1. Thus $|\mathsf{N}(b)| < |\mathsf{N}(a)|$ and $|\mathsf{N}(c)| < |\mathsf{N}(a)|$. By induction both b and c can be written as units times a product of irreducible elements, and we can combine these expressions into such an expression for a. □

3.5 Local and global fields

3.5.a Local fields

There are two types of *local fields*: local function fields, and p-adic fields. They are discussed in more detail in Section 5.6.c (where local fields are defined) but here are the basic definitions. If F is a field, then the (local) *function field* $F((x))$ consists of all *formal Laurent series* $\sum_{i=-k}^{\infty} a_i x^i$, with $a_i \in F$. Such a series has finitely many terms of negative degree and possibly infinitely many terms of positive degree. Its ring of "integers" is the subring $F[[x]]$ of formal power series, that is, sums with no terms of negative degree. The ring $F[[x]]$ is a local ring with unique maximal ideal (x). The field $F((x))$ is the fraction field of $F[[x]]$. That is, every formal Laurent series $a(x) \in F((x))$ may be expressed as a quotient $a(x) = f(x)/g(x)$ of two formal power series $f, g \in F[[x]]$ (and in fact the denominator $g(x)$ may be chosen to be a power of x). Addition and multiplication in $F((x))$ are performed in a way that is analogous to the addition and multiplication of polynomials. One must check that only finitely many terms contribute to any term in a product.

Let p be a prime number. The p-adic field $\mathbb{Q}_p$ consists of all formal Laurent series $\sum_{i=-k}^{\infty} a_i p^i$ (with finitely many terms of negative degree and possibly infinitely many terms of positive degree), where $0 \leq a_i \leq p-1$, and where addition and multiplication are performed "with carry". It contains a ring $\mathbb{Z}_p$ of "integers" consisting of formal power series with no terms of negative degree, which is a local ring with maximal ideal (p). The field $\mathbb{Q}_p$ is the fraction field of $\mathbb{Z}_p$: every $a \in \mathbb{Q}_p$ can be expressed as a fraction f/g with $f, g \in \mathbb{Z}_p$ and in fact the denominator g may be chosen to be a power of p. A *p-adic field* is a finite degree extension of $\mathbb{Q}_p$.

3.5.b Global fields

There are also two types of *global fields*: function fields and algebraic number fields. The algebraic number fields (= finite degree extensions of $\mathbb{Q}$) have been previously discussed in Section 3.4. Let F be a field. A *global function field* over F is any finite degree extension of the field $F(x)$ of rational functions. The field $F(x)$ is the fraction field of the ring $F[x]$ of polynomials, that is, every element of $F(x)$ is of the form f/g where f and g are polynomials. If K is a finite degree extension of $F(x)$ then there exists n so that $K \cong F[x_1, x_2, \cdots, x_n]/I$ where I is an appropriate maximal ideal in the ring $F[x_1, x_2, \cdots, x_n]$ of polynomials in n variables. One normally assumes that K has transcendance degree one over F, in other words, the set

$$V(I) = \{(x_1, x_2, \cdots, x_n) \in F^n : \ h(x) = 0 \text{ for all } h \in I\}$$

is a one dimensional algebraic variety, or an *algebraic curve*. Then K is called the field of rational functions on the algebraic curve $V(I)$.

3.6 Exercises

1. Lemma 3.2.14: Let d and e be positive integers with d dividing e. Prove that if $a \in \mathbb{F}_{p^e}$, then $a + a^{p^d} + a^{p^{2d}} + \cdots + a^{p^{e-d}} \in \mathbb{F}_{p^a}$.

2. Suppose p is prime and c, d, and e are integers with $c|d|e$. Prove that $\mathsf{Tr}_{p^c}^{p^d} \circ \mathsf{Tr}_{p^d}^{p^e} = \mathsf{Tr}_{p^c}^{p^e}$.

3. Develop an alternate definition of the trace function for a finite field F in terms of embeddings of F in its algebraic closure. Prove that your definition agrees with the previous one.

4. Let $F = \mathbb{Q}[\sqrt{5}]$. Show that the full ring of integers of F is $K = \mathbb{Z}[(1 \pm \sqrt{5})/2]$ and that the norm of an element $a + b\sqrt{5}$ is $a^2 + 5b^2$.

5. (*continued*) Let $L = \mathbb{Z}[\sqrt{5}]$. It is an order in K. Let $a = 2 \in L$. The inclusion $L \subset K$ induces a homomorphism of quotient rings $L/aL \to K/aK$. According to Lemma 3.4.8, both rings have 4 elements. Show that $L/aL \cong \mathbb{F}_2[x]/(x^2)$, that $K/aK \cong \mathbb{F}_4$, and that the mapping $L/aL \to K/KaK$ is neither injective nor surjective. (*Hint:* The elements of L/aL are represented by 0, 1, $\sqrt{5}$, $1 + \sqrt{5}$, from which multiplication and addition tables can be constructed. The ring K/aK can be similarly analyzed.)

Chapter 4 Finite Local Rings and Galois Rings

Local rings have a single maximal ideal. In algebraic geometry they are used to understand the local geometry at a single point on an algebraic variety. These rings, and especially the special case of Galois rings (see Section 4.5), generalize finite fields. They have recently been used in several constructions of error correcting codes and families of sequences with interesting correlation properties. They are useful models of multiphase signals.

4.1 Finite local rings

In this section we examine the structure of a commutative ring (with identity) which has finitely many elements. The standard reference for this section is [141]. During the last decade a considerable amount of effort has been directed towards developing linear feedback shift register sequences based on a finite local ring R. The analysis of these sequences depends on an understanding of the units in R (see Theorem 6.6.4).

Let R be a commutative ring. Recall from Definition 2.2.13 that R is a *local ring* if it contains a unique maximal ideal $\mathfrak{m}$. In this case (see Section 2.2.a), the maximal ideal $\mathfrak{m}$ consists precisely of the non-units of R. The quotient $F = R/\mathfrak{m}$ is a field and is called the *residue field* of R. For each $i \geq 0$ the quotient $\mathfrak{m}^{i-1}/\mathfrak{m}^i$ is naturally a vector space over F, because R acts on this quotient by multiplication, and $\mathfrak{m}$ acts trivially. The following are examples of finite local rings.

- any finite (Galois) field.
- $\mathbb{Z}/(p^n)$ for any prime number p, with maximal ideal (p) and residue field $\mathbb{Z}/(p)$.
- $\mathbb{F}[x]/(f^n)$, where $\mathbb{F}$ is a finite field and f is an irreducible polynomial, with maximal ideal (f) and residue field $\mathbb{F}[x]/(f)$.
- $R[x]/(f^n)$ where R is a finite local ring and f is a basic irreducible polynomial (see below).

Any commutative finite ring may be expressed as a direct sum of finite local rings. For the remainder of this section we assume that R is a finite local ring.

Basic irreducible polynomials: Let R be a finite local ring with maximal ideal $\mathfrak{m}$. Let $\mu : R \to F = R/\mathfrak{m}$ be the projection. Applying μ to each coefficient of a polynomial gives a mapping which we also denote by $\mu : R[x] \to F[x]$. A polynomial $f(x) \in R[x]$ is *regular* if it is not a zero divisor, which holds if and only if $\mu(f) \neq 0$. Let $f(x) \in R[x]$. If $\mu(f)$ is nonzero and is irreducible in $F[x]$ then f is irreducible in $R[x]$, and we refer to f as a *basic irreducible polynomial*. In this case $R[x]/(f^n)$ is again a local ring for any $n > 0$ (see ([141], XIV.10). Its maximal ideal is $\mathfrak{m}[x] + (f)$

and its residue field is $F[x]/(\mu(f))$, where $\mathfrak{m}[x]$ is the collection of those polynomials $f \in R[x]$ all of whose coefficients are in $\mathfrak{m}$.

If the leading term of a basic irreducible polynomial $f(x) \in R[x]$ is in the maximal ideal $\mathfrak{m}$ then the degree of the reduction $\mu(f) \in F[x]$ will be less than $\deg(f)$. If $f(x)$ is a *monic* polynomial then $\deg(f) = \deg(\mu(f))$ since the leading term is 1. For this reason we will often consider monic basic irreducible polynomials.

Lemma 4.1.1. *Let $f \in R[x]$ be a regular polynomial and suppose $\bar{\alpha} \in F$ is a simple zero of $\mu(f) \in F[x]$. Then f has one and only one root $\alpha \in R$ such that $\mu(\alpha) = \bar{\alpha}$.*

Proof. This is proven in Lemma (XV.1) of [141]. □

Further properties of polynomials over R are described in Section 4.3. The following is a powerful tool for studying local rings.

Theorem 4.1.2. *(Nakayama's Lemma for local rings [141], [139, p. 11]) Let R be a finite local ring with maximal ideal $\mathfrak{m}$. Let M be a module over R.*

1. *If M is finite and $\mathfrak{m}M = M$, then $M = 0$.*
2. *If N is a submodule of M and $M = N + \mathfrak{m}M$, then $N = M$.*

4.1.a Units in a finite local ring

Let R be a finite local ring with maximal ideal $\mathfrak{m}$ and residue field F. Let $R^\times$ be the set of invertible elements in R. Let $1 + \mathfrak{m} = \{1 + a : \ a \in \mathfrak{m}\}$. By [141], Theorem (V.1) and Proposition (IV.7),

- the ideal $\mathfrak{m}$ consists precisely of the non-units of R,
- for every $a \in R$, at least one of a and $1 + a$ is a unit, and
- there is a positive integer n such that $\mathfrak{m}^n = 0$.

The details are left as an exercise.

An element a is *nilpotent* if for some natural number k we have $a^k = 0$. It follows from the above that every element a of R is either a unit or is nilpotent. In fact we can take the same k for all a.

Proposition 4.1.3. *There exists an isomorphism of Abelian groups*

$$R^\times \cong F^\times \times (1 + \mathfrak{m}) \tag{4.1}$$

Proof. Let n be the smallest integer such that $\mathfrak{m}^n = 0$. It is called the *degree of nilpotency* of $\mathfrak{m}$. As in [141] Exercise (V.9), we have a sequence of surjective ring homomorphisms

$$R = R/\mathfrak{m}^n \xrightarrow{\sigma_n} R/\mathfrak{m}^{n-1} \xrightarrow{\sigma_{n-1}} \cdots \xrightarrow{\sigma_2} R/\mathfrak{m} = F.$$

For $2 \leq i \leq n$, the kernel $\mathrm{Ker}(\sigma_i) = \mathfrak{m}^{i-1}/\mathfrak{m}^i$ is a vector space over F. If $|F| = q$ it follows by induction that there exists an integer j such that

$$|\mathfrak{m}| = q^j \quad \text{and} \quad |R| = q^{j+1}. \tag{4.2}$$

The natural ring homomorphism $\mu : R \to F = R/\mathfrak{m}$ gives an exact sequence of (multiplicative) Abelian groups,

$$1 \to 1 + \mathfrak{m} \to R^\times \to F^\times \to 1.$$

The Abelian group $F^\times$ is cyclic of order $q-1$, and $1+\mathfrak{m}$ has order q^j, which is relatively prime to $q-1$. It follows (from the structure theorem for finite Abelian groups, Theorem 2.1.16) that there is a splitting $\iota : F^\times \to R^\times$ and this gives the isomorphism (4.1). □

The structure of $1+\mathfrak{m}$ is often very complicated. However it is possible to identify the cyclic group $F^\times$ as a subgroup of $R^\times$.

Lemma 4.1.4. *There is a unique (group homomorphism) splitting $\iota : F^\times \to R^\times$ of the projection μ, and its image consists of all elements $\alpha \in R$ such that $\alpha^{q-1} = 1$.*

Proof. Every element $a \in F^\times$ satisfies $a^{q-1} = 1$ so if ι exists, the same must be true of $\iota(a)$. Let $g(x) = x^{q-1} - 1$. Then every element of $F^\times$ is a (simple) root of $\mu(g) \in F[x]$. Therefore g is a regular polynomial, and Lemma 4.1.1 implies that every element $a \in F^\times$ has a unique lift $\iota(a) \in R$ such that $\iota(a)^{q-1} = 1$. Hence the splitting ι exists, and there is only one such. □

4.2 Examples

4.2.a $\mathbb{Z}/(p^m)$

Fix a prime number $p \in \mathbb{Z}$ and let $R = \mathbb{Z}/(p^m)$. This is a finite local ring with maximal ideal $\mathfrak{m} = (p)$ and residue field $F = \mathbb{Z}/(p)$. The multiplicative group $F^\times$ is cyclic, of order $p-1$. By Proposition 4.1.3 the group of units $R^\times$ is the product $F^\times \times (1+\mathfrak{m})$.

Proposition 4.2.1. *If $p > 2$ then $1+\mathfrak{m}$ is a cyclic group of order p^{m-1} so $R^\times \cong \mathbb{Z}/(p-1) \times \mathbb{Z}/(p^{m-1}) \cong \mathbb{Z}/(p^{m-1}(p-1))$. If $p = 2$ and if $m \geq 3$ then $1+\mathfrak{m}$ is a product of two cyclic groups, one of order 2 (generated by the element -1), the other of order 2^{m-2} (generated by the element 5).*

Proof. The order of the group of units is easy to calculate: since every pth integer is a multiple of p, there are $p^m/p = p^{m-1}$ non-invertible elements in R. So there are $p^m - p^{m-1} = (p-1)p^{m-1}$ units. It follows that $1 + \mathfrak{m}$ contains p^{m-1} elements.

Now consider the case $p \geq 3$. Define $E : \mathbb{Z} \to R = \mathbb{Z}/(p^m)$ by $E(a) = \exp(pa) \pmod{p^m}$. That is,

$$E(a) = 1 + pa + \frac{p^2a^2}{2!} + \frac{p^3a^3}{3!} + \cdots \pmod{p^m} \tag{4.3}$$

Consider the nth term, $a^n p^n/n!$. The number $n!$ is not necessarily invertible in $\mathbb{Z}/(p^m)$ but the number $p^n/n!$ does make sense in $\mathbb{Z}/(p^n)$ if we interpret it to mean that the factor p^e which occurs in the prime decomposition of $n!$ should be canceled with the same factor p^e which occurs in the numerator. In fact, the prime p occurs in the prime decomposition of $n!$ fewer than $n/p + n/p^2 + n/p^3 \cdots = n/(p-1)$ times. Since it occurs in the numerator n times, it is possible to cancel all occurrences of p from the denominator. This leaves a denominator which is relatively prime to p and hence is invertible in $\mathbb{Z}/(p^m)$. It follows, moreover, that after this cancellation the numerator still has at least $n(p-2)/(p-1)$ factors of p. So if $n \geq m(p-1)/(p-2)$ the term $a^n p^n/n!$ is 0 in $\mathbb{Z}/(p^m)$. Therefore the sum (4.3) is finite.

Since $E(a+b) = E(a)E(b)$, the mapping E is a group homomorphism. Moreover $E(a) = 1$ if and only if a is a multiple of p^{m-1}. So E induces to an injective homomorphism

$$E : \mathbb{Z}/(p^{m-1}) \to 1 + \mathfrak{m}.$$

This mapping is also surjective because both sides have p^{m-1} elements.

Now consider the case $R = \mathbb{Z}/(2^m)$ with $m \geq 3$. The element $\{-1\}$ generates a cyclic subgroup of order 2. The element 5 generates a cyclic subgroup of order 2^{m-2}. To show this, first verify by induction that

$$5^{2^{m-3}} = (1+2^2)^{2^{m-3}} \equiv 1 + 2^{m-1} \pmod{2^m}$$

so this number is not equal to 1 in $\mathbb{Z}/(2^m)$. However

$$5^{2^{m-2}} \equiv (1+2^{m-1})^2 \equiv 1 \pmod{2^m}.$$

So 5 has order 2^{m-2} in R. Since -1 is not a power of 5 (mod 4) it is also not a power of 5 (mod 2^m). Therefore the product of cyclic groups $\langle -1 \rangle \langle 5 \rangle$ has order 2^{m-1}, and it consequently exhausts all the units. □

4.2.b $F[x]/(x^m)$

Let F be a finite field and let $R = F[x]/(x^m)$. Then R is a finite local ring with maximal ideal $\mathfrak{m} = (x)$ and with residue field F. The mapping $\mu : R \to F$ (which associates to each polynomial its constant term) takes $R^\times$ surjectively to $F^\times$. This mapping has a splitting $F^\times \to R^\times$ which assigns

to any nonzero $a \in F$ the polynomial $a+0x$. This gives an isomorphism $R^\times \cong F^\times \times (1+\mathfrak{m})$, where $1+\mathfrak{m}$ is the (multiplicative) group of all polynomials of the form $1+xh(x)$, $h(x)$ a polynomial of degree $\leq m-2$. (So we have recovered Proposition 4.1.3.) It is fairly difficult to determine the exact structure of the group $1+\mathfrak{m}$ but the following proposition describes its general form. Let $q = |F| = p^d$ with p prime.

Proposition 4.2.2. *The (multiplicative) group $1+\mathfrak{m}$ is isomorphic to a product of (additive) cyclic groups,*

$$1+\mathfrak{m} \cong \mathbb{Z}/(p^{r_1}) \times \cdots \times \mathbb{Z}/(p^{r_k}) \tag{4.4}$$

where each $r_i \leq \lceil \log_p(m) \rceil$ and where

$$\sum_{i=1}^{k} r_i = d(m-1).$$

Proof. The Abelian group $1+\mathfrak{m}$ is finite. It is thus isomorphic to a product of cyclic groups whose orders divide $|1+\mathfrak{m}| = q^{m-1}$ which is a power of p. This gives an abstract isomorphism (4.4). Counting the number of elements on each side of this equation gives

$$q^{m-1} = p^{r_1+\cdots+r_k}$$

so

$$d(m-1) = \sum_{i=1}^{k} r_i.$$

Let r be the smallest integer such that $p^r \geq m$. Then $y^{p^r} = 1$ for any $y \in 1+\mathfrak{m}$, because expressing $y = 1+xh(x)$ and computing in $F[x]$ we find that

$$y^{p^r} = 1 + x^{p^r} h^{p^r} \equiv 1 \pmod{x^m}.$$

Thus the cyclic groups occurring in (4.4) each have $r_i \leq r = \lceil \log_p(m) \rceil$. □

4.2.c $F[x]/(f^m)$

Let F be a finite field and let $f \in F[x]$ be an irreducible polynomial. Fix $m \geq 1$. The ring $R = F[x]/(f^m)$ is a finite local ring with maximal ideal (f) and with quotient field $K = F[x]/(f)$. The next result identifies the ring R with that of Section 4.2.b. Let

$$\mu : R = F[x]/(f^m) \to K = F[x]/(f)$$

be reduction modulo f. It is a surjective ring homomorphism.

Proposition 4.2.3. *There is a unique splitting of μ. That is, there is a unique injective ring homomorphism $\varphi : K \to R$ so that $\mu(\varphi(a)) = a$ for all $a \in K$. Moreover the mapping φ extends to a mapping $\varphi : K[y] \to R$ by setting $\varphi(y) = f$. The resulting mapping*

$$\bar{\varphi} : K[y]/(y^m) \to R$$

is an isomorphism of rings.

Proof. Let q denote the number of elements in F and let $Q = q^d$ denote the number of elements in K, where $d = \deg(f)$. First we show that the set

$$Z_m = \left\{ g \in F[x]/(f^m) : \ g^Q = g \right\}$$

is a *lift*[1] of the field K to R. It is therefore a candidate for the image of φ.

The set Z_m is closed under addition and multiplication, because if $g_1, g_2 \in Z_m$ then

$$(g_1 + g_2)^Q = g_1^Q + g_2^Q = g_1 + g_2.$$

Moreover the restriction $\mu : Z_m \to K$ is an injection, for if $g \in Z_m$ lies in the kernel of μ and if $\dot{g} \in F[x]$ is any lift of g, then f divides $\dot{g}$. However f^m divides $\dot{g}^Q - \dot{g} = (\dot{g}^{Q-1} - 1)(\dot{g})$. Since these two factors are relatively prime, it follows that f^m divides $\dot{g}$, which says that $g = 0$ in R. Now let us show that the restriction $\mu : Z_m \to K$ is surjective. Fix $a \in K$. We need to find $g \in Z_m$ so that $\mu(g) = a$. We use induction on m, and the case $m = 1$ holds trivially. So let m be arbitrary and consider the mapping

$$\mu_m : F[x]/(f^m) \to F[x]/(f^{m-1}).$$

By induction, there exists $g' \in F[x]/(f^{m-1})$ so that $(g')^Q = g'$ and so that g' maps to the given element $a \in K$, that is, $g' \pmod f = a$. Let $\dot{g}' \in F[x]$ be any lift of g' to $F[x]$. Then f^{m-1} divides $(\dot{g}')^Q - \dot{g}'$, or

$$(\dot{g}')^Q - \dot{g}' = f^{m-1}h$$

for some polynomial $h \in F[x]$. Set $g = \dot{g}' + hf^{m-1}$. Then

$$(g)^Q - g = (\dot{g}')^Q - \dot{g}' + h^Q f^{(m-1)Q} - hf^{m-1} = h^Q f^{(m-1)Q}$$

which is divisible by f^m. This says that the class $[g] \in F[x]/(f^m)$ lies in the set Z_m and that $g \pmod f = a$ as needed.

The splitting φ is unique because every element g in the image of a splitting must satisfy $g^Q = g$. This function $\varphi : K \to R.$ extends to a function $\varphi : K[y] \to R$ by mapping y to f. We

[1] If $\tau : A \to B$ is a set function, then a *lift* of a subset $C \subset B$ is a subset of $D \subset A$ that is mapped by τ one to one and onto C. A lift of an element $y \in B$ is an element $x \in A$ so that $\tau(x) = y$.

claim that the kernel of φ is (y^m) and that φ is onto. The kernel contains (y^m) since $f^m = 0$ in R. Let

$$g(y) = \sum_{i=0}^{m-1} g_i y^i$$

with $\varphi(g) = 0$. Thus

$$\sum_{i=0}^{m-1} g_i f^i = 0. \tag{4.5}$$

As a vector space over K the ring R has dimension m since $|R| = Q^m = |K|^m$. R is spanned over K by $\{1, f, f^2, \cdots, f^{m-1}\}$ (this can be proved by induction on m). Therefore these elements form a basis. As we have seen in the preceding paragraph, the projection $\mu_m : F[x]/(f^m) \to F[x]/(f^{m-1})$ takes Z_m to Z_{m-1} (both of which are lifts of the field K). Applying the projection μ_m to equation (4.5) gives

$$\sum_{j=0}^{m-2} g_j f^j = 0$$

and by induction we conclude that $g_0 = g_1 = \ldots = g_{m-2} = 0$. This leaves $g_{m-1} f^{m-1} = 0$ in the ring R, which means that f^m divides $g_{m-1} f^{m-1}$ in the polynomial ring $F[x]$. But $F[x]$ is an integral domain, so we conclude that f divides g_{m-1}, hence $g_{m-1} = 0$ as an element of K.

In conclusion, we obtain a well defined surjective ring homomorphism $K[y] \to R$ by sending y to f. The kernel of this homomorphism is the ideal (y^m) so we obtain an isomorphism $K[y]/(y^m) \to R$. □

4.2.d Equal characteristics

Suppose that R is a finite local ring with maximal ideal $\mathfrak{m}$ and quotient field $F = R/\mathfrak{m}$. Recall that there is a unique homomorphism $\mathbb{Z} \to R$ (taking m to $1 + \cdots + 1$, m times). Its kernel is an ideal $(t) \subset \mathbb{Z}$ where $t = \mathrm{char}(R)$ is the characteristic of R. Since R is finite, its characteristic is nonzero. If the characteristic of R were divisible by two distinct primes, say p and q, then neither p nor q would be a unit, hence both would be in $\mathfrak{m}$. It would follow that 1 is in $\mathfrak{m}$, since 1 is an integer linear combination of p and q. Hence $\mathrm{char}(R) = p^e$ is a power of a prime p. Moreover, the image of $\mathbf{Z}$ in F is a quotient of its image $\mathbf{Z}/(p^e)$ in R. Thus $\mathrm{char}(F) = p$.

Let $\mu : R \to F$ be the quotient mapping. Set $q = |F|$. Let $\iota : F^\times \to R^\times$ be the homomorphism described in Lemma 4.1.4. Extend ι to F by mapping 0 to 0.

Proposition 4.2.4. *The function $\iota : F \to R$ is a ring homomorphism if and only if R and F have the same characteristic (so $t = p$ and $e = 1$).*

Proof. Whenever $f : A \to B$ is a ring homomorphism, $\mathrm{char}(B)$ divides $\mathrm{char}(A)$ because the kernel of the homomorphism from $\mathbf{Z}$ to B contains the kernel of the homomorphism from $\mathbf{Z}$ to A. If ι is a ring homomorphism then $\mathrm{char}(R)|\mathrm{char}(F)$ so $e = 1$.

For the convese, suppose that R and F have the same characteristic, p. Since ι is multiplicative, to show that ι is a ring homomorphism we need only show that for every $a, b \in F$ we have $\iota(a) + \iota(b) = \iota(a + b)$. The polynomial $x^q - x$ is regular over F and has only simple roots. Thus $\iota(a)$ can be defined to be the unique element of R so that $\mu(\iota(a)) = a$ and that is a root of $x^q - x$. This element exists by Lemma 4.1.1. We have

$$\mu(\iota(a) + \iota(b)) = \mu(\iota(a)) + \mu(\iota(b)) = a + b = \mu(\iota(a + b)),$$

so that $\iota(a) + \iota(b)$ and $\iota(a + b)$ are congruent modulo $\mathfrak{m}$. Also,

$$(\iota(a) + \iota(b))^q = \iota(a)^q + \iota(b)^q = \iota(a) + \iota(b),$$

because $(x + y)^p = x^p + y^p$ in any ring of prime characteristic p. Thus $\iota(a) + \iota(b)$ is in the image of ι and must equal $\iota(a + b)$. □

Proposition 4.2.5. *Let R be a finite local ring with quotient field $F = R/\mathfrak{m}$. Then* $\mathrm{char}(R) = \mathrm{char}(F)$ *if and only if there exists r, k so that R is isomorphic to the quotient $F[x_1, x_2, \cdots, x_r]/I$ where I is an ideal that contains every monomial of degree $\geq k$.*

Proof. Suppose $\mathrm{char}(R) = \mathrm{char}(F)$. Use ι to identify F as a subring of R. Let M be a set of variables in one to one correspondence with the elements of $\mathfrak{m}$. Then $R \cong F[M]/I$, where I is the set of all polynomials in M that vanish when the elements of M are replaced by the corresponding elements of $\mathfrak{m}$. Since $\mathfrak{m}^k = (0)$ for some k, every monomial of degree $\geq k$ is contained in I. Conversely, if F is a finite field, M is a finite set of variables, and I is an ideal in $F[M]$ containing every monomial of degree $\geq k$, then $R = F[M]/I$ is a finite local ring with maximal ideal generated by M whose characteristic equals that of its quotient field. □

4.3 Divisibility in $R[x]$

Throughout this subsection, R denotes a finite local ring with $\mu : R \to F = R/\mathfrak{m}$ the projection to its residue field. Let $f, g \in R[x]$.

1. f is *nilpotent* if $f^n = 0$ for some $n \geq 0$.
2. f is a *unit* if there exists $h \in R[x]$ so that $fh = 1$.
3. f is *regular* if f is not a zero divisor.
4. f is *prime* if the ideal (f) is a proper prime ideal.
5. f is *irreducible element* if f is not a unit and, whenever $f = gh$ then g or h is a unit.

6. f and g are *coprime* if $R[x] = (f) + (g)$.

In [141] the following results are proven.

Theorem 4.3.1. *Let $f = a_0 + a_1 x + \cdots + a_d X^d \in R[x]$. Then*

1. *The following are equivalent:*
   (a) *f is a unit.*
   (b) *$\mu(f) \in F[x]$ is a unit.*
   (c) *a_0 is a unit and the remaining coefficients $a_1, \cdots, a_d$ are nilpotent.*
2. *The following are equivalent:*
   (a) *f is nilpotent.*
   (b) *$\mu(f) = 0$.*
   (c) *All the a_i are nilpotent.*
   (d) *f is a zero divisor.*
   (e) *there exists $a \neq 0$ in R such that $af = 0$.*
3. *The following are equivalent:*
   (a) *f is regular.*
   (b) *$\mu(f) \neq 0$.*
   (c) *a_i is a unit for some i $(0 \leq i \leq d)$.*
4. *f and g are coprime if and only if $\mu(f)$ and $\mu(g)$ are coprime. In this case, f^i and g^j are coprime for all $i, j \geq 1$.*
5. *If $\mu(f)$ is irreducible then f is irreducible. If f is irreducible then $\mu(f) = ag^n$ where $a \in F$ and $g \in F[x]$ is a monic irreducible polynomial.*
6. *(Euclidean algorithm) If $f \neq 0$ and if $g \in R[x]$ is regular then there exist (not necessarily unique) elements $q, r \in R[x]$ such that $\deg r < \deg g$ and $f = gq + r$.*
7. *If f and g are monic and regular and if $(f) = (g)$ then $f = g$.*

Recall that an ideal $I \subset R[x]$ is *primary* if $I \neq R[x]$ and whenever $ab \in I$, then either $a \in I$ or $b^n \in I$ for some $n \geq 1$. An element $g \in R[x]$ is primary if (g) is primary.

Proposition 4.3.2. *An element $f \in R[x]$ is a primary regular non-unit if and only if $f = ug^n + h$ where $u \in R[x]$ is a unit, $g \in R[x]$ is a basic irreducible, $n \geq 1$, and $h \in \mathfrak{m}[x]$ (that is, all the coefficients of h lie in $\mathfrak{m}$).*

Although $R[x]$ is not necessarily a unique factorization domain, the following theorem ([141] Thm. XIII.11) states that regular polynomials have unique factorization.

Theorem 4.3.3. *Let $f \in R[x]$ be a regular polynomial. Then there exist unique (up to reordering and multiplication by units) regular coprime primary polynomials $g_1, g_2, \cdots, g_n \in R[x]$ so that $f = g_1 g_2 \cdots g_n$.*

4.4 Tools for local rings

In this section we develop several tools for the analysis of finite local rings – Galois theory, the trace and norm, and primitive elements. These are all generalizations of the similarly named tools for analyzing finite fields, and in most cases we use the finite field versions to help construct the finite local ring version.

4.4.a Galois theory of local rings

In the next few paragraphs we see that a finite local ring R has a distinguished collection of *Galois extensions* $\mathrm{GR}(R, n)$, one for each positive integer n, which are themselves local rings and for which many of the familiar properties of Galois fields continue to hold.

Extensions. Let R be a finite local ring. An *extension* ring is a finite local ring S which contains R. Any extension S of R is an R-algebra. A ring homomorphism $\varphi : S \to S$ is said to be an R-algebra automorphism of S provided it is both surjective and injective, and provided $\varphi(ac) = a\varphi(c)$ for all $a \in R$ and $c \in S$. Define the *Galois group*

$$G = \mathrm{Gal}(S/R) = \mathrm{Aut}_R(S)$$

to be the set of R-algebra automorphisms of S. The Galois group G acts on S. Let S^G denote the set of elements which are fixed under the action of G (hence $R \subset S^G$). Then S^G is an R-algebra. If $\mathfrak{M}$ is the maximal ideal of S, then S^G is a finite local ring with maximal ideal $S^G \cap \mathfrak{M}$, hence is an extension of R. An extension S of R is *unramified* if the maximal ideal $\mathfrak{m}$ of R generates the maximal ideal of S; otherwise it is said to be *ramified* If S is an unramified extension of R then $\mathfrak{m}^i$ generates $\mathfrak{M}^i$ so the degree of nilpotency of $\mathfrak{m}$ equals the degree of nilpotency of $\mathfrak{M}$. An unramified extension $R \subset S$ is said to be a *Galois extension* if $R = S^G$.

Example Let R be a finite local ring with maximal ideal $\mathfrak{m}$. Let $f \in R[x]$ be a monic basic irreducible polynomial. The extension $S = R[x]/(f^m)$ is again a finite local ring (see Section 4.1). Its maximal ideal is $\mathfrak{M} = \mathfrak{m} + (f)$. If $m > 1$ then S is a *ramified* extension of R. If $m = 1$ then S is an unramified extension and $\mathfrak{M} = \mathfrak{m}S$ is generated by $\mathfrak{m}$.

The following result is the main theorem in the Galois theory of finite local rings. The proof may be found in [141].

Theorem 4.4.1. *Let R be a finite local ring. Then every unramified extension $R \subset S$ is a Galois extension. Suppose $R \subset S$ is such an extension, with corresponding maximal ideals $\mathfrak{m} \subset \mathfrak{M}$. Then*

the following diagram

$$\begin{array}{ccc} S & \xrightarrow{\nu} & K = S/\mathfrak{M} \\ \bigcup & & \bigcup \\ R & \xrightarrow{\mu} & F = R/\mathfrak{m} \end{array} \tag{4.6}$$

induces an isomorphism $Gal(S/R) \cong Gal(K/F)$ *which is therefore a cyclic group. There exists* $h \in S$ *so that* $S = R[h]$*. The mapping determined by* $h \mapsto h^{|F|}$ *generates* $Gal(S/R)$*. Let* $h = h_1, h_2, \ldots, h_d$ *be the distinct images of* h *under* $Gal(S/R)$*. Then the following polynomial*

$$f(x) = (x - h_1)(x - h_2) \cdots (x - h_d) \tag{4.7}$$

actually lies in $R[x]$*. It is a (monic) basic irreducible polynomial of degree* $d = |Gal(S/R)|$*. The mapping* $R[x]/(f) \to S$ *which takes* $x \in R[x]$ *to* $h \in S$ *is an isomorphism of rings (and of* R*-algebras). The ring* S *is a free module of rank* d *over the ring* R*, hence* $|S| = |R|^d$ *and we say that* S *is an extension of* degree d*. The above diagram induces a (combinatorial) lattice preserving bijection between the Galois extensions of* R *which are contained in* S *and the field extensions of* F *which are contained in* K*. The ring* S *is a field if and only if the ring* R *is a field. If* $f' \in R[x]$ *is another monic basic irreducible polynomial of the same degree* d *then there exists an* R*-algebra isomorphism* $S \cong R[x]/(f')$*. In particular,* f' *also splits into linear factors over* S*.*

Corollary 4.4.2. *Let* R *be a finite local ring, let* S *be an unramified degree* d *extension of* R*, and let* $f \in R[x]$ *be a monic basic irreducible polynomial of degree* d*. Let* $\alpha \in S$ *be a root of* f*. Then the collection*

$$\left\{1, \alpha, \alpha^2, \cdots, \alpha^{d-1}\right\}$$

is a basis of S *over* R*. The element* α *is invertible in* S*.*

Proof. According to Theorem 4.4.1, we may replace S with $R[x]/(f)$ and we may replace α with x. By the division theorem for polynomials, the set

$$\left\{1, x, x^2, \cdots, x^{d-1}\right\}$$

is a basis of $R[x]/(f)$ over R. If $f(x) = a_0 + a_1x + \cdots + a_dx^d$ then $\mu(a_0) \neq 0$ since $\mu(f)$ is irreducible. Therefore a_0 is invertible in S and

$$x^{-1} = \frac{-1}{a_0}(a_1 + a_2x^2 + \cdots + a_dx^{d-1})$$

in $R[x]/(f)$. □

4.4.b The Trace and the norm

Let $R, \mathfrak{m}, F = R/\mathfrak{m}$ be a finite local ring with $\mu : R \to F$ the reduction map. Let $S, \mathfrak{M}, K = S/\mathfrak{M}$ be a Galois extension of degree d with $\nu : S \to K$ the reduction map. Let $a \in S$. The *trace* $\mathsf{Tr}_{S/R}(a) \in R$ and *norm* $N_{S/R}(a) \in R$ of a are defined to be

$$\mathsf{Tr}_{S/R}(a) = \sum_{\sigma \in \mathrm{Gal}(S/R)} \sigma(a)$$

and

$$N_{S/R}(a) = \prod_{\sigma \in \mathrm{Gal(S/R)}} \sigma(a).$$

Consider the mapping $\kappa_a : S \to S$ which is given by multiplication by a. Since S is a free module over R it has a basis consisting of d elements, and the mapping κ_a may be expressed as a $d \times d$ matrix M_a. Then the trace and norm of a equal the trace and determinant (respectively) of this matrix (which are thus independent of the choice of basis).

Lemma 4.4.3. *$\mathsf{Tr}_{S/R}(a)$ equals the trace of M_a and $N_{S/R}(a)$ equals the determinant of M_a. Also, we have $\mu \circ \mathsf{Tr}_{S/R} = \mathsf{Tr}_{K/F} \circ \nu$ and $\mu \circ N_{S/R} = N_{K/F} \circ \nu$.*

Proof. The last statement of the theorem follows from Theorem 4.4.1. We know the first statement concerning the trace is true for the fields K and F by Proposition 3.2.14. Let N be the set of elements a of S such that the trace of a equals the trace of M_a. Then N is an R-submodule of S since the mapping from a to the trace of M_a is R-linear. Moreover $S = N + \mathfrak{M}S = N + \mathfrak{m}S$. By Nakayama's lemma (Theorem 4.1.2) we have $S = N$, which proves the claim.

Next we consider the norm. Let us denote the determinant of M_a by $D(a)$. We want to show that $D(a) = N_{S/R}(a)$ for every $a \in S$. Since both $N_{S/R}$ and D are multiplicative, it suffices to show this for a set V such that every element of S is a product of elements of V.

If $a \in R$, then $M_a = aI$ so $D(a) = a^d$, and $N_{S/R}(a) = a^d$.

Suppose that $a \in S$ reduces to a primitive element of K modulo $\mathfrak{M}$. If N is the R-submodule of S spanned by $1, a, \cdots, a^{d-1}$, then $S = N + \mathfrak{M}$, so by Nakayama's lemma $S = N$. That is, $1, a, \cdots, a^{d-1}$ is an R-basis for S. With respect to this basis M_a has the form described in the proof of Proposition 3.2.14. If

$$f(x) = x^d + \sum_{i=0}^{e/d-1} a_i x^i$$

is the minimal polynomial of a over R, then $D(a) = a_0 = N_{S/R}(a)$. Thus $D(a^i) = N_{S/R}(a^i)$ for every i. If n is the degree of nilpotency of S and R, then $|S| = |K|^n$. We have thus far accounted for the $(|K| - 2)|K|^{n-1}$ elements of S that are congruent to some a^i, $i = 1, \cdots, |K| - 2$. We also

have $D((a+b)/a) = N_{S/R}((a+b)/a)$ if $b \in \mathfrak{M}$. This accounts for the $|\mathfrak{M}| = |K|^{n-1}$ elements in $1+\mathfrak{M}$, and hence for all the units. Finally, since $\mathfrak{M} = \mathfrak{m}S$, every element of $\mathfrak{M}$ can be written in the form cb with $c \in \mathfrak{m}^i$ for some i and b a unit. Using multiplicativity again completes the proof. □

Corollary 4.4.4. *The trace $\mathsf{Tr}_{S/R} : S \to R$ is surjective.*

Proof. First we show there exists an element $s \in S$ so that $\mathsf{Tr}(s)$ is invertible in R. If this were false, then we would have $\mathsf{Tr}(s) \in \mathfrak{m}$ for all $s \in S$ which would imply that the induced mapping $S/\mathfrak{M} \to R/\mathfrak{m}$ is 0. This would contradict the above lemma which states that this induced mapping is the trace, $\mathsf{Tr}_{K/F}$, which is surjective. So choose $c \in S$ so that $\mathsf{Tr}_{S/R}(c)$ is invertible and let $a \in R$ denote its inverse. Then for any $b \in R$ we have $\mathsf{Tr}_{S/R}(bac) = baTr_{S/R}(c) = b$. □

Theorem 4.4.5. *Let $\sigma \in Gal(S/R)$ be a generator of the Galois group. Then $\mathsf{Tr}_{S/R}(a) = 0$ if and only if there exists $c \in S$ such that $a = c - \sigma(c)$, and $N_{S/R}(a) = 1$ if and only if there is a unit $b \in S$ so that $a = b\sigma(b)^{-1}$.*

Proof. The ring S is a free module over R of rank d. First we prove the statement about the trace. Let $\phi : S \to S$ be defined by $\phi(x) = x - \sigma(x)$. The kernel of ϕ is R since S is a Galois extension of R. As a homomorphism of R modules the rank of ϕ is $d-1$ because its kernel is 1-dimensional. Therefore the image of ϕ contains $|R|^{d-1}$ elements. The image of ϕ is contained in Ker(Tr) which by Corollary 4.4.4 also contains $|R|^{d-1}$ elements, so they coincide. Thus $Tr(a) = 0$ if and only if $a = b - \sigma(b)$ for some $b \in S$.

The statement concerning the norm is similar, but it uses the function $\psi : S^\times \to S^\times$ defined by $\psi(x) = x\sigma(x)^{-1}$. □

Suppose $L : S \to R$ is any R-linear mapping. Then for any $i \geq 1$ we have $L(\mathfrak{M}^i) \subset \mathfrak{m}^i$. (Since $\mathfrak{M} = \mathfrak{m}S$, any element in $\mathfrak{M}^i$ may be expressed as ac with $a \in \mathfrak{m}^i$ and $c \in S$, in which case $L(ac) = aL(c) \in \mathfrak{m}^i$.) In particular, L induces an F-linear mapping $\bar{L} : K = S/\mathfrak{M} \to F = R/\mathfrak{m}$ and the diagram

$$\begin{array}{ccc} S & \xrightarrow{\nu} & K = S/\mathfrak{M} \\ {\scriptstyle L}\downarrow & & \downarrow{\scriptstyle \bar{L}} \\ R & \xrightarrow{\mu} & F = R/\mathfrak{m} \end{array} \tag{4.8}$$

commutes. Let us say that L is *nonsingular* if this mapping $\bar{L}$ is surjective. This is equivalent to saying that $\bar{L}$ is not the zero map.

Theorem 4.4.6. *Let $L : S \to R$ be an R linear mapping. Then*

1. *The mapping $L : S \to R$ is surjective if and only if L is nonsingular. (In particular, the trace $\mathsf{Tr}_{S/R}$ is nonsingular.)*

2. *If L is nonsingular, then $L(\mathfrak{M}^i) = \mathfrak{m}^i$ for any $i \geq 1$.*
3. *If L is nonsingular, $b \in S$ and $L(ab) = 0$ for all $a \in S$, then $b = 0$.*
4. *There exists $b \in S$ so that $L(a) = \mathsf{Tr}(ba)$ for all $a \in S$. The element b is invertible if and only if L is nonsingular.*

Proof. If L is surjective then it is nonsingular by diagram (4.8). On the other hand, if L is nonsingular then (as above) there exists $b \in S$ such that $L(b)$ is invertible in R. If $a = L(b)^{-1}$ then, for any $c \in R$, $L(cab) = c$ so L is surjective. This proves (1). We already know that $L(\mathfrak{M}^i) \subset \mathfrak{m}^i$ so let $c \in \mathfrak{m}^i$ and, by part (1), let $a_0 \in S$ be an element such that $L(a_0) = 1$. Then $ca_0 \in \mathfrak{M}^i$ and $L(ca_0) = c$, which proves (2).

To prove (3), let n be the degree of nilpotency of $\mathfrak{m}$. That is, $\mathfrak{m}^n = 0$ but $\mathfrak{m}^{n-1} \neq 0$. Then n is also the degree of nilpotency of $\mathfrak{M}$. Let $b \neq 0 \in S$ and suppose that $L(ab) = 0$ for all $a \in S$. Let $m < n$ be the largest integer so that $b \in \mathfrak{M}^m$. Then $b = db_1$ with $d \in \mathfrak{m}^m - \mathfrak{m}^{m+1}$ and b_1 a unit in S. Therefore for all $a \in S$ we have $0 = L(da) = dL(a)$. But $m < n$ so we must have $L(a) \in \mathfrak{M}$ which contradicts the nonsingularity of L, proving (3).

To prove (4), consider the mapping $S \to \mathrm{Hom}_R(S, R)$ which assigns to any $b \in S$ the R linear mapping $a \mapsto \mathsf{Tr}_{S/R}(ab)$. This mapping is injective, for if $b' \in S$ and $\mathsf{Tr}_{S/R}(ab) = \mathsf{Tr}_{S/R}(ab')$ for all $a \in S$, then by part (3) this implies $b = b'$. Since S is a free module over R of some rank d, there are $|R|^d$ elements in $\mathrm{Hom}_R(S, R)$. But this is the same as the number of elements in S. Therefore every R-linear mapping $L : S \to R$ is of the form $a \mapsto \mathsf{Tr}_{S/R}(ab)$ for some $b \in S$. If b is invertible, then the mapping L is nonsingular, whereas if $b \in \mathfrak{M}$ then $L(ab) \in \mathfrak{m}$ so the resulting mapping $\bar{L} : S/\mathfrak{M} \to R/\mathfrak{m}$ is zero. □

4.4.c Primitive polynomials

Let R be a finite local ring with maximal ideal $\mathfrak{m}$ and residue field $\mu : R \to F = R/\mathfrak{m}$. Let S be a degree d Galois extension of R, with maximal ideal $\mathfrak{M}$ and residue field $\nu : S \to K = S/\mathfrak{M}$ as in (4.6). Let $f \in R[x]$ be a basic irreducible polynomial of degree d. Then f is said to be *primitive* if the polynomial $\bar{f} = \mu(f) \in F[x]$ is primitive. That is, if for some (and hence for any) root $\bar{a} \in K$ of $\bar{f}$, the distinct powers of $\bar{a}$ exactly account for all the nonzero elements in K. Unfortunately this is not enough to guarantee that each root $a \in S$ of f generates the cyclic group $\iota(K^\times) \subset S$.

Lemma 4.4.7. *Let $f \in R[x]$ be a basic irreducible polynomial of degree d and let S be a degree d Galois extension of R, so that f splits into linear factors over S. Let $a \in S$ be a root of f. If $\mu(f)$ is primitive (in $F[x]$) then the elements $\{1, a, a^2, \cdots, a^{Q-2}\}$ are distinct, where $Q = |K| = |F|^d$. The roots of f lie in $\iota(K^\times) \subset S^\times$ if and only if f divides $x^Q - 1$. Thus, if $\mu(f)$ is primitive and f divides $x^Q - 1$, then $\iota(K^\times) \subset S^\times$ consists of the $Q - 1$ distinct powers $\{1, a, a^2, \cdots, a^{Q-2}\}$ of a.*

Proof. The element $\mu(a) \in K$ is a root of $\mu(f) \in F[x]$. If $\mu(f)$ is primitive, then $\mu(a)$ is a primitive element in K and the elements $\mu(a)^i$ $(0 \leq i \leq Q-2)$ are distinct, so the same is true of the elements

a^i $(0 \le i \le Q-2)$. By 4.1.4 the polynomial $g(x) = x^{Q-1} - 1$ factors completely in S as

$$g(x) = \prod_{b \in K^\times} (x - \iota(b)).$$

Since f also factors completely over S, we see that the roots of f lie in $\iota(K^\times)$ if and only if f divides $g(x)$. □

4.5 Galois rings

Let $p \in \mathbb{Z}$ be a prime number. According to Theorem 4.4.1, for each $n, d \ge 1$ the ring $\mathbb{Z}/(p^n)$ has a unique Galois extension of degree d. This extension $S = GR(p^n, d)$ is called the *Galois ring* of degree d over $\mathbb{Z}/(p^n)$. For $n = 1$ it is the Galois field $\mathbb{F}_{p^d}$. For $d = 1$ it is the ring $\mathbb{Z}/(p^n)$. Let us review the general facts from Section 4.4 for the case of a Galois ring S.

The Galois ring $S = GR(p^n, d)$ is isomorphic to the quotient ring $\mathbb{Z}/(p^n)[x]/(f)$ where $f \in \mathbb{Z}/(p^n)[x]$ is a monic basic irreducible polynomial. That is, it is a monic polynomial such that its reduction $f \pmod p \in \mathbb{Z}/(p)[x]$ is irreducible. The ring S contains p^{nd} elements. For each divisor e of d the Galois ring S contains the ring $GR(p^n, e)$ and this accounts for all the subrings of S. For any $m \le n$ there is a projection $S \to GR(p^m, d)$ whose kernel is the ideal (p^m), and this accounts for all the nontrivial ideals in S. In particular the maximal ideal $\mathfrak{M} = (p) = pS$ consists of all multiples of p. The quotient $S/\mathfrak{M} \cong \mathbb{F}_{p^d}$ is isomorphic to the Galois field with p^d elements. If μ denotes the projection to this quotient, then it is compatible with the trace mapping in the sense that the following diagram commutes,

$$\begin{array}{ccc} S = GR(p^n, d) & \xrightarrow{\mu} & K = \mathbb{F}_q \\ \mathsf{Tr}\downarrow & & \downarrow\mathsf{Tr} \\ \mathbb{Z}/(p^n) & \xrightarrow[\mu]{} & \mathbb{F}_p \end{array}$$

where $q = p^d$. There is a natural (multiplication-preserving) splitting $\iota : K \to S$ of the mapping μ whose image is the set all elements $x \in S$ such that $x^q = x$. The group of units of S is the product

$$S^\times = \iota(K^\times) \times (1 + \mathfrak{M}).$$

If $p \ge 3$ then

$$1 + \mathfrak{M} \cong \mathbb{Z}/(p^{n-1}) \times \cdots \times \mathbb{Z}/(p^{n-1}) \quad (d \text{ times}).$$

If $p = 2$ and $n \ge 3$ then

$$1 + \mathfrak{M} \cong \big(\mathbb{Z}/(2^{n-1})\big)^{d-1} \times \mathbb{Z}/(2^{n-2}) \times \mathbb{Z}/(2)$$

If $p = 2$ and $n = 1, 2$ then in this equation, each factor $\mathbb{Z}/(2^m)$ should be dropped whenever $m \leq 0$.

It follows that, in general, $S^\times$ contains cyclic subgroups of order $(p^d - 1)p^{n-1}$ and that $|S^\times| = (p^d - 1)p^{d(n-1)}$.

Lemma 4.5.1. *For any $x \in S$ there are unique elements $a_0, a_1, \cdots, a_{n-1} \in \iota(K)$ such that*

$$x = a_0 + a_1 p + \cdots + a_{n-1} p^{n-1}. \tag{4.9}$$

The coefficients $a_0, a_1, \cdots a_{n-1}$ in (4.9) are called the coordinates *of x, and the expansion (4.9) is called the p-*adic expansion *of x.*

Proof. First note that if $t \in \iota(K)$ and if $1 - t$ is not a unit, then $t = 1$. Next, according to the comments in the first paragraph of this section, $|\mathfrak{M}^i/\mathfrak{M}^{i+1}| = q$ for $1 \leq i \leq n-1$. We claim that every element of $\mathfrak{M}^i/\mathfrak{M}^{i+1}$ has a unique representative of the form ap^i where $a \in \iota(K)$. Certainly $ap^i \in \mathfrak{M}^i$ and there are no more than q such elements, so we need to show these elements are distinct modulo $\mathfrak{M}^{i+1}$. Suppose $ap^i \equiv bp^i \pmod{\mathfrak{M}^{i+1}}$ with $a, b \in \iota(K)$. Then $p^i(1 - ba^{-1}) \in \mathfrak{M}^{i+1}$ from which it follows that $1 - ba^{-1} \in \mathfrak{M}$. But $ba^{-1} \in \iota(K)$ so the above note implies that $a = b$.

It now follows by induction that every $x \in \mathfrak{M}^i$ has a unique expression $x = p^i(a_0 + a_1 p + \cdots + a_{n-i-1}p^{n-i-1})$ with $a_i \in \iota(K)$. The coefficient a_0 is the unique representative of $x \pmod{\mathfrak{M}^{i+1}}$, while the inductive step applies to $x - p^i a_0 \in \mathfrak{M}^{i+1}$. □

The advantage of Lemma 4.5.1 is that multiplication by elements in $\iota(K)$ is described coordinatewise. That is, if $b \in \iota(K)$ and if x is given by 4.9, then $ba_0 + ba_1 p + \cdots + ba_{n-1}p^{n-1}$ is the p-adic expansion of bx. Multiplication by p is given by a "shift" of the coefficients a_i. However addition is described using a generalized "carry" procedure: if $a, b \in \iota(K)$ and if $a+b = c_0+c_1 p+\cdots+c_{n-1}p^{n-1}$ is the p-adic expansion of $a + b$ then we may think of the coefficient c_0 as the "sum" and the coefficients c_i (for $i \geq 1$) as being higher "carries".

4.6 Exercises

1. Let R be a finite local ring with maximal ideal $\mathfrak{m}$. Show that

   a.the ideal $\mathfrak{m}$ consists precisely of the non-units of R,
   b.for every $a \in R$, at least one of a and $1 + a$ is a unit, and
   c.there is a positive integer n such that $\mathfrak{m}^n = 0$.

2. Let R be a finite local ring with maximal ideal $\mathfrak{m}$ and residue field $F = R/\mathfrak{m}$. Show that $\mathfrak{m}^{i-1}/\mathfrak{m}^i$ naturally admits the structure of a vector space over F.

3. If R is a local ring and $g \in R[x]$ is regular, then use Nakayama's Lemma to show that for every $f \in R[x]$ there exist $q, r \in R[x]$ with $f = gq + r$ and $\deg(r) < \deg(g)$.

4. Show that for $p = 3$ and $m = 3$, the mapping $E : \mathbb{Z}/(3^2) \to \mathbb{Z}/(3^3)$ of Section 4.2.a is given by

$$E(a) = 1 + 3a + 18a^2 + 18a^3.$$

5. Let S/R be a Galois extension of finite local rings $\sigma \in \mathrm{Gal}(S/R)$ be a generator of the Galois group. Prove that $N_{S/R}(a) = 1$ if and only if there is a unit $b \in S$ so that $a = b\sigma(b)^{-1}$.

Chapter 5 Sequences, Power Series and Adic Rings

The central theme of this work is the design and analysis of sequences by identifying them with algebraic structures. The most common example associates to a sequence $\mathbf{a}$ its *generating function*, the formal power series whose coefficients are the elements of the sequence. This idea has been extremely fruitful, with applications to many disparate areas including probability theory, cryptography, combinatorics, random number generation, and algebraic topology. However, an infinite sequence may also be associated to a p-adic number, a π-adic number, or a reciprocal power series. Despite their differences, these algebraic structures can all be described in terms of a single general construction known as "completion". In this chapter these structures are individually described and an outline of the general theory is given.

5.1 Sequences

In this section we describe basic combinatorial notions concerning sequences. See also Section 11.2.

5.1.a Periodicity

Let A be a set and let $\mathbf{a} = (a_0, a_1, a_2, \cdots)$ be a sequence of elements $a_i \in A$, also called a *sequence over* A. If the set A is discrete (meaning that it is finite or countable) then we refer to A as the *alphabet* from which the *symbols* a_i are drawn. If N is a natural number and $A = \{0, 1, \cdots, N-1\}$, then we refer to $\mathbf{a}$ as an *N-ary sequence*. The sequence $\mathbf{a}$ is *periodic* if there exists an integer $T > 0$ so that

$$a_i = a_{i+T} \tag{5.1}$$

for all $i = 0, 1, 2, \cdots$. Such a T is called *a period* of the sequence $\mathbf{a}$ and the least such T is called *the period*, or sometimes the *least period* of $\mathbf{a}$. The sequence $\mathbf{a}$ is *eventually periodic* if there exists $N > 0$ and $T > 0$ so that equation (5.1) holds for all $i \geq N$. To emphasize the difference, we sometimes refer to a periodic sequence as being *purely periodic* or *strictly periodic*. A *period* (resp. the *least period*) of an eventually periodic sequence refers to a period (resp. least period) of the periodic part of $\mathbf{a}$.

Lemma 5.1.1. *Suppose* $\mathbf{a}$ *is a periodic (or eventually periodic) sequence with least period* T*. Then every period of* $\mathbf{a}$ *is a multiple of* T*.*

Proof. If T' is a period of $\mathbf{a}$, then dividing by T gives $T' = qT + r$ for some quotient $q \geq 1$ and remainder r with $0 \leq r \leq T-1$. Since both T and T' are periods, $a_{i+T'} = a_{i+qT+r} = a_{i+r}$ for all $i \geq 0$. Therefore r is a period also. Since $r < T$, the minimality of T implies $r = 0$. □

5.1.b Distinct sequences

Let A be an alphabet and let $\mathbf{a} = (a_0, a_1, \cdots)$ and $\mathbf{b} = (b_0, b_1, \cdots)$ be sequences of elements of A. We say that $\mathbf{b}$ is a *shift* of $\mathbf{a}$ if there exists $\tau \geq 0$ so that $b_i = a_{i+\tau}$ for all $i \geq 0$. We write $\mathbf{b} = \mathbf{a}^\tau$. If no such shift τ exists then we say that $\mathbf{a}$ and $\mathbf{b}$ are *shift distinct*. If $\mathbf{a}$ and $\mathbf{b}$ are periodic with the same period and $\mathbf{b}$ is a shift of $\mathbf{a}$, then we say that $\mathbf{b}$ is a *left shift* of $\mathbf{a}$. If no such shift exists then $\mathbf{a}$ and $\mathbf{b}$ are *shift distinct*. More generally, if $\mathbf{a}$ is a sequence over an alphabet A and $\mathbf{b}$ is a sequence over an alphabet B, we say that $\mathbf{a}$ and $\mathbf{b}$ are *isomorphic* if there exists an isomorphism of sets $\sigma : A \to B$ so that $b_i = \sigma(a_i)$ for all $i \geq 0$. (If $A = B$ are the same alphabet then σ is just a permutation of the symbols in the alphabet.) We say the sequences $\mathbf{a}$ and $\mathbf{b}$ are *isomorphic up to a shift* if there exists an isomorphism $\sigma : A \to B$ and a shift τ such that $b_i = \sigma(a_{i+\tau})$ for all $i \geq 0$. If no such pair σ, τ exists then we say that $\mathbf{a}$ and $\mathbf{b}$ are *non-isomorphic, even after a shift*.

5.1.c Sequence generators and models

The sequences described in this book are generated by algebraic methods involving rings. We formalize constructions of this type by defining a *sequence generator*. In the models we encounter, the state space of the sequence generator usually corresponds to a cyclic subgroup of the group of units in a ring.

Definition 5.1.2. *A* sequence generator*, or* discrete state machine with output

$$F = (U, \Sigma, f, g)$$

consists of a set U *of* states*, an alphabet* Σ *of* output values*, a state transition function* $f : U \to U$ *and an output function* $g : U \to \Sigma$*.*

Such a generator is depicted as follows:

$$f \circlearrowright U \xrightarrow{\text{g}} \Sigma.$$

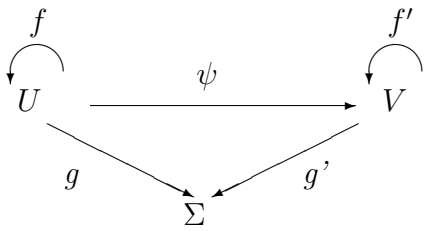
The set U of states is assumed to be *discrete*, meaning that it is either finite or countably infinite. We also assume the alphabet Σ of possible output values is discrete. Given an initial state $\mathbf{s} \in U$, such a sequence generator outputs an infinite sequence

$$F(\mathbf{s}) = g(\mathbf{s}), g(f(\mathbf{s})), g(f^2(\mathbf{s})), \cdots$$

with elements in Σ. A state $\mathbf{s} \in U$ is *aperiodic* if, starting from $\mathbf{s}$, the generator never returns to this state. The state $\mathbf{s}$ is *periodic* of period L if starting from $\mathbf{s}$, after L steps, the generator returns to the state $\mathbf{s}$. That is, if $f^L(\mathbf{s}) = \mathbf{s}$. The *least period* of such a periodic state is the least such $L \geq 1$. A state $\mathbf{s}$ is *eventually periodic* if, starting from $\mathbf{s}$, after a finite number of steps, the generator arrives at a periodic state. If U is finite then every state is eventually periodic. We say a set of states is *closed* if it is closed under state change. It is *complete* if it consists of all the periodic states. If a state $\mathbf{s}$ is periodic (resp., eventually periodic), then the output sequence $F(\mathbf{s})$ is periodic (resp., eventually periodic) as well. The converse is false, however. For example, let $F = (\mathbb{N}, \{0,1\}, f, g)$ with $f(n) = n + 1$ and $g(n) = 0$ for all n. Then the output sequence from every state is periodic but no state is even eventually periodic.

Definition 5.1.3. *Let $F = (U, \Sigma, f, g)$ and $G = (V, \Sigma, f', g')$ be sequence generators. A* homomorphism *from F to G is a partial function ψ from U to V so that*

1. *for all $\mathbf{s} \in U$, if a is in the domain of ψ, then $f(a)$ is also in the domain of ψ, and*
2. *the following diagram commutes:*



That is, $g'(\psi(a)) = g(a)$ and $\psi(f(a)) = f'(\psi(a))$ for all a in the domain of ψ.

If R is a ring and $b \in R$, then let $h_b : R \to R$ denote multiplication by b. That is, $h_b(x) = bx$. Then for any function $T : R \to \Sigma$, the 4-tuple $R^{b,T} = (R, \Sigma, h_b, T)$ is a sequence generator.

Definition 5.1.4. *Let $F = (U, \Sigma, f, g)$ be a sequence generator. An* algebraic model *or simply a* model *for F is a homomorphism ψ of sequence generators between F and $R^{b,T}$ for some ring R, $b \in R$, and $T : R \to \Sigma$. The model is* injective *if $\psi : R^{b,T} \to F$ and the model is* projective *if $\psi : F \to R^{b,T}$.*

In the case of an injective model, if a is in the domain of ψ, then the output sequence generated from $\psi(a)$ is described by the *exponential representation*,

$$T(a), T(ba), T(b^2a), \cdots.$$

In the case of a projective model, if $\mathbf{s}$ is in the domain of ψ, then the output sequence generated from b is described by the *exponential representation*,

$$T(\psi(\mathbf{s})), T(b\psi(\mathbf{s})), T(b^2\psi(\mathbf{s})), \cdots.$$

If the ring R is a finite field, then every such sequence is strictly periodic (because $b^k a = b^{k+r} a$ implies that $a = b^r a$). We say that the model is *complete* if every periodic state $\mathbf{s} \in \Sigma$ is in the range (in the injective case) or domain (in the projective case) of ψ. A complete model, if one exists, allows us to analyze the behavior of the sequence generator using the algebraic structure of the ring R. In this book we encounter many different types of sequence generators and their models.

If ψ is a one to one mapping on its domain, then it can be inverted (possibly resulting in a partial function), allowing us to replace a projective model with an injective model, or vice versa. In practice, however it may require a nontrivial amount of computation to describe the inverse mapping, particularly when attempting to describe the initial state of the generator, cf. (6.5), (7.4), (8.5). Thus one or the other version may be a more natural way to describe a model.

5.2 Power series

5.2.a Definitions

Throughout this section we fix a commutative ring R (with identity 1).

Definition 5.2.1. *A (formal)* power series over R *is an infinite expression*

$$a(x) = a_0 + a_1 x + a_2 x^2 + \cdots,$$

where x is an indeterminate and $a_0, a_1, \cdots \in R$. As with polynomials, the a_is are called coefficients. *The sequence $(a_0, a_1, \cdots)$ of coefficients of a power series $a(x)$ is denoted* $\mathbf{seq}(a)$. *If $b(x) = b_0 + b_1 x + b_2 x^2 + \cdots$ is a second power series over R, then define*

$$(a+b)(x) = a(x) + b(x) = \sum_{i=0}^{\infty} (a_i + b_i) x^i$$

and

$$(ab)(x) = a(x)b(x) = \sum_{i=0}^{\infty} (\sum_{j=0}^{i} a_j b_{i-j}) x^i.$$

The set of power series over R is denoted $R[[x]]$. The least degree *of a nonzero power series $a(x) = \sum_{i=0}^{\infty} a_i x^i$ is the least index i such that $a_i \neq 0$. The least degree of 0 is ∞.*

These operations make $R[[x]]$ into a ring with identity given by the power series $1 = 1 + 0x + 0x^2 + \cdots$. The following lemma concerns a remarkable property of the ring of power series: most elements have inverses in $R[[x]]$ and it is easy to determine when an element is invertible.

Lemma 5.2.2. *Let $b(x) = \sum_{i=0}^{\infty} b_i x^i \in R[[x]]$ be a power series. Then (1) b is invertible in $R[[x]]$ if and only if (2) the constant term $b_0 \in R$ is invertible in R.*

Proof. The constant term of a product is the product of the constant terms, so (1) $\Rightarrow$ (2). We claim that (2) $\Rightarrow$ (1). If b_0 is invertible then the equation $b(x)c(x) = 1$ may be solved inductively for $c(x) = \sum_{i=0}^{\infty} c_i x^i$ because $c_0 = b_0^{-1}$ and

$$c_i = -b_0^{-1} \left(b_1 c_{i-1} + b_2 c_{i-1} + \cdots + b_i c_0 \right). \qquad \square$$

The set of polynomials over R is the subring of $R[[x]]$ consisting of those power series with finitely many nonzero coefficients. In fact there is a chain of subrings,

$$\begin{array}{ccccccccccc} R & \subset & R[x] & \subset & E & \subset & R_0(x) & \subset & R[[x]] & \subset & R((x)) \\ & & & & & & \cap & & & & \\ & & & & & & R(x) & & & & \end{array}$$

which is described in the next few sections. The ring $R((x))$ of *formal Laurent series* consists of infinite sums

$$a(x) = a_{-m} x^{-m} + a_{-m+1} x^{-m+1} + \cdots + a_0 + a_1 x + \cdots$$

with coefficients $a_i \in R$ and at most finitely many nonzero terms of negative degree. Addition and multiplication are defined as with power series. The ring of *rational functions* $R(x)$ consists of all fractions $f(x)/g(x)$ where $f, g \in R[x]$ and g is not a zero divisor. So $R(x) = S_1^{-1} R[x]$ is the full ring of fractions, obtained by inverting the set $S_1 \subset R[x]$ consisting of all nonzero-divisors, cf. Section 2.2.h. (The ring $R(x)$ is usually of interest only when R is a field, in which case $S_1 = R[x] - \{0\}$ consists of the nonzero polynomials, cf. Section 5.2.d.) The rings $R_0(x)$ and E merit special attention.

5.2.b Recurrent sequences and the ring $R_0(x)$ of fractions

Let $S_0 \subset R[x]$ denote the multiplicative subset consisting of all polynomials $b(x)$ such that the constant term $b_0 = b(0) \in R$ is invertible in R and define (cf. Section 2.2.h)

$$R_0(x) = S_0^{-1} R[x]$$

to be the ring of fractions $a(x)/b(x)$ with $b(x) \in S_0$. We obtain an injective homomorphism $\psi : R_0(x) \to R[[x]]$ by mapping $a(x)/b(x)$ to the product $a(x)c(x)$ where $c(x) \in R[[x]]$ is the power series inverse of $b(x)$ which was constructed in Lemma 5.2.2. The series

$$\psi(a(x)/b(x)) = a_0 + a_1 x + a_2 x^2 + \cdots \in R[[x]]$$

is referred to as the *power series expansion* of the fraction $a(x)/b(x)$, and we write

$$\mathbf{a} = \mathbf{seq}(a(x)/b(x)).$$

Henceforth we identify $R_0(x)$ with its image in $R[[x]]$.

A sequence $\mathbf{a} = a_0, a_1, \cdots$ of elements of R is *linearly recurrent* (of degree d) if there exist $q_1, \cdots, q_d \in R$ (with $q_1 \neq 0, q_d \neq 0$) such that for all $n \geq d$ we have

$$a_n = q_1 a_{n-1} + \cdots + q_d a_{n-d}. \tag{5.2}$$

More generally, we say that $\mathbf{a}$ satisfies a recurrence of degree d for $n \geq N$ if equation (5.2) holds for all $n \geq N$. The following theorem characterizes the ring $R_0(x)$ as consisting of those power series $a(x)$ having linearly recurrent coefficient sequences.

Theorem 5.2.3. *Let $a = a_0 + a_1 x + \cdots \in R[[x]]$ be a formal power series. Fix $N \geq d > 1$. The following statements are equivalent.*

1. *There exist polynomials $f(x), g(x) \in R[x]$ such that $g(0)$ is invertible, $\deg(g) = d$, $\deg(f) < N$, and $a(x) = f(x)/g(x)$.*
2. *For all $n \geq N$ the sequence of coefficients $a_n, a_{n+1}, a_{n+2}, \cdots = \mathbf{seq}(f/g)$ satisfies a linear recurrence, of degree d.*

Proof. First suppose that statement (1) holds, say $a(x) = f(x)/g(x)$ with $g(x) = g_0 + g_1 x + \cdots + g_d x^d$ and $g_d \neq 0$. Then $f(x) = a(x)g(x)$ which gives

$$f_n = \sum_{i=0}^{d} g_i a_{n-i}$$

for $n \geq d$. Since $f(x)$ is a polynomial, these coefficients vanish for $n > \deg(f)$. Consequently, if $n \geq N \geq \max(d, \deg(f) + 1)$ we have,

$$a_n = -g_0^{-1} \left(g_1 a_{n-1} + g_2 a_{n-2} + \cdots + g_d a_{n-d} \right)$$

which is a linear recurrence (of degree d). Conversely, suppose the coefficients of f satisfy a linear recurrence $a_n = g_1 a_{n-1} + \cdots + g_d a_{n-d}$ (with $g_d \neq 0$) for all $n \geq N$. Let $g(x) = -1 + g_1 x + \cdots + g_d x^d$ and set $g_0 = -1$. Then the product $f(x) = g(x)a(x)$ is a polynomial of degree less than N, because for $n \geq N$ its term of degree n is

$$\sum_{i=0}^{d} g_i a_{n-i} = 0.$$

Consequently $a(x) = f(x)/g(x)$, g_0 is invertible, and $\deg(f) < N$. □

5.2.c Eventually periodic sequences and the ring E

Definition 5.2.4. *The ring $E \subset R[[x]]$ is the collection of all power series $a(x) = \sum_{i=0}^{\infty} a_i x^i$ such that the sequence of coefficients $\mathbf{seq}(a) = (a_0, a_1, \cdots)$ is eventually periodic.*

Theorem 5.2.5. *Let $a(x) = \sum_{i=0}^{\infty} a_i x^i$ be a power series over a ring R and let $n \geq 1$. Then the following are equivalent. (See also Lemma 2.4.6.)*

1. *The sequence $\mathbf{seq}(a) = (a_0, a_1, \cdots)$ is eventually periodic and n is a period of $\mathbf{seq}(a)$.*
2. *$a(x) = h(x)/(x^n - 1)$ for some $h(x) \in R[x]$.*
3. *$a(x) = f(x)/g(x)$ for some $f, g \in R[x]$ such that $g(x)$ is monic and $g(x)|(x^n - 1)$.*
4. *$a(x) = f(x)/g(x)$ for some $f, g \in R[x]$ such that $g(x)|(x^n - 1)$.*

These statements imply

5. *$a(x) = f(x)/g(x)$ for some $f, g \in R[x]$ such that $g(0)$ is invertible in R.*

Hence $E \subseteq R_0(x)$. The eventual period is the least n for which (2), (3), or (4) holds. If R is finite then statement (5) implies the others (for some $n \geq 1$), so $E = R_0(x)$. (In other words, if R is finite then a sequence over R satisfies a linear recurrence if and only if it is eventually periodic.)

The sequence $\mathbf{seq}(a)$ is purely periodic if and only if (2) holds with $\deg(h(x)) < n$ or equivalently, if (3) or (4) holds with $\deg(f(x)) < \deg(g(x))$.

Proof. To see that condition (1) implies condition (2), suppose $a(x)$ is eventually periodic with $a_i = a_{i+n}$ for all $i \geq N$. Then we have

$$\begin{aligned} a(x) &= \sum_{i=0}^{N-1} a_i x^i + x^N \sum_{j=0}^{\infty} \left(\sum_{k=0}^{n-1} a_{nj+i+N} x^i \right) x^{nj} \\ &= \frac{(x^n - 1)(\sum_{i=0}^{N-1} a_i x^i) - x^N \sum_{k=0}^{n-1} a_{nj+i+N} x^i}{x^n - 1}. \end{aligned}$$

This can be written as a rational function with denominator $x^n - 1$.

That conditions (2), (3) and (4) are equivalent is left to the reader. In case (3) or (4), if $b(x)g(x) = x^n - 1$, then $\deg(b(x)f(x)) < n$ if and only if $\deg(f(x)) < \deg(g(x))$, which reduces the statements about purely periodic power series to the statement about purely periodic power series in case (2).

To see that condition (2) implies condition (1), suppose $a(x) = h(x)/(x^n - 1)$ with $h(x) \in R[x]$. By the division theorem we can write $h(x) = (x^n - 1)u(x) + v(x)$ with $u(x), v(x) \in R[x]$ and

$\deg(v(x)) < n$. Thus

$$\begin{aligned} a(x) &= u(x) + \frac{v(x)}{x^n - 1} \\ &= u(x) + (v(x) + x^n v(x) + x^{2n} v(x) + \cdots). \end{aligned}$$

The power series $v(x) + x^n v(x) + x^{2n} v(x) + \cdots$ is strictly periodic since there is no overlap among the degrees of the monomials in any two terms $x^{in} v(x)$ and $x^{jn} v(x)$. The addition of $u(x)$ only affects finitely many terms, so the result is eventually periodic. Also, the sequence is periodic if and only if $u(x) = 0$, which is equivalent to $\deg(h(x)) < n$.

It follows immediately that the eventual period is the least n for which (2), (3), or (4) holds. Lemma 2.4.6 says that (4) implies (5), and if R is finite, then (5) implies (4) (for some n). □

It is not always true that $E = R_0(x)$: take $R = \mathbb{Z}$, $g(x) = 1 - 2x$, and $f(x) = 1$. Then $a(x) = 1 + 2x + 4x^2 + \cdots$ which is not eventually periodic.

5.2.d When R is a field

Theorem 5.2.6. *If R is a field, then $R(x) \subset R((x))$ and both of these are fields. (The former is called the* field of rational functions *over R; it is a* global field*). They are the fraction fields of $R[x]$ and $R[[x]]$ respectively. The only non-trivial ideals in $R[[x]]$ are the principal ideals (x^m) for $m \geq 1$.*

Proof. The only nontrivial statement in this theorem concerns the ideal structure of $R[[x]]$. Suppose that I is a nonzero ideal in $R[[x]]$. Let $a(x)$ be an element of I whose least degree nonzero term has the smallest possible degree, n. Then we have $a(x) = x^n b(x)$ for some $b(x) \in R[[x]]$, and the constant term of $b(x)$ is nonzero. By Lemma 5.2.2, $b(x)$ is invertible in $R[[x]]$. Hence $x^n \in I$. Moreover, every element of I has least degree $\geq n$, so can be written as $x^n c(x)$ for some $c(x) \in R[[x]]$. Hence $I = (x^n)$. □

5.2.e $R[[x]]$ as an inverse limit

The quotient ring $R[x]/(x^i)$ may be (additively, but not multiplicatively) identified with the collection of all polynomials of degree $\leq i - 1$. Let

$$\psi_i : R[[x]] \to R[x]/(x^i)$$

be the homomorphism that associates to each $a = \sum_{i=0}^{\infty} a_i x^i$ the partial sum (that is, the polynomial)

$$\psi_i(a) = a_0 + a_1 x + \cdots + a_{i-1} x^{i-1}.$$

These homomorphisms are compatible in the sense that if $k \leq i$ then

$$\psi_{i,k}(\psi_i(a)) = \psi_k(a)$$

where

$$\psi_{i,k} : R[x]/(x^i) \to R[x]/(x^k)$$

is reduction modulo x^k. The next lemma says that every element of $R[[x]]$ can be described in terms of such a sequence of partial sums.

Lemma 5.2.7. *Suppose $s_1, s_2, \cdots$ is a sequence with $s_i \in R[x]/(x^i)$. Assume these elements are* compatible *in the sense that $\psi_{i,k}(s_i) = s_k$ for every pair $k \leq i$. Then there is a unique element $a \in R[[x]]$ such that $\psi_i(a) = s_i$ for all $i \geq 1$.*

Proof. The element $a = \sum_{i=0}^{\infty} a_i x^i$ is given by $a_i = (\psi_{i+1}(a) - \psi_i(a))/x^i$. □

This lemma implies that $R[[x]] = \varprojlim \{R[x]/(x^i)\}$, is the inverse limit of the system of rings $R[x]/(x^i)$. See Section 2.2.1 to recall the definition of inverse limits. Specifically, the set of rings $\{R[x]/(x^i)\}$ is a directed system indexed by the positive integers, with the reduction functions $\psi_{i,k}$. Thus there is a homomorphism ψ from $R[[x]]$ to $\varprojlim \{R[x]/(x^i)\}$ so that if

$$\varphi_i : \varprojlim \{R[x]/(x^i)\} \to R[x]/(x^i)$$

is the projection function, then $\psi_i = \varphi_i \circ \tau$. This is shown in Figure 5.1.

We claim that ψ is an isomorphism. Lemma 5.2.7 says that ψ is surjective. If $a(x) = \sum_{i=0}^{\infty} a_i x^i \in R[[x]]$ is nonzero, then $a_i \neq 0$ for some i. Then $\psi_i(a) \neq 0$, so also $\psi(a) \neq 0$. Thus ψ is also injective, and thus an isomorphism.

Corollary 5.2.8. *Let M be an Abelian group. For $i = 1, 2, \cdots$ let $\tau_i : M \to R[x]/(x^i)$ be group homomorphisms satisfying $\tau_i = \psi_{j,i} \circ \tau_j$ whenever $i \leq j$. Then there is a unique homomorphism $\tau : M \to R[[x]]$ so that $\tau_i = \psi_i \circ \tau$ for $i = 1, 2, \cdots$. If M is also a module over R (respectively, an algebra) and the τ_i are R-module homomorphisms (resp., ring homomorphisms), then so is τ.*

5.3 Reciprocal Laurent series

Let K be a field, let $g(x) \in K[x]$ be a polynomial of degree d. The *reciprocal polynomial* is the polynomial $g^*(y) = y^d g(1/y)$.

It is straightforward to check that $(gh)^* = g^* h^*$ for any $h \in K[x]$, and that $(g^*)^* = g$ if and only if $g(0) \neq 0$. If the polynomial g has nonzero constant term, then it is irreducible (resp. primitive)

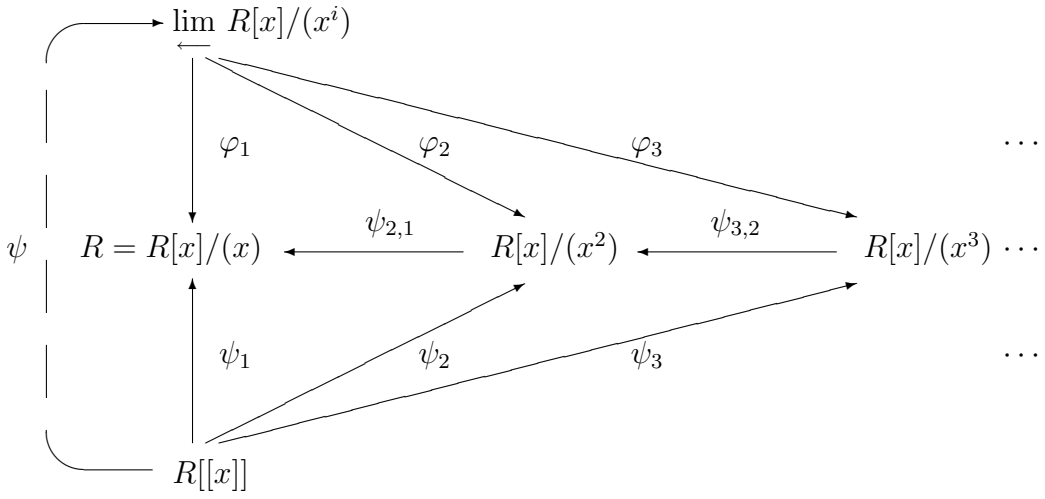


Figure 5.1: $R[[x]]$ as an inverse limit

if and only if the same is true of g^*. If $\alpha \neq 0$ is a root of g (possibly in some extension field of K) then α^{-1} is a root of g^*. Let

$$F = K((x^{-1})) = \left\{ \sum_{i=k}^{\infty} a_i x^{-i} : k \in \mathbb{Z}, a_i \in K \right\}$$

be the ring of formal Laurent series in x^{-1}. According to Theorem 5.2.6, F is a field. Every polynomial $f(x) \in K[x]$ is in F. Therefore F also contains every rational function $f(x)/g(x)$ (where $f, g \in K[x]$). Such a function can therefore be expanded as a Laurent series in x^{-1}. Some of these will be a power series in x^{-1}. It is sometimes helpful, by analogy with the real numbers, to think of $\sum_{i=k}^{0} a_i x^{-i}$ as the integer part of $a = \sum_{i=k}^{\infty} a_i x^{-i}$, and to think of $\sum_{i=1}^{\infty} a_i x^{-i}$ as the fractional part. The degree of a is $-k$ if $a_k \neq 0$.

Proposition 5.3.1. *Let $f, g \in K[x]$ be polynomials. Then $\deg(f) \leq \deg(g)$ if and only if the rational function $f(x)/g(x)$ is actually a power series in x^{-1}, that is,*

$$\frac{f(x)}{g(x)} = a_0 + a_1 x^{-1} + a_2 x^{-2} + \cdots \in K[[x^{-1}]], \tag{5.3}$$

and $\deg(f) < \deg(g)$ if and only if $a_0 = 0$. The sequence $\mathbf{a} = a_0, a_1, a_2, \cdots$ satisfies a linear recurrence with connection polynomial g^*. Conversely, any element of $K[[x^{-1}]]$ whose coefficients

satisfy a linear recurrence with connection polynomial g^* may be expressed as a rational function as in equation (5.3). The sequence $\mathbf{a}$ *is eventually periodic if and only if there exists N so that $g(x)|(x^N-1)$ (which always holds if K is a finite field). The sequence* $\mathbf{a}$ *is strictly periodic if and only if it is eventually periodic and $f(0)=0$ (meaning that $f(x)$ is divisible by x).*

Proof. Let $a(x)=a_0+a_1x^{-1}+\cdots$, let $f(x)=f_0+f_1x+\cdots+f_rx^r$ and $g(x)=g_0+g_1x+\cdots+g_dx^d$. If equation (5.3) holds then $g(x)a(x)=f(x)$ is a polynomial of degree $r=\deg(g)+\deg(a)\le d$, whose degree d term is $f_d=g_da_0$, which vanishes when $a_0=0$. The absence of negative powers of x in $f(x)$ exactly says that the sequence $\mathbf{a}$ satisfies a linear recurrence with connection polynomial g^*. Next, suppose the sequence $\mathbf{a}$ is eventually periodic. Then there exists N such that $g^*(y)|(y^N-1)$, from which it follows that $g(x)|(x^N-1)$ (and vice versa). Finally, let $\widehat{f}(y)=y^df(1/y)=f_0y^d+f_1y^{d-1}+\cdots+f_ry^{d-r}$ so that

$$\frac{\widehat{f}(y)}{g^*(y)}=\frac{f(1/y)}{g(1/y)}=a_0+a_1y+a_2y^2+\cdots.$$

The sequence $\mathbf{a}$ is strictly periodic when $\deg(\widehat{f})<\deg(g^*)=d$, i.e., when $f_0=0$. □

5.4 N-Adic numbers

5.4.a Definitions

The p-adic numbers were discovered by K. Hensel around 1900. He was pursuing the idea that numbers were like functions – the p-ary expansion of an integer is like a polynomial, so what number corresponds to a power series? His ideas led to many far-reaching discoveries that have shaped much of the modern approach to number theory. Fix an integer $N\ge 2$.

Definition 5.4.1. *An* N-adic integer *is an infinite expression*

$$a=a_0+a_1N+a_2N^2+\cdots, \tag{5.4}$$

where $a_0,a_1,\cdots\in\{0,1,\cdots,N-1\}$. The set of N-adic integers is denoted by $\mathbb{Z}_N$. The least degree *of a nonzero N-adic integer $a=\sum_{i=0}^{\infty}a_iN^i$ is the least index i such that $a_i\neq 0$. The least degree of 0 is ∞.*

The a_i are called *coefficients*.When writing N-adic integers we may omit terms whose coefficients are zero. We may also write the terms in a different order. A series such as equation (5.4) does not converge in the usual sense. Nevertheless it can be manipulated as a formal object, just

as with power series, but with slightly different algebraic rules. Addition and multiplication are defined so as to take into account the "carry" operation. To be precise, the statement

$$\sum_{i=0}^{\infty} a_i N^i + \sum_{i=0}^{\infty} b_i N^i = \sum_{i=0}^{\infty} c_i N^i \tag{5.5}$$

with $a_i, b_i, c_i \in \{0, 1, \cdots, N-1\}$ means that there exist integers $t_0, t_1, \cdots \in \{0, 1\}$ so that

$$a_0 + b_0 = c_0 + N t_0 \tag{5.6}$$

and for all $i \geq 1$,

$$a_i + b_i + t_{i-1} = c_i + N t_i. \tag{5.7}$$

The quantity t_i is called the *carry* and it is 0 or 1 since (by induction) $a_i + b_i + t_{i-1} \leq 2(N-1)+1 < 2N$. Also by induction the numbers t_i, c_i are determined by the a_k, b_k. In fact,

$$c_n = (a_n + b_n + t_{n-1}) \pmod{N} \quad \text{and} \quad t_n = \lfloor (a_n + b_n + t_{n-1})/N \rfloor$$

(with $t_{-1} = 0$). The product $ab = c$ is defined similarly with

$$\sum_{i=0}^{n} a_i b_{n-i} + t_{n-1} = c_n + N t_n, \tag{5.8}$$

although in this case the carry t_i may be greater than 1. (Some readers may find it easier to think in terms of power series in some indeterminate, say, Y, and to use the rule that $NY^i = Y^i + Y^i + \cdots + Y^i = Y^{i+1}$. But this notation quickly becomes cumbersome. The use of N instead of Y facilitates many computations.)

It is easy to see that these operations make $\mathbb{Z}_N$ into a ring (the ring axioms hold in $\mathbb{Z}_N$ because they hold modulo N^k for every k). As with power series, we refer to the sequence $(a_0, a_1, \cdots)$ of coefficients as $\mathbf{seq}_N(a)$. It is an N-ary sequence (that is, a sequence over the alphabet $\{0, 1, \cdots, N-1\}$). We say that a is periodic (resp. eventually periodic) if the sequence $\mathbf{seq}_N(a)$ of coefficients is periodic (resp. eventually periodic).

If $a = \sum_{i=0}^{\infty} a_i N^i$ is an N-adic integer, then the coefficient a_0 is called the *reduction of a modulo* N and it is denoted $a_0 = a \pmod{N}$. This gives a ring homomorphism $\mathbb{Z}_N \to \mathbb{Z}/(N)$. We also define the *integral quotient* of a by N to be

$$a \text{ (div } N) = \sum_{i=0}^{\infty} a_{i+1} N^i = \frac{a - a_0}{N}.$$

Thus $a = a \pmod{N} + N(a \text{ (div } N))$.

In the ring $\mathbb{Z}_N$ we have an identity, $-1 = (N-1) + (N-1)N + (N-1)N^2 + \cdots$, which can be verified by adding 1 to both sides. Similarly, there is an explicit formula for multiplication by -1. If $a = \sum_{i=d}^{\infty} a_i N^i$ with $1 \le a_d \le N-1$, then

$$-a = (N - a_d)N^d + \sum_{i=1}^{\infty} (N - a_i - 1)N^i. \tag{5.9}$$

It follows that $\mathbb{Z}_N$ contains the integers as a subring, and in fact there is a chain of rings, similar to that of Section 5.2.a,

$$\begin{array}{ccccccc} \mathbb{Z} & \subset & \mathbb{Z}_{N,0} & \subset & \mathbb{Z}_N & \subset & \mathbb{Q}_N \\ & & \cap & & & & \\ & & \mathbb{Q} & & & & \end{array}$$

The following analog to Lemma 5.2.2 characterizes the invertible elements of $\mathbb{Z}_N$.

Lemma 5.4.2. *Let $a = \sum_{i=0}^{\infty} a_i N^i \in \mathbb{Z}_N$. Then a is invertible in $\mathbb{Z}_N$ if and only if a_0 is relatively prime to N.*

Proof. The proof is essentially the same as that of Lemma 5.2.2. Recall from Section 2.2.d that $a_0 \in \mathbb{Z}$ is relatively prime to N if and only if a_0 is invertible in $\mathbb{Z}/(N)$. We want to find $b = \sum_{i=0}^{\infty} b_i N^i$ so that $ab = 1$, and $0 \le b_i \le N-1$. By equation (5.8) this means $a_0 b_0 = 1 + N t_0$ (which has the unique solution $b_0 = a_0^{-1} \pmod N$ and $t_0 = a_0 b_0 - 1 \text{ (div } N)$) and

$$\sum_{i=0}^{n} a_i b_{n-i} + t_{n-1} = c_n + N t_n,$$

which has the unique solution recursively given by

$$\begin{aligned} b_n &= a_0^{-1} \left(c_n - t_{n-1} - \sum_{i=1}^{n} a_i b_{n-i} \right) \pmod N \\ t_n &= \left(\sum_{i=0}^{n} a_i b_{n-i} - c_n \right) \text{ (div } N). \end{aligned}$$

□

5.4.b The ring $\mathbb{Q}_N$

The ring $\mathbb{Q}_N$ of *N-adic numbers* is the analog of formal Laurent series; it consists of infinite sums

$$a(x) = a_{-m} N^{-m} + a_{-m+1} N^{-m+1} + \cdots + a_0 + a_1 N + \cdots$$

with coefficients $0 \le a_i \le N-1$ and at most finitely many nonzero terms of negative degree. Addition and multiplication are defined as with N-adic integers. We have $\mathbb{Q}_N = S^{-1}\mathbb{Z}_N$ where $S = \{N, N^2, N^3 \cdots\}$.

It follows from Lemma 5.4.2 that if $N = p$ is a prime number, then $\mathbb{Z}_p$ is an integral domain and $\mathbb{Q}_p$ is its fraction field, that is, $\mathbb{Q}_p = S^{-1}\mathbb{Z}_p$ where $S = \mathbb{Z}_p^\times$ consists of all nonzero elements (cf. Section 2.2.h). For composite N, the ring $\mathbb{Z}_N$ has zero divisors and the ring $\mathbb{Q}_N$ is not a field. However in Corollary 5.4.9 we show that $\mathbb{Q}_N = S^{-1}\mathbb{Z}_N$ is the "full" ring of fractions, meaning that the set S consists of all nonzero-divisors in $\mathbb{Z}_N$. In Theorem 5.4.8 we show that $\mathbb{Z}_N$ and $\mathbb{Q}_N$ can be described in terms of $\mathbb{Z}_p$ and $\mathbb{Q}_p$ as p ranges over the prime divisors of N. For these reasons, the rings $\mathbb{Z}_N$ and $\mathbb{Q}_N$ (with N composite) are seldom encountered in the mathematical literature. However we make use of them when studying sequences generated by an FCSR in Chapter 7.

5.4.c The ring $\mathbb{Z}_{N,0}$

Definition 5.4.3. *The ring $\mathbb{Z}_{N,0}$ consists of the set of all rational numbers $a/b \in \mathbb{Q}$ (in lowest terms) such that b is relatively prime to N. That is, $\mathbb{Z}_{N,0} = S^{-1}\mathbb{Z}$, where S is the multiplicative set $\{b \in \mathbb{Z} : \gcd(b, N) = 1\}$.*

Lemma 5.4.2 says that $\mathbb{Z}_{N,0}$ is naturally contained in the N-adic integers $\mathbb{Z}_N$ and it is a subring. The next theorem identifies $\mathbb{Z}_{N,0}$ as the collection of N-adic integers $a \in \mathbb{Z}_N$ such that $\mathbf{seq}_N(a)$ is eventually periodic.

Theorem 5.4.4. *Let $a = \sum_{i=0}^{\infty} a_i N^i \in \mathbb{Z}_N$ and let $n \ge 1$. Then the following statements are equivalent.*

1. *$a = f/g$ for some $f, g \in \mathbb{Z}$ such that $g > 0$ is relatively prime to N and $\mathrm{ord}_g(N)$ divides n.*
2. *$a = f/g$ for some $f, g \in \mathbb{Z}$ such that $g > 0$ and $g|(N^n - 1)$.*
3. *$a = h/(N^n - 1)$ for some $h \in \mathbb{Z}$.*
4. *$\mathbf{seq}_N(a)$ is eventually periodic and n is a period of a.*

The eventual period is the least n for which (1), (2) or (3) holds. The N-adic integer a is purely periodic if and only if $-(N^n - 1) \le h \le 0$ in case (3) or $-g \le f \le 0$ in cases (1) and (2).

Proof. Recall from Section 2.2.d that the integer g is relatively prime to N if and only if there exists $n \ge 0$ so that $g|(N^n - 1)$, and the smallest such is $n = \mathrm{ord}_g(N)$. (See also Lemma 2.4.6.) Hence (1) and (2) are equivalent. That (2) and (3) are equivalent is left to the reader.

To see that (4) implies (3) let us first consider the special case when a is strictly periodic with period n. Set $h = a_0 + a_1 N + \cdots + a_{n-1}N^{n-1}$. Then $0 \le h \le N^n - 1$ and

$$a = h(1 + N^n + N^{2n} + \cdots) = h/(1 - N^n) = -h/(N^n - 1) \tag{5.10}$$

as claimed. (Notice that no carries occur in the above product.) If a is eventually periodic, suppose it becomes periodic after the mth term. Then we can write $a = H + N^m b$ for some integer $H \geq 0$, where $b \in \mathbb{Z}_N$ is strictly periodic. Applying the special case to b and taking a common denominator gives $a = h'/(N^n - 1)$ for some $h' \in \mathbb{Z}$.

To see that (3) implies (4), let us first consider the special case when $1 - N^n \leq h \leq 0$. Then $0 \leq -h \leq N^n - 1$ so $-h$ can be uniquely expressed as a sum, $-h = a_0 + a_1 N + \cdots + a_{n-1} N^{n-1}$ with $0 \leq a_i < N$. Consequently equation (5.10) holds and since there are no carries in the product that occurs there, the sequence $\mathbf{seq}_N(a)$ is strictly periodic.

Now we show how to reduce to the case that $1 - N^n < h \leq 0$. If $h > 0$ then $-h < 0$ and according to equation (5.9), multiplication by -1 does not affect the eventual periodicity (nor the eventual period) of an N-adic number. So we may assume $h \leq 0$. If $h \leq 1 - N^n < 0$ then we can write $a = H + h'/(N^n - 1)$ where $H \in \mathbb{Z}$, $H < 0$, and $1 - N^n < h' \leq 0$. Therefore $-a = (-H) + (-h')/(N^n - 1)$ is eventually periodic because the addition of the positive integer $-H$ to the eventually periodic expansion of $(-h')/(N^n - 1)$ still leaves an eventually periodic series. Using equation (5.9) again, it follows that the N-adic expansion of a is also eventually periodic.

It follows immediately that the eventual period is the least n for which (1), (2) or (3) holds. □

Corollary 5.4.5. *Let $f, g, \in \mathbb{Z}$ with $\gcd(g, N) = 1$. If $\gcd(f, g) = 1$, then the period of the N-adic expansion $\mathbf{seq}_N(f/g)$ is the multiplicative order of N modulo g.*

5.4.d $\mathbb{Z}_N$ as an inverse limit

Let $\psi_\ell : \mathbb{Z}_N \to \mathbb{Z}/(N^\ell)$ be the homomorphism that associates to each $a = \sum_{i=0}^{\infty} a_i N^i$ the partial sum

$$\psi_\ell(a) = \sum_{i=0}^{\ell-1} a_i N^i.$$

These homomorphisms are compatible in the sense that if $k \leq \ell$ then

$$\psi_{\ell,k}(\psi_\ell(a)) = \psi_k(a)$$

where

$$\psi_{\ell,k} : \mathbb{Z}/(N^\ell) \to \mathbb{Z}/(N^k)$$

is reduction modulo N^k. In the language of Section 2.2.l, the family of rings $\mathbb{Z}/(N^\ell)$ is a directed system indexed by the positive integers with the maps $\psi_{\ell,k}$. The next lemma says that every N-adic integer can be described as such a sequence of partial sums. It is an exact parallel of Lemma 5.2.7.

Lemma 5.4.6. *For all $N > 1$, the mappings $\psi_{\ell,k}$ induce an isomorphism of rings,*

$$\mathbb{Z}_N \cong \varprojlim \{\mathbb{Z}/(N^i)\}.$$

In other words, there is a one to one correspondence between $\mathbb{Z}_N$ and the set of all sequences $(s_0, s_1, \cdots)$ with $s_i \in \mathbb{Z}/(N^i)$ such that for all pairs $i \leq j$ we have $\psi_{j,i}(s_j) = s_i$.

Proof. By Theorem 2.2.33 there is a unique induced map $\psi : \mathbb{Z}_N \to \varprojlim \{\mathbb{Z}/(N^i)\}$. It suffices to construct an inverse for this function. Suppose $s = (s_1, s_2, \cdots) \in \varprojlim \{\mathbb{Z}/(N^i)\}$. That is, $s_i \in \mathbb{Z}/(N^i)$ and for all pairs $i \leq j$ we have $\psi_{j,i}(s_j) = s_i$. Let a_i be the coefficient of N^i in the N-adic expansion of s_{i+1}. By the commutativity assumptions, this is also the coefficient of N^i in s_j for all $j > i$. Define $\tau(s) = \sum_{i=0}^{\infty} a_i N^i \in \mathbb{Z}_N$. Then $\psi(\tau(s)) = s$, so τ is the desired inverse. This is illustrated in Figure 5.2. □

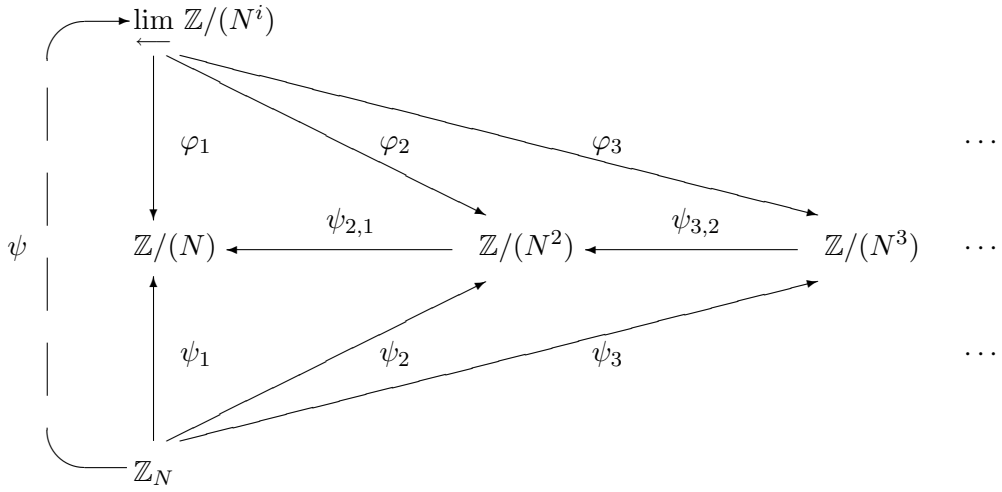


Figure 5.2: $\mathbb{Z}_N$ as an inverse limit

Corollary 5.4.7. *Let M be an Abelian group. For $i = 1, 2, \cdots$ let $\tau_i : M \to \mathbb{Z}/(N^i)$ be group homomorphisms satisfying $\tau_i = \psi_{j,i} \circ \tau_j$ whenever $i \leq j$. Then there is a unique homomorphism $\tau : M \to \mathbb{Z}_N$ so that $\tau_i = \psi_i \circ \tau$ for $i = 1, 2, \cdots$. If M is also a ring and the τ_i are homomorphisms, then so is τ.*

5.4.e Structure of $\mathbb{Z}_N$

In this section we suppose the prime factorization of N is $N = p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$, with distinct primes p_i. The ring $\mathbb{Z}_N$ can be expressed in terms of the p-adic integers $\mathbb{Z}_{p_i}$.

Theorem 5.4.8. *With N as above, the ring $\mathbb{Z}_N$ is isomorphic to the ring $\mathbb{Z}_{p_1} \times \mathbb{Z}_{p_2} \times \cdots \times \mathbb{Z}_{p_k}$. Similarly, $\mathbb{Q}_N \cong \mathbb{Q}_{p_1} \times \cdots \times \mathbb{Q}_{p_k}$.*

Proof. To simplify the notation slightly set $q_i = p_i^{n_i}$. For each j and ℓ, we have a homomorphism from $\mathbb{Z}/(N^\ell)$ to $\mathbb{Z}/(q_i^\ell)$. This induces a homomorphism from $\mathbb{Z}_N$ to $\mathbb{Z}/(q_i^\ell)$, and all the appropriate functions commute. Thus by the universal property of inverse limits, there are homomorphisms

$$\gamma_i : \mathbb{Z}_N \to \mathbb{Z}_{q_i}$$

with appropriate commutativity. This gives us a homomorphism

$$\gamma : \mathbb{Z}_N \to \prod_{i=1}^{k} \mathbb{Z}_{q_i},$$

which we now show to be an isomorphism by constructing an inverse. For every positive ℓ there is a reduction homomorphism

$$\delta_\ell : \prod_{i=1}^{k} \mathbb{Z}_{q_i} \to \prod_{i=1}^{k} \mathbb{Z}/(q_i^\ell).$$

By the Chinese Remainder Theorem (Theorem 2.2.18), the latter ring is isomorphic to $\mathbb{Z}/(N^\ell)$. Everything commutes appropriately, so there is an induced map

$$\delta : \prod_{i=1}^{k} \mathbb{Z}_{q_i} \to \mathbb{Z}_N.$$

It is straightforward to see that γ and δ are inverses (it follows, for example, from the uniqueness of the induced map into the universal object $\mathbb{Z}_N$).

This reduces the theorem to the case where $N = p^n$ for some prime p. Let

$$a = \sum_{i=0}^{\infty} a_i p^{ni} \in \mathbb{Z}_{p^n}, \tag{5.11}$$

with $0 \le a_i < p^n$. Each coefficient can be uniquely expressed as $a_i = \sum_{j=0}^{n-1} a_{i,j} p^j$ with $0 \le a_{i,j} < p$. Substituting this into (5.11) gives a p-adic integer. It is straightforward to verify that the resulting mapping $\mathbb{Z}_{p^n} \to \mathbb{Z}_p$ is a ring isomorphism.

The ring $\mathbb{Q}_N$ is obtained from $\mathbb{Z}_N$ by inverting N, which is equivalent to inverting $p_1, p_2, \cdots, p_k$ simultaneously. It follows that $\mathbb{Q}_N \cong \prod_{i=1}^{k} \mathbb{Q}_{p_i}$. □

Corollary 5.4.9. *The ring $\mathbb{Q}_N = S^{-1}\mathbb{Z}_N$ is obtained by inverting the set S consisting of all non-zero-divisors in $\mathbb{Z}_N$. The ring $\mathbb{Z}_N$ has no zero divisors if and only if N is prime.*

Proof. Use the isomorphism $\mathbb{Z}_N \cong \prod_{i=1}^{k} \mathbb{Z}_{p_i}$ of Theorem 5.4.8. Then an element $a = (a_1, a_2, \cdots, a_k)$ in this product is a zero divisor if and only if at least one of the coordinates $a_j = 0$ (since then, if b has a 1 in the jth position and zeroes elsewhere, we have $ab = 0$). So the set S of non-zero-divisors consists of all such k-tuples where all of the a_j are nonzero. Hence, $S = S_1 \times S_2 \times \cdots \times S_k$ is the product of the sets $S_i = \mathbb{Z}_{p_i}^{\times}$ of nonzero elements in $\mathbb{Z}_{p_i}$. So inverting this set gives $\prod_{i=1}^{k} \mathbb{Q}_{p_i} \cong \mathbb{Q}_N$. □

There are many irrational algebraic numbers in $\mathbb{Z}_N$. For example, suppose that $u(x)$ is a polynomial with integer coefficients that has a root modulo N. Then $u(x)$ has a root in $\mathbb{Z}_N$. This is proved in the next section using Hensel's Lemma.

5.5 π-Adic numbers

In this section we put the constructions from Subsections 5.2 and 5.4 into a larger context that enables us to build very general algebraic sequence generators. Let R be an integral domain. Let $\pi \in R$.

In the case of power series, we took coefficients from the underlying ring. In the case of N-adic integers we took coefficients from $\{0, 1, \cdots, N-1\}$. When we construct π-adic numbers, the generalizations of power series and N-adic integers, there may be no such natural set to use for coefficients so we take a slightly different approach.

5.5.a Construction of R_π

Definition 5.5.1. *A* pre-π-adic number over R *is an infinite expression*

$$a = a_0 + a_1\pi + a_2\pi^2 + \cdots,$$

with $a_0, a_1, \cdots \in R$. Let $\widehat{R}_\pi$ denote the set of pre-π-adic numbers.

The a_i are the *coefficients*, and the sequence $(a_0, a_1, \cdots)$ is referred to as $\mathbf{seq}(a)$ or $\mathbf{seq}_\pi(a)$. When writing pre-π-adic numbers we may omit terms whose coefficient is zero. We may also write the terms in a different order. The coefficients are arbitrary and may even be multiples of π. In fact a pre-π-adic number is just a power series over R, so $\widehat{R}_\pi$ is a commutative ring.

We want to think of certain pre-π-adic numbers as representing the same element. For example, $\pi \cdot 1 + 0 \cdot \pi + 0 \cdot \pi^2 \cdots$ and $0 \cdot 1 + 1 \cdot \pi + 0 \cdot \pi^2 \cdots$ should be equal. We accomplish this by taking a

quotient by an appropriate ideal. For each positive integer n we have a function $\widehat{\varphi}_n : \widehat{R}_\pi \to R/(\pi^n)$ defined by discarding terms of degree $\geq n$ and mapping the resulting element of R to $R/(\pi^n)$,

$$\widehat{\varphi}_n(\sum_{i=0}^{\infty} a_i\pi^i) = \sum_{i=0}^{n-1} a_i\pi^i \pmod{\pi^n}.$$

Let $I = \bigcap_{n=1}^{\infty} \mathrm{Ker}(\widehat{\varphi}_n)$. (This ideal contains many nonzero elements; see Exercise 9.)

Definition 5.5.2. *The ring of* π-adic integers over R *is the quotient ring* $R_\pi = \widehat{R}_\pi/I$.

If the context is clear we may simply refer to a π-adic integer. The homomorphism $h : R \to R_\pi$ (given by $a \mapsto a\pi^0 + 0 + 0\cdots$) is injective if and only if

$$\bigcap_{i=0}^{\infty}(\pi^i) = (0). \tag{5.12}$$

because its kernel is the set of $a \in R$ such that $\pi^n | a$ for all n. In studying sequences we often focus on rings that satisfy equation (5.12) since we can replace R by $R/\cap_{i=0}^{\infty}(\pi^i)$ without changing R_π.

The element π generates an ideal in R_π, and the homomorphism $h : R \to R_\pi$ induces an isomorphism $R/(\pi^n) \cong R_\pi/(\pi^n)$ for all n. (First check that h induces an injection from $R/(\pi^n)$ to $R_\pi/(\pi^n)$. But any $a = \sum_{i=0}^{\infty} a_i\pi^i \in R_\pi/(\pi^n)$ is the image of $\sum_{i=0}^{n-1} a_i\pi^i \in R/(\pi^n)$, so it is also a surjection.)

A more convenient way of representing π-adic integers is the following. By a *complete set of representatives for* R *modulo* π we mean a set S such that for all $a \in R$ there is a unique $b \in S$ so that $a \equiv b \pmod{\pi}$. The set S is not necessarily closed under addition or multiplication, however it often happens that additively or multiplicatively closed sets S can be found.

Theorem 5.5.3. *Let* R *be an integral domain, let* $\pi \in R$ *and let* S *be a complete set of representatives for* R *modulo* π. *Then every* π*-adic integer* $a \in R_\pi$ *has a unique* π*-adic expansion* $a = \sum_{i=0}^{\infty} b_i\pi^i$ *with all* $b_i \in S$.

Proof. Let $a = \sum_{i=0}^{\infty} a_i\pi^i \in R_\pi$. We need to construct a sequence $b_0, b_1, \cdots \in S$ so that for all n

$$\pi^n | \sum_{i=0}^{n-1}(a_i - b_i)\pi^i, \tag{5.13}$$

for then equation (5.12) will imply that $a = \sum b_i\pi^i$. Let $b_0 \in S$ be the unique element so that $a_0 \equiv b_0 \pmod{\pi}$. Inductively assume that we have found $b_0, \cdots, b_{n-1}$ so that equation (5.13) holds. Then

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

for some $c \in R$. Let b_n be the unique element of S such that π divides $a_n - b_n + c$. There is a $d \in R$ such that $a_n + c = b_n + \pi d$. Then

$$\begin{aligned}\sum_{i=0}^{n}(a_i - b_i)\pi^i &= (a_n - b_n)\pi^n + \sum_{i=0}^{n-1}(a_i - b_i)\pi^i \\ &= (a_n - b_n)\pi^n + c\pi^n \\ &= d\pi^{n+1}.\end{aligned}$$

This proves the existence part of the theorem.

Suppose $c_0, c_1, \cdots \in S$ is a second set of coefficients such that

$$\pi^n \left| \sum_{i=0}^{n-1}(a_i - c_i)\pi^i \right.$$

for all n. Then also

$$\pi^n \left| \sum_{i=0}^{n-1}(b_i - c_i)\pi^i \right.$$

for all n. Then $\pi|(b_0 - c_0)$ which implies $b_0 = c_0$. Inductively suppose that $b_i = c_i$ for $i < n$. Then

$$\pi^{n+1} \,|(b_n - c_n)\pi^n.$$

But R is an integral domain, so $\pi|(b_n - c_n)$, so $b_n = c_n$. □

For example, the power series ring $A[[x]]$ over a ring A is the ring R_π where $R = A[x]$ and $\pi = x$, so it is the ring of x-adic integers over $A[x]$, in the terminology of this section. Also, the ring $\mathbb{Z}_N$ of N-adic integers (in the terminology of the preceding section) is the ring R_π where $R = \mathbb{Z}$ and $\pi = N$, so it is the ring of N-adic integers over $\mathbb{Z}$.

5.5.b Divisibility in R_π

Relative to a fixed complete set of representatives S for R modulo π, there is a well defined notion of the reduction of an element of R_π modulo π in R, and of the integral quotient of an element of R_π by π. If

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

is a π-adic integer with $a_0, a_1 \cdots \in S$, then we write $a_0 = a \pmod{\pi}$ and refer to it as the *reduction of a modulo π*. The *integral quotient* of a by π is

$$a \ (\mathrm{div}_S \ \pi) = \sum_{i=0}^{\infty} a_{i+1}\pi^i = (a - a_0)/\pi.$$

If $a \in R$, then $a\ (\mathrm{div}_S\ \pi) \in R$ also. If the set S is understood, we simply write $a\ (\mathrm{div}\ \pi)$. Thus in general $a = a\ \pmod{\pi} + \pi(a\ (\mathrm{div}\ \pi))$.

Recall from Section 2.2.e that $q, \pi \in R$ are *relatively prime* if the following equivalent conditions hold.

- $(q) + (\pi) = R$
- The image of q in $R/(\pi)$ is invertible.
- The image of π in $R/(q)$ is invertible.

Proposition 5.5.4. *Let R be an integral domain with fraction field F. Let $\pi \in R$ and assume equation (5.12) holds. Then an element $q \in R$ is invertible in R_π if and only if q and π are coprime. Thus $R_\pi \cap F$ consists of all fractions u/q such that q, π are coprime.*

Proof. If q is invertible in R_π then there exists $u \in R_\pi$ so that $qu = 1$. Reducing this equation modulo π implies that q is invertible in $R_\pi/(\pi) \cong R/(\pi)$ so q and π are coprime. To verify the converse, let $S \subset R$ be a complete set of representatives for $R/(\pi)$ and let $q = \sum_{i=0}^{\infty} q_i \pi^i$ with $q_i \in S$. By hypothesis, q_0 is invertible in $R/(\pi)$. We seek $u = \sum_{i=0}^{\infty}$ with $u_i \in S$ such that $qu = 1$, which is to say that $q_0 u_0 \equiv 1\ \pmod{\pi}$ and, for all $n \geq 1$,

$$q_0 u_n + q_1 u_{n-1} + \cdots + q_n u_0 \equiv 0\ \pmod{\pi}.$$

These equations may be solved recursively for u_n, using the fact that q_0 is invertible in $R/(\pi)$. □

5.5.c The example of $\pi^d = N$

Fix integers $N, d > 0$ such that the polynomial $x^d - N$ is irreducible over the rational numbers $\mathbb{Q}$. This occurs precisely when (1) for any prime number k dividing d, the integer N is not a kth power of an integer, and (2) if 4 divides d, then N is not of the form $-4x^2$ where x is an integer; see [119, p. 221]. Let $\pi \in \mathbb{C}$ be a fixed root of this polynomial; it can be chosen to be a positive real number. The ring $R = \mathbb{Z}[\pi]$ consists of all polynomials in π, with integer coefficients. It is an integral domain in which every prime ideal is maximal.

We claim that the set $S = \{0, 1, \cdots, N-1\} \subset R = \mathbb{Z}[\pi]$ is a complete set of representatives for the quotient $R/(\pi)$. The mapping $\mathbb{Z}[\pi] \to \mathbb{Z}[\pi]/(\pi)$ throws away all the terms of degree ≥ 1 in any polynomial $u \in \mathbb{Z}[\pi]$. Consequently the composition $\mathbb{Z} \to R = \mathbb{Z}[\pi] \to R/(\pi)$ is surjective. So it suffices to show that $R/(\pi)$ contains N elements. In fact the ring $\mathbb{Z}[\pi]$ is an order (but not necessarily the maximal order) in its fraction field $F = \mathbb{Q}(\pi)$ so Lemma 3.4.8 gives:

$$|R/(\pi)| = \left|\mathsf{N}_{\mathbb{Q}}^{F}(\pi)\right|,$$

which we now compute.

The field F is a degree d extension of the rational numbers $\mathbb{Q}$ and it is the smallest field extension of $\mathbb{Q}$ containing π. Having fixed $\pi \in \mathbb{C}$, we obtain an embedding $F \subset \mathbb{C}$. However it actually admits d different embeddings into the complex numbers, $\sigma_i : F \to \mathbb{C}$ which are determined by setting $\sigma_i(\pi) = \zeta^i \pi$ (for $0 \leq i \leq d-1$) where $\zeta \in \mathbb{C}$ is a primitive d-th root of unity. The norm $\mathsf{N}(u)$ of an element $u \in F$ is the product of the images of u under these embeddings. (These facts use the irreducibility of the polynomial $x^d - N$.) Hence, $\mathsf{N}(\pi) = \pi^d \zeta^{d(d-1)/2} = \pm N$, which proves the claim. (We remark in passing that $1/\pi = \pi^{d-1}/N$ so the field $F = \mathbb{Q}(\pi) = \mathbb{Q}[\pi]$ consists of polynomials in π with rational coefficients.)

Having found a complete set of representatives for $R/(\pi)$ we can now describe the completion R_π. Each $a \in R_\pi$ can be uniquely represented as a power series $a = a_0 + a_1\pi + \cdots$ with coefficients $a_i \in S = \{0, 1, 2, \cdots, N-1\}$, however we must remember that $N = \pi^d$. Consequently, addition of π-adic integers may be described as termwise addition with a "delayed carry": each carried quantity is delayed d steps before adding it back in. In other words, if $b = b_0 + b_1\pi + \cdots$ then

$$a + b = \sum_{i=0}^{\infty} e_i \pi^i,$$

with $0 \leq e_i \leq N-1$, means that there exist $c_d, c_{d+1}, \cdots \in \{0, 1\}$ with

$$a_i + b_i + c_i = e_i + N c_{i+d}.$$

That is, c_i is the carry to the ith position. Similarly the difference $a - b = \sum_{i=0}^{\infty} f_i \pi^i$ is obtained by subtracting the coefficients symbol by symbol, using a "borrow" operation which is delayed d steps. The "borrow" operation is actually the same as the "carry" operation, but the carried quantity is negative. That is,

$$a_i - b_i + c_i = e_i + N c_{i+d}$$

from which it also follows immediately that the amount c_i to be carried to the ith place is either 0 or -1. In this case it is possible to improve on Proposition 5.5.4.

Proposition 5.5.5. *As a subset of F, the intersection $R_\pi \cap F$ consists of all elements u/q such that $u, q \in F$ and q, π are coprime. As a subset of R_π the intersection $R_\pi \cap F$ consists of all elements $a = a_0 + a_1\pi + \cdots$ whose coefficient sequence $\mathbf{seq}_\pi(a) = a_0, a_1, \cdots$ is eventually periodic.*

Proof. The first statement is Proposition 5.5.4. If $a \in R_\pi$ is an element whose coefficient sequence is eventually periodic with period m, then using the geometric series, it follows that $a = h/(1-\pi^m) \in F$ (for some $h \in R$). On the other hand, suppose that $u/q \in F$ (and q is relatively prime to π). The ring $\mathbb{Z}[\pi]/(q)$ is finite by Lemma 3.4.8. Therefore the elements $\{1, N, N^2, \cdots\}$ are not all distinct (mod q) which implies that $N^r \equiv 1 \pmod{q}$ for some $r \geq 1$. Hence there exists $a \in \mathbb{Z}[\pi]$ such that $aq = 1 - N^r$ so $u/q = ua/(1 - N^r)$. Set $ua = v_0 + v_1\pi + \cdots + v_{d-1}\pi^{d-1}$ with $v_i \in \mathbb{Z}$.

Each $v_i/(1-N^r) \in \mathbb{Z}_N$ is an N-adic integer whose coefficient sequence is eventually periodic, of period (a divisor of) r. These series exactly interleave in the sum

$$\frac{u}{q} = \sum_{i=0}^{d-1} v_i(1 + \pi^{dr} + \pi^{2dr} + \cdots)\pi^i \in R_\pi$$

giving a π-adic number whose coefficient sequence is eventually periodic of period rd. □

Even if the coefficient sequences of $a, b \in R_\pi$ are strictly periodic of the same period, say T, the same is not necessarily true for the coefficient sequences of $a \pm b$. Since the carries are delayed for d steps, the periodic part of $a \pm b$ might begin only after d symbols have passed.

Lemma 5.5.6. *Let*

$$a = \sum_{i=0}^{\infty} a_i\pi^i \quad \text{and} \quad b = \sum_{i=0}^{\infty} b_i\pi^i$$

be π-adic integers whose coefficient sequences are eventually periodic with period (a divisor of) n. Then the coefficient sequence of $a \pm b$ is eventually periodic with period (a divisor of) n.

Proof. (We consider the case of $a - b$; the case of the sum is similar.) It suffices to show that the sequence of carries $c_0, c_1, \ldots$ is eventually periodic with period dividing n. Suppose a carry occurs in the ith place, i.e. $c_i = -1$. This occurs if and only if for some positive $k \leq i/d$ we have $a_{i-jd} = b_{i-jd}$ for $1 \leq j \leq k-1$, $a_{i-kd} = 0$, and $b_{i-kd} = 1$. But if this occurs then the same is true with i replaced by $i + rn$ for every positive integer r.

Thus there are two possibilities for any i. Either for all r there is no carry to position $i + rn$, or for r large enough there is a carry to position $i + sn$ for every $s \geq r$. Therefore the sequence of carries has eventual period dividing n, and the lemma follows. □

5.6 Other constructions

In this section we describe other ways to define the π-adic integers over an integral domain R. We include this material for completeness but the results in this section will not be used in the sequel.

5.6.a R_π as an inverse limit

The collection of rings $\{R^i = R/(\pi^i) : 1 \leq i < \infty\}$ forms a directed system with (the reduction moduli π^i) homomorphisms $\psi_{j,i} : R^j \to R^i$ for $i \leq j$. So the limit $\varprojlim \{R^i\}$ exists (see Section 2.2.l), and there are projections $\varphi_i : \varprojlim \{R^i\} \to R^i$ such that $\phi_i = \psi_{j,i} \circ \phi_j$ whenever $i \leq j$. Similarly, the ring R_π comes with (reduction modulo π^i) homomorphisms $\psi_i : R_\pi \to R^i$ such that $\psi_i = \psi_{j,i} \circ \psi_j$.

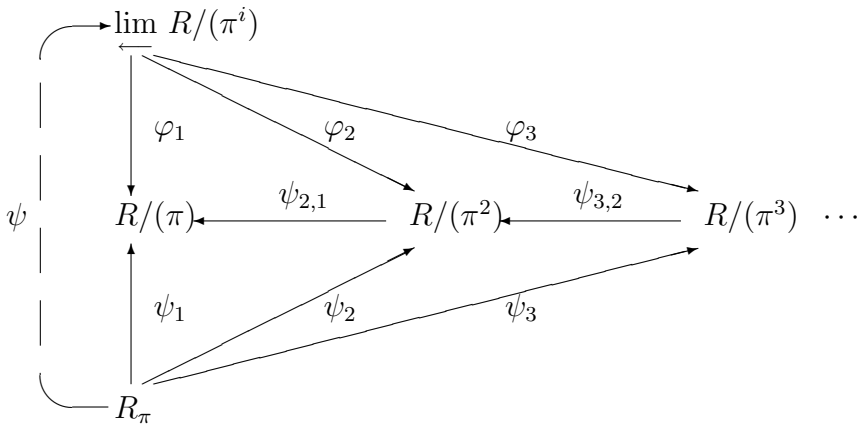


Figure 5.3: R_π as an inverse limit

Consequently there exists a homomorphism $\psi : R_\pi \to \varprojlim \{R_i\}$ such that $\psi_i = \varphi_i \circ \psi$, see Figure 5.3.

Proposition 5.6.1. *The function $\psi : R_\pi \to \varprojlim \{R/(\pi^i)\}$ is an isomorphism of rings.*

Proof. If $a = \sum_{i=0}^{\infty} a_i \pi^i \in R_\pi$ is nonzero, then π^n does not divide $\sum_{i=0}^{n-1} a_i \pi^i$ for some n. Thus $\psi_n(a) \neq 0$, and therefore $\psi(a) \neq 0$. This implies ψ is injective. Let $b = (b_0, b_1, \cdots) \in \varprojlim \{R/(\pi^i)\}$. For each i let $c_i \in R$ reduce to b_i modulo π^i. Thus $\pi^i | (c_i - c_{i-1})$. Let $a_i = (c_i - c_{i-1})/\pi^i$. Then $a = \sum_{i=0}^{\infty} a_i \pi^i$ reduces to b_i modulo π^{i+1} for every i. That is, $\psi(a) = b$ and ψ is a surjection and thus an isomorphism. □

Corollary 5.6.2. *Let M be an Abelian group. For $i = 1, 2, \cdots$ let $\tau_i : M \to R/(\pi^i)$ be group homomorphisms satisfying $\tau_i = \psi_{j,i} \circ \tau_j$ whenever $i \leq j$. Then there is a unique homomorphism $\tau : M \to R_\pi$ so that $\tau_i = \psi_i \circ \tau$ for $i = 1, 2, \cdots$. If M is also a module over R (respectively, a ring) and the τ_i are R-module homomorphisms (resp., ring homomorphisms), then so is τ.*

Corollary 5.6.3. *Let R and S be commutative rings with nonunits $\pi \in R$ and $\rho \in S$. Suppose that $\mu : R \to S$ is a ring homomorphism such that $\mu(\pi)$ is divisible by ρ. Then μ extends to a homomorphism $\mu : R_\pi \to S_\rho$.*

Proof. By hypothesis, for each i there is a series of homomorhisms

$$R_\pi \to R_\pi/(\pi^i) = R/(\pi^i) \to S/(\rho^i)$$

so that all the usual diagrams commute. Thus the universal property of S_ρ implies that μ extends to $\mu : R_\pi \to S_\rho$. □

5.6.b Valuations

In many cases the ring R_π may be described as a *completion* with respect to a *discrete valuation*. This important notion is central in much of modern number theory and algebraic geometry.

Definition 5.6.4. *Let A be a ring.* A valuation on A *(sometimes called a "discrete exponential valuation") is a function $\nu : A \to \mathbb{Z} \cup \{\infty\}$ such that for all $a, b \in A$*

1. $\nu(a+b) \geq \min(\nu(a), \nu(b))$.
2. $\nu(ab) = \nu(a) + \nu(b)$.
3. $\nu(a) = \infty$ *if and only if* $a = 0$.

It follows that $\nu(1) = 0$ so $\nu(a^{-1}) = -\nu(a)$ if $a \in A$ is invertible. The valuation is *nontrivial* if there exists a nonzero $a \in A$ such that $\nu(a) > 0$. Let (A, ν) be a ring with a nontrivial valuation. Then A is an integral domain (for if $ab = 0$ then $\infty = \nu(0) = \nu(ab) = \nu(a) + \nu(b)$ so at least one of $\nu(a), \nu(b)$ is ∞). If K is the field of fractions of A, then the valuation extends to a valuation on K by $\nu(a/b) = \nu(a) - \nu(b)$.

Conversely, if (F, ν) is a field with a (nontrivial discrete exponential) valuation ν then the following statements can be checked.

1. The set $F_{\geq 0} = \{a \in F : \nu(a) \geq 0\}$ is a ring with valuation, called the *valuation ring* of the field. For every $a \in F$, at least one of $a \in F_{\geq 0}$ or $a^{-1} \in F_{\geq 0}$.
2. If $S^{-1}F_{\geq 0}$ denotes the fraction field of $F_{\geq 0}$ then the mapping $h : S^{-1}F_{\geq 0} \to F$ given by $h(a/b) = ab^{-1}$ is an isomorphism of fields (with valuation).
3. There is exactly one maximal ideal in $F_{\geq 0}$ and it is the set $I = F_{>0} = \{a : \nu(a) > 0\}$. Consequently $F_{\geq 0}$ is a *local ring*, and an element $a \in F_{\geq 0}$ is invertible if and only if $\nu(a) = 0$. Moreover, $\cap_{i=0}^{\infty} I^n = \{0\}$, cf. equation (5.12).
4. The quotient $F_{\geq 0}/F_{>0}$ is a field, called the *residue field*,.

Examples:

1. Let K be a field and $F = K((x))$ its field of formal Laurent series. If

$$a(x) = a_m x^m + a_{m+1}x^{m+1} + a_{m+2}x^{m+2} + \cdots \in F$$

is a series with leading term $a_m \neq 0$, set $\nu(a(x)) = m$ (which can be positive, zero, or negative). Then ν is a discrete valuation on F and its valuation ring is the ring of formal power series $F[[x]]$, with maximal ideal (x). The residue field is K.

2. Let p be a prime integer. If $a \in \mathbb{Z}$ is an integer then we have $a = p^n b$ for some nonnegative integer n and some integer b that is relatively prime to p. Define $\nu_p(a) = n$. Then ν_p is a valuation on $\mathbb{Z}$. This valuation extends to the fraction field $\mathbb{Q}$ and its valuation ring is $\mathbb{Q}_{\geq 0} = \mathbb{Z}_{p,0}$, the set of fractions a/b (in lowest terms) such that b is not divisible by p, cf. Section 5.4.c. The maximal ideal in $\mathbb{Z}_{p,0}$ is (p) and the residue field is $\mathbb{F}_p = \mathbb{Z}/(p)$. We remark that this procedure does not give a valuation if p is replaced by a composite integer, say, $N = ab$ with $a, b \in \mathbb{Z}$. For then $\nu_N(a) = \nu_N(b) = 0$ (since N does not divide a or b), but $\nu(N) = 1$.

3. Let $p \in \mathbb{Z}$ be a prime integer and let $\mathbb{Q}_p$ be the field of p-adic numbers. If $a = a_m p^m + a_{m+1} p^{m+1} + \cdots$ with leading term $a_m \neq 0$, set $\nu_p(a) = m$. Then ν_p is a valuation on $\mathbb{Q}_p$; its restriction to $\mathbb{Q} \subset \mathbb{Q}_p$ agrees with the valuation ν_p described in item (2) above. The valuation ring is $\mathbb{Z}_p$, the p-adic integers, and the residue field is $\mathbb{F}_p = \mathbb{Z}/(p)$.

4. More generally, let R be a UFD with fraction field F, and let $\pi \in R$ be prime. If $a \in R$, then $a = \pi^n b$ for some nonnegative integer n and some $b \in R$ not divisible by π. If we define $\nu_\pi(a) = n$, then ν_π is a valuation on R. It extends to a valuation on F by $\nu(c/d) = \nu(c) - \nu(d)$. Similarly, it extends to valuations on R_π and F_π. Moreover, $R_\pi = (F_\pi)_{\geq 0}$.

5. Let (F, ν) be a discretely valued field with valuation ring $R_\nu = F_{\geq 0}$ and maximal ideal $I_\nu = F_{>0}$. Let $\pi \in I_\nu$ be an element whose valuation is minimal, say $\nu(\pi) = c$. Then π is prime in R_ν, and every element $y \in F$ is of the form $y = \pi^a x$ with $\nu(x) = 0$, and $\nu(y) = ac$. In other words, the construction in example (4) is completely general. To see this, first suppose that $\nu(y) = ac + d$ with $0 < d < c$. Then $\nu(y/\pi^a) = d$ which contradicts the minimality of c, hence $d = 0$. Therefore $b = y/\pi^a$ is a unit, and $y = b\pi^a$ as claimed. Similarly, if $\pi = uv$ is a nontrivial product (with $u, v \in R_\nu$, neither of which is a unit) then $\nu(\pi) = \nu(u) + \nu(v) \geq 2c$ which is false, so π is prime in R_ν.

5.6.c Completions

A *metric space* is a set X with a *metric* or "distance function" $\delta : X \times X \to \mathbb{R}$ such that $\delta(a, b) = \delta(b, a)$; $\delta(a, b) = 0$ if and only if $a = b$; and $\delta(a, b) \leq \delta(a, c) + \delta(c, b)$ (triangle inequality) for all $a, b, c \in X$. The *metric topology* on X is the topology generated by the open "balls" $B_\epsilon(x) = \{y \in X : \delta(x, y) < \epsilon\}$ for all $x \in X$ and all $\epsilon > 0$. The metric δ is continuous with respect to this topology. A mapping $f : (X, \delta_X) \to (Y, \delta_Y)$ between metric spaces is *isometric* if $\delta_Y(f(x_1), f(x_2)) = \delta_X(x_1, x_2)$ for all $x_1, x_2 \in X$. Such a mapping is continuous with respect to the metric topologies on X and Y. An *isometry* $f : (X, \delta_X) \to (Y, \delta_Y)$ is an isometric mapping that has an isometric inverse. (In particular, it is one to one and onto.)

A sequence of points $x_1, x_2, \cdots$ in a metric space X is a *Cauchy sequence* if for every $\epsilon > 0$ there exists a k so that $\delta(x_i, x_j) < \epsilon$ if $i, j \geq k$. A metric space is *complete* if every Cauchy

sequence converges. A *completion* $(\widehat{X}, \widehat{\delta})$ of a metric space (X, δ) is a complete metric space that contains X as a dense subset, such that the restriction of $\widehat{\delta}$ to X equals δ. Every metric space has a completion (constructed below). If $(\widehat{X}_1, \widehat{\delta}_1)$ and $(\widehat{X}_2, \widehat{\delta}_2)$ are two completions of a metric space (X, δ) then the identity mapping $X \to X$ extends in a unique way, to a continuous mapping $\widehat{f} : (\widehat{X}_1, \widehat{\delta}_1) \to (\widehat{X}_2, \widehat{\delta}_2)$ and moreover, the mapping $\widehat{f}$ is an isometry.

Thus, the completion of (X, δ) is unique up to isometry. It may be constructed as follows. The points in $\widehat{X}$ are equivalence classes of Cauchy sequences in X, two sequences $\mathbf{x} = x_1, x_2, \cdots$ and $\mathbf{y} = y_1, y_2, \cdots$ being equivalent if

$$\lim_{i \to \infty} \delta(x_i, y_i) = 0.$$

If $\mathbf{z} = z_1, z_2, \cdots$ is another point in $\widehat{X}$ then the extended metric is defined by $\widehat{\delta}(\mathbf{x}, \mathbf{z}) = \lim \delta(x_i, z_i)$ (which exists because $\mathbf{x}, \mathbf{z}$ are Cauchy sequences). The space X is contained in $\widehat{X}$ as the set of constant sequences.

Two metrics δ_1, δ_2 on a set X are *equivalent* if the set of Cauchy sequences for δ_1 coincides with the set of Cauchy sequences for δ_2. In this case, the identity mapping $I : X \to X$ has a unique continuous extension to the completions, $\widehat{I} : (\widehat{X}, \widehat{\delta}_1) \to (\widehat{X}, \widehat{\delta}_2)$ and $\widehat{I}$ is a homeomorphism.

Lemma 5.6.5. *Let ν be a discrete valuation on a field F and let $q > 1$ be a positive real number. Then $\delta(a, b) = q^{-\nu(a-b)}$ defines a metric on F. A different choice of $q > 1$ determines an equivalent metric. Moreover, for any Cauchy sequence $x_1, x_2, \cdots$ the limit $\lim \nu(x_i) \in \mathbb{Z} \cup \infty$ exists, and if the limit is not ∞ then it is attained after finitely many terms.*

Proof. If $\delta(a, b) = 0$ then $\nu(a - b) = \infty$ so $a = b$. If $a, b, c \in F$ then

$$\delta(a, b) = q^{-\nu(a-c+c-b)} \leq \max(q^{-\nu(a-c)}, q^{-\nu(c-b)}) \leq q^{-\nu(a-c)} + q^{-\nu(c-b)} = \delta(a, c) + \delta(c, b)$$

so δ is a metric. Now let $x_1, x_2, \cdots$ be a Cauchy sequence. This means that for any $T \geq 1$ there is a $k \geq 1$ such that

$$\nu(x_i - x_j) \geq T \quad \text{whenever} \quad i, j \geq k. \tag{5.14}$$

(So the particular choice of q does not matter.) There are now two possibilities. The first is that for infinitely many values of i, the values of $\nu(x_i)$ grow without bound. This in fact implies that $\nu(x_i) \to \infty$ (so $x_i \to 0$) because, for all i, j,

$$\nu(x_j) = \nu(x_i + (x_j - x_i)) \geq \min\{\nu(x_i), \nu(x_j - x_i)\}. \tag{5.15}$$

Since $\nu(x_i)$ can be chosen to be arbitrarily large, and since $\nu(x_i - x_j)$ grows without bound, it follows that $\nu(x_j)$ grows without bound.

The second possibility, therefore, is that the values of $\nu(x_i)$ remain bounded for all i. Consequently there exists $0 \leq M < \infty$ so that $\nu(x_j) \leq M$ for all j sufficiently large, and so that

$\nu(x_i) = M$ for infinitely many values of i. So there is an index i_0 with the property that $\nu(x_{i_0}) = M$ and if $i, j \geq i_0$ then $\nu(x_i - x_j) > M$. Hence, by equation (5.15), $\nu(x_j) = M$ whenever $j \geq i_0$, that is, the sequence $\nu(x_j)$ converges to M and it equals M after finitely many terms have passed. □

If the discretely valued field (F, ν) is complete with respect to the metric defined by ν, then we say that (F, ν) is a *complete discretely valued field* or *local field*.

Theorem 5.6.6. *Let (F, ν) be a field with a discrete valuation, with associated metric δ as in Lemma 5.6.5, and with valuation ring $R = F_{\geq 0}$. Then the valuation ν extends to a valuation $\widehat{\nu}$ on the completion $\widehat{F}$. Moreover, $\widehat{F}$ is again a field (so it is a local field) and its valuation ring $\widehat{F}_{\geq 0}$ naturally identifies with the completion $\widehat{R}$. In particular, $\widehat{F}$ is the fraction field of $\widehat{R}$. We say that $\widehat{F}$ is the* completion *of F with respect to ν.*

Proof. Let $\mathbf{x} = x_1, x_2, \cdots$ be a Cauchy sequence. By Lemma 5.6.5 the sequence $\nu(x_1), \nu(x_2), \cdots$ converges so we may define $\widehat{\nu}(\mathbf{x}) = \lim_{i\to\infty} \nu(x_i)$. If $\mathbf{y} = y_1, y_2, \cdots$ is an equivalent Cauchy sequence then $\widehat{\nu}(\mathbf{y}) = \widehat{\nu}(\mathbf{x})$ so $\widehat{\nu}$ is a well defined valuation on the completion $\widehat{F}$.

Let T be the set of all Cauchy sequences in F. Then T is a subring of the product of infinitely many copies of F, that is, addition and multiplication of two sequences is defined termwise. We need to check that these arithmetic operations are preserved by the equivalence relation on Cauchy sequences. The set of Cauchy sequences with limit 0 is an ideal I in T. Observe that two Cauchy sequences $\mathbf{x} = x_1, x_2, \cdots$ and $\mathbf{y} = y_1, y_2, \cdots$ are equivalent if and only if

$$0 = \lim_{i\to\infty} \delta(x_i, y_i) = q^{-\nu(x_i - y_i)}$$

which holds if and only if $\nu(x_i - y_i) \to \infty$, that is, $x_i - y_i \in I$. Thus, the completion $\widehat{F}$ is exactly T/I, which is a ring. To see that it is a field we need to show that every nonzero element has an inverse. Let $\mathbf{x} = x_1, x_2, \cdots$ be a Cauchy sequence that does not converge to 0. By Lemma 5.6.5 the sequence $\nu(x_i)$ converges to some number M and it equals M after some finite point. Therefore if i, j are sufficiently large,

$$\nu\left(\frac{1}{x_i} - \frac{1}{x_j}\right) = \nu\left(\frac{x_j - x_i}{x_i x_j}\right) = \nu(x_i - x_j) - \nu(x_i) - \nu(x_j) = \nu(x_i - x_j) - 2M \to \infty.$$

This shows that the sequence $x_1^{-1}, x_2^{-1} \cdots$ is a Cauchy sequence, and it therefore represents $\mathbf{x}^{-1} \in \widehat{F}$.

In a similar way we obtain the completion $\widehat{R} \subset \widehat{F}$ of the valuation ring $R = F_{\geq 0}$ and Lemma 5.6.5 implies that $\widehat{R} \subset \widehat{F}_{\geq 0}$. Conversely, if $\mathbf{x} \in \widehat{F}$ and if $\widehat{\nu}(\mathbf{x}) \geq 0$ then, again by Lemma 5.6.5, this implies that $\nu(x_i) \geq 0$ for all sufficiently large i, say, $i \geq i_0$. Therefore, if we replace $\mathbf{x}$ by the equivalent Cauchy sequence $\mathbf{x}' = x_{i_0}, x_{i_0+1}, \cdots$ then $\mathbf{x}' \in \widehat{R}$, which proves that $\widehat{F}_{\geq 0} = \widehat{R}$. □

Remarks. Completeness is not a "purely topological" invariant: it depends on a choice of metric as well. The real numbers $\mathbb{R}$ is complete (with the usual metric) but the open interval $(0,1)$, which is homeomorphic to $\mathbb{R}$, is not complete. Its completion is the closed interval. The set of real numbers is the completion of the rational numbers with respect to the Euclidean metric $\delta(x,y) = |x-y|$, however this metric does not arise from a *discrete* valuation. The following theorem says that the π-adic numbers as constructed in Section 5.5 is an example of a completion.

Theorem 5.6.7. *Let R be a UFD with field of fractions F. Let $\pi \in R$ be prime and let ν_π be the corresponding discrete valuation as in example 4 of Section 5.6.b. Then (R_π, ν_π) is isomorphic to its completion $(\widehat{R}, \widehat{\nu}_\pi)$, and F_π is isomorphic to its completion $(\widehat{F}, \widehat{\nu}_\pi)$.*

Proof. Let $\mathbf{x} = x_1, x_2, \cdots \in \widehat{R}$ be a Cauchy sequence. For each fixed n the sequence of reductions $x_i \pmod{\pi^n} \in R/(\pi^n)$ eventually stabilizes, giving a collection of compatible homomorphism $\widehat{R} \to R/(\pi^n)$. By Corollary 5.6.2 this gives a homomorphism $\widehat{R} \to R_\pi$. The inverse homomorphism associates to each power series $\sum_{i=0}^{\infty} a_i \pi^i$ its sequence of partial sums $a_0, a_0 + a_1\pi, \cdots$ which is a Cauchy sequence. Thus ψ is an isomorphism of complete valued rings, so its canonical extension $\psi : \widehat{F} \to F_\pi$ is an isomorphism of the corresponding fraction fields. □

A basic property of local fields is expressed in *Hensel's Lemma* which allows us to factor a polynomial over the residue field, and to lift the factorization to the local field. Let F be a complete discretely valued field with discrete valuation ν, valuation ring $R = F_{\geq 0}$, maximal ideal $I = F_{>0}$ and residue field $K = F_{\geq 0}/F_{>0}$ (all depending on ν) as in Section 5.6.b. If $f(x)$ is a polynomial over R, we denote by $\bar{f}(x)$ the reduction of $f(x)$ modulo the ideal I. The proof of the following may be found, for example, in [85, pp. 573-4].

Theorem 5.6.8. *(Hensel's Lemma) Suppose $f(x) \in R[x]$ is a monic polynomial and $\bar{f}(x) = g_0(x)h_0(x)$ in $K[x]$, where $g_0(x)$ and $h_0(x)$ are monic and relatively prime. Then there exist monic polynomials $g(x)$ and $h(x)$ in $R[x]$ such that $f(x) = g(x)h(x)$, $\bar{g}(x) = g_0(x)$, and $\bar{h}(x) = h_0(x)$.*

Corollary 5.6.9. *With the same hypotheses, if $\bar{f}(x)$ has a simple root a_0, then $f(x)$ has a simple root a such that $a \pmod{I} = a_0$.*

5.6.d Adic topology

The construction of R_π using valuations only works when R is a UFD and π is prime. But a similar construction works for more general ideals in more general rings. Let R be an integral domain and let $I \subset R$ be an ideal. Suppose that R is *separable* with respect to I, that is, $\cap_{n=1}^{\infty} I^n = \{0\}$. If $x \in I$ define $V(x) = \sup\{n : \ x \in I^n\} \in \mathbb{Z} \cup \infty$. If $x, y \in I$ then $V(x+y) \geq \min\{V(x), V(y)\}$ and $V(xy) \geq V(x) + V(y)$ (compare with Section 5.6.4). Fix $q > 1$ and define

$$\delta(x,y) = \begin{cases} q^{-V(x-y)} & \text{if } \ x - y \in I \\ \infty & \text{otherwise.} \end{cases}$$

Then V is almost a valuation, although it is not defined on all of R. However, the same method as in Lemma 5.6.5 and Theorem 5.6.6 shows that δ is a metric on R. It determines a topology on R, a basis of which is given by the open sets of the form $B_n(x) = x + I^n$ for $x \in R$ and $n \geq 1$. If $I = (\pi)$ is principal we refer to δ as a π-adic metric, and to the resulting topology as the π-adic topology.

Theorem 5.6.10. *Let R be an integral domain and let $\pi \in R$. Suppose R is separable with respect to the ideal (π). Then the ring R_π of π-adic integers may be naturally identified with the completion of R in the π-adic metric.*

Proof. The proof is the same as that of Theorem 5.6.7. □

5.7 Continued fractions

Continued fraction expansion provides an alternate way to represent certain algebraic objects. Every real number x has a continued fraction expansion. The continued fraction expansion of a rational number a/b is equivalent to Euclid's algorithm (300 BC) for (a, b). Specific examples of continued fractions were known to Bombelli and Cataldi around 1600. The first systematic treatment of continued fractions was by John Wallis in Opera Mathematica (1695). The subject was intensively studied in the nineteenth century. Like the Euclidean algorithm, the continued fraction expansion is optimal in two ways: (a) the successive terms, or "convergents" in this expansion give best-possible rational approximations to x, see Theorem 5.7.4; and (b) the terms in the expansion can be computed with very little effort. There are many wonderful applications of continued fractions to problems in mathematics, science, and engineering. For example, in [16] a constant is estimated, using a hand calculator, to be 2.1176470588. The CF expansion for this number is $[2, 8, 2, 147058823]$, suggesting that the actual number is $[2, 8, 2] = 36/17$, which turns out to be correct. In [48], continued fractions are used to describe the efficacy of the twelve-tone equal tempered musical scale, with the next best equal tempered scale having 19 tones.

Standard references for continued fractions include [75], [90] and [153]. We shall not explore this topic in detail, but we develop enough of the theory to understand the relation between continued fractions and the Berlekamp-Massey algorithm for linear feedback shift register synthesis (see Section 18.2.d).

5.7.a Continued fractions for rational numbers

The continued fraction representation for a rational number a/b (with a, b positive integers) is defined by the iterative procedure in Figure 5.4. Let $a_0 = a, a_1, a_2, \cdots$ and $b_0 = b, b_1, b_2, \cdots$ be the

RatContFrac(a, b)
begin
$n = 0$
while $b \neq 0$ **do**
Let $\frac{a}{b} = c_n + \frac{a'}{b}$ with $c_n, a' \in \mathbb{Z}$ and $0 \leq a' < b$
$a = b$
$b = a'$
$n = n + 1$
od
return $\langle c_0, c_1, \cdots, c_{n-1} \rangle$
end

Figure 5.4: Rational Continued Fraction Expansion.

sequences of as and bs generated by the algorithm, then we have successively

$$\frac{a}{b} = c_0 + \frac{a'}{b} = c_0 + \frac{b_1}{a_1} = c_0 + \frac{1}{\frac{a_1}{b_1}} = c_0 + \frac{1}{c_1 + \frac{b_2}{a_2}} = c_0 + \frac{1}{c_1 + \frac{1}{c_2 + \frac{a_3}{b_3}}} = \cdots$$

which we denote $[c_0, c_1, c_2 \cdots]$. If $a/b = 10/7$, this gives $\frac{10}{7} = 1 + \frac{3}{7} = 1 + \frac{1}{2 + \frac{1}{3}}$, so the continued fraction expansion of $10/7$ is $[1, 2, 3]$.

Proposition 5.7.1. *The procedure in algorithm* RatContFrac *halts after finitely many steps.*

Proof. We always have $b > a'$, so after the first iteration $a' < a$. Therefore, for every i we have $\max(a_{i+2}, b_{i+2}) < \max(a_i, b_i)$. It follows that eventually $b = 0$ and the algorithm halts. □

The sequence of non-negative integers $[c_0, c_1, \cdots, c_{n-1}]$ is called the *continued fraction expansion of* a/b. It is uniquely defined and gives an exact representation of a/b. Similarly, we can generate continued fraction expansions of real numbers. If $z > 0$ is real, let $\{z\}$ denote the fractional part of z and let $[z]$ denote the integer part or floor of z. Thus $z = [z] + \{z\}$. Then the continued fraction expansion of z is the sequence generated by the recursive definition

$$c_0 = [z], \quad r_0 = \{z\} = z - c_0$$

and for $n \geq 1$,

$$z_n = \frac{1}{r_{n-1}}, \quad c_n = [z_n], \quad r_n = \{z_n\}, \quad \text{so } z_n = c_n + r_n \tag{5.16}$$

where we continue only as long as $r_n \neq 0$. If z is irrational, then this recursion does not halt, but outputs an infinite sequence of integers $[c_0, c_1, \cdots]$ which is called the *continued fraction expansion of* z. For example, the continued fraction expansion of $z_0 = \sqrt{7}$ is:

$$\begin{aligned} z_0 &= 2 + (\sqrt{7} - 2) & z_1 &= \frac{1}{\sqrt{7}-2} = 1 + \frac{\sqrt{7}-1}{3} \\ z_2 &= \frac{3}{\sqrt{7}-1} = 1 + \frac{\sqrt{7}-1}{2} & z_3 &= \frac{2}{\sqrt{7}-1} = 1 + \frac{\sqrt{7}-2}{3} \\ z_4 &= \frac{3}{\sqrt{7}-2} = 4 + (\sqrt{7} - 2) & z_5 &= z_1. \end{aligned}$$

Thus the expansion repeats from here on, and the continued fraction is $[2, 1, 1, 1, 4, 1, 1, 1, 4, \cdots]$.

The *nth convergent* of the continued fraction $[c_0, c_1, \cdots]$ is the rational number f_n/q_n that is obtained from the finite continued fraction $[c_0, c_1, \cdots, c_n]$. The convergents f_n/q_n form a sequence of rational approximations to z. The following proposition, when combined with the algorithm of equation (5.16) or Figure 5.4 provides an efficient way to compute the convergents.

Proposition 5.7.2. *Let $z > 0$ be a real number with continued fraction expansion $z = [c_0, c_1, \cdots]$. Let r_n be the remainder as in equation (5.16). Then the convergents may be obtained from the following recursive rule: if $r_n \neq 0$ then*

$$f_{n+1} = c_{n+1} f_n + f_{n-1} \quad \text{and} \quad q_{n+1} = c_{n+1} q_n + q_{n-1}. \tag{5.17}$$

The initial conditions are: $f_0 = c_0$, $q_0 = 1$, $f_{-1} = 1$, and $q_{-1} = 0$. Moreover, for any n,

$$z = \frac{f_n + f_{n-1} r_n}{q_n + q_{n-1} r_n}. \tag{5.18}$$

Proof. At the nth stage of the recursion we have a representation

$$z = c_0 + (c_1 + (c_2 + \cdots + (c_{n-1} + (c_n + r_n)^{-1})^{-1} \cdots)^{-1})^{-1}.$$

The dependence on the innermost quantity, $(c_n + r_n)$, is fractional linear:

$$z = \frac{u_n(c_n + r_n) + w_n}{x_n(c_n + r_n) + y_n}, \tag{5.19}$$

where u_n, w_n, x_n, y_n are multilinear expressions in $c_0, \cdots, c_{n-1}$. Then f_n/q_n is obtained by setting $r_n = 0$ in equation (5.19), so

$$\frac{f_n}{q_n} = \frac{u_n c_n + w_n}{x_n c_n + y_n}.$$

That is, $f_n = u_n c_n + w_n$ and $q_n = x_n c_n + y_n$. Now consider equation (5.19) with n replaced by $n+1$. This gives the same result as equation (5.19) with $r_n \neq 0$ replaced by $(c_{n+1} + r_{n+1})^{-1}$. Thus

$$\begin{aligned}
\frac{u_{n+1}(c_{n+1} + r_{n+1}) + w_{n+1}}{x_{n+1}(c_{n+1} + r_{n+1}) + y_{n+1}} &= \frac{u_n(c_n + (c_{n+1} + r_{n+1})^{-1}) + w_n}{x_n(c_n + (ca_{n+1} + r_{n+1})^{-1}) + y_n} \\
&= \frac{u_n(c_n(c_{n+1} + r_{n+1}) + 1) + w_n(c_{n+1} + r_{n+1})}{x_n(c_n(c_{n+1} + r_{n+1}) + 1) + y_n(c_{n+1} + r_{n+1})} \\
&= \frac{(u_n c_n + w_n)(c_{n+1} + r_{n+1}) + u_n}{(x_n c_n + y_n)(c_{n+1} + r_{n+1}) + x_n}.
\end{aligned}$$

This being an equality of rational functions, we may conclude that $u_{n+1} = u_n c_n + w_n$ and that $w_{n+1} = u_n$. But $u_n c_n + w_n = f_n$ hence $u_{n+1} = f_n$ (and therefore $u_n = f_{n-1}$). Similarly $x_{n+1} = x_n c_n + y_n = q_n$, and $y_{n+1} = x_n = q_{n-1}$. Equation (5.17) follows immediately and equation (5.19) becomes (5.18). □

Lemma 5.7.3. *If $n \geq 0$ and $r_n \neq 0$ then $f_{n+1}q_n - q_{n+1}f_n = (-1)^{n+1}$ so f_n and q_n are relatively prime.*

Proof. The proof is by induction on n. The initial conditions give $f_0 q_{-1} - q_0 f_{-1} = 1$. If $n > 1$, then using equations (5.17) we have

$$f_n q_{n+1} - q_n f_{n+1} = f_n(c_{n+1}q_n + q_{n-1}) - q_n(c_{n+1}f_n + f_{n-1}) = -(f_{n-1}q_n - q_{n-1}f_n) = -(-1)^n. \quad \square$$

Theorem 5.7.4. *Let f_n/q_n denote the nth convergent ($n \geq 1$) of $z \in \mathbb{R}$ ($z > 0$). Then*

$$\left| z - \frac{f_n}{q_n} \right| < \frac{1}{q_n q_{n+1}}. \tag{5.20}$$

If $r_n \neq 0$, then $q_{n+1} > q_n$. If f, q are positive integers and if $|z - f/q| < 1/q_n q_{n+1}$, then $q > q_n$ unless $f/q = f_n/q_n$.

Proof. If $r_n = 0$, then $z = f_n/q_n$, so we may assume $r_n \neq 0$. By equation (5.18) we have

$$\begin{aligned}
z - \frac{f_n}{q_n} &= \frac{f_n + f_{n-1}r_n}{q_n + q_{n-1}r_n} - \frac{f_n}{q_n} \\
&= \frac{(f_{n-1}q_n - q_{n-1}f_n)r_n}{q_n(q_n + q_{n-1}r_n)} \\
&= \frac{(-1)^n r_n}{q_n(q_n + q_{n-1}r_n)} \\
&= \frac{(-1)^n}{q_n((1/r_n)q_n + q_{n-1})}.
\end{aligned}$$

For $i \geq 1$ we have $c_i \geq 0$. Thus by equation (5.17), $q_i \geq 0$ for all i. Therefore

$$\frac{1}{r_n} q_n + q_{n-1} \geq c_{n+1} q_n + q_{n-1} = q_{n+1}$$

which proves equation (5.20). If $0 \neq r_n < 1$ then $z_{n+1} > 1$ so $c_{n+1} \geq 1$. Since $q_{n-1} > 0$, equation (5.17) gives $q_n > q_{n-1}$. Although it is not difficult, the proof of the last statement is tedious, and may be found in [75], [90], or [153]. □

5.7.b Continued fractions for (reciprocal) Laurent Series

A theory of continued fractions can be developed whenever we have a subset R (the analog of the "integers") of a field F and a subset $U \subset F$ (of "fractions") so that every element of F can be uniquely written in the form $a + y$ with $a \in R$ and $y \in U$. Let $z \mapsto \{z\}$ be a function from F to U so that for every z we have $z - \{z\} \in R$. The sequences of elements $c_0, c_1, \cdots \in R$ and $r_0, r_1, \cdots \in U$ are defined exactly as in equation (5.16). This point of view is developed in [191], following work of [147] and [127]. In many cases (but not always: see the example in Section 5.7.c) the resulting "rational" approximations converge and they are often the "best" possible. In this section we describe the approach of [191].

Let K be a field. We wish to develop continued fraction expansions for *rational functions* $f(x)/g(x)$ where $f, g \in R = K[x]$ are polynomials. Unfortunately there is no apparent analog to the "integer part" of such a function, which we would like to be a polynomial in x. However a formal Laurent series $h(x) \in K((x))$ is the sum of two pieces: the (finitely many) terms with negative powers of x, plus the infinite series of terms with positive powers of x. We are thus led to consider continued fractions for the field

$$F = K((x^{-1})) = \left\{ \sum_{i=k}^{\infty} a_i x^{-i} : k \in \mathbb{Z}, a_i \in K \right\},$$

of "reciprocal Laurent series", or formal Laurent series in x^{-1}, because the "integer part" will now be a polynomial in x (with positive powers). As in Section 5.3, the field F contains all quotients of polynomials $f(x)/g(x)$. Thus we define

$$\left[\sum_{i=k}^{\infty} a_i x^{-i} \right] = \sum_{i \leq 0} a_i x^{-i} \in R = K[x]$$

and

$$\left\{ \sum_{i=k}^{\infty} a_i x^{-i} \right\} = \sum_{i \geq 1}^{\infty} a_i x^{-i} \in x^{-1} K[[x^{-1}]].$$

That is, the polynomial part of this Laurent series is the sum of the monomials with nonnegative exponents. The fractional part is the sum of the monomials with negative exponents. It is an element in the unique maximal ideal (x^{-1}) in $K[[x^{-1}]]$. With these definitions we can carry out continued fraction expansions just as for the real numbers in equation (5.16). That is,

$$z_n = \frac{1}{r_{n-1}}, \quad c_n = [z_n], \quad r_n = \{z_n\}, \quad \text{so } z_n = c_n + r_n$$

with $c_0 = [z]$ and $r_0 = \{z\}$. (See exercises 16 and 17.) The associated *convergent* f_n/q_n is obtained by stopping at stage n and replacing r_n with 0.

Proposition 5.7.5. *Let $z \in K((x^{-1}))$, let $n \geq 1$ and let f_n/q_n be its nth convergent. Then*

$$f_{n+1} = c_{n+1}f_n + f_{n-1} \quad \textit{and} \quad q_{n+1} = c_{n+1}q_n + q_{n-1}. \tag{5.21}$$

The initial conditions are $f_0 = c_0 = [z]$; $q_0 = 1$; and $f_{-1} = 1$; $q_{-1} = 0$. Moreover,

$$f_{n-1}q_n - q_{n-1}f_n = (-1)^n \tag{5.22}$$

so f_n and q_n are relatively prime. If $z = u/v$ with $u, v \in K[x]$ then the continued fraction expansion of z is finite and its length is at most the degree of v.

Proof. The proof of the first two statements is exactly the same as that of Proposition 5.7.2 and Lemma 5.7.3. For $n \geq 1$, the element r_n can be expressed as a quotient of polynomials with the degree of the numerator less than the degree of the denominator. The numerator at the nth stage is the denominator at the $(n+1)$st stage. Thus the degrees of the denominators are strictly decreasing. This implies the length of the expansion is no more than $\deg(v)$. □

Recall from Section 5.6.b that the field $K((x^{-1}))$ of formal Laurent series admits a metric,

$$\delta(z, w) = 2^{-\nu(z-w)}$$

for $z, w \in K((x^{-1}))$, where ν is the discrete valuation

$$\nu\left(\sum_{i=k}^{\infty} a_i x^{-i}\right) = \min\{i : \ a_i \neq 0\}.$$

If $u \in K[x]$ is a polynomial then $\nu(u) = -\deg(u)$.

Now fix $z \in K((x^{-1}))$. Let f_n/q_n be its nth convergent and set $e_n = \deg(q_n)$. The following theorem says that the continued fraction expansion converges, and that the convergents provide the best rational approximation to z.

Theorem 5.7.6. *For any $n \geq 1$, the power series expansion of z equals that of f_n/q_n in all terms involving x^k for $k > -(e_n + e_{n+1})$. That is, $f_n/q_n \equiv z \pmod{x^{-e_n - e_{n+1}}}$ or equivalently,*

$$\delta\left(z, \frac{f_n}{q_n}\right) \leq 2^{-e_n - e_{n+1}}. \tag{5.23}$$

If $r_n \neq 0$ then $e_{n+1} > e_n$. If $f, q \in K[x]$ are relatively prime and if $\delta(z, f/q) < 2^{-2e_n}$ then $\deg(q) > \deg(q_n)$ unless $f/q = f_n/q_n$.

Proof. If $r_n = 0$ then $z = f_n/q_n$ so to prove (5.23), we may assume that $r_n \neq 0$. As in the proof of Theorem 5.7.4,

$$z - \frac{f_n}{q_n} = \frac{(-1)^n}{q_n((1/r_n)q_n + q_{n-1})} = \frac{(-1)^n}{q_n(q_{n+1} + r_{n+1}q_n)}.$$

We will use the fact that $\nu(x+y) \geq \min(\nu(x), \nu(y))$ and that equality holds if $\nu(x) \neq \nu(y)$. By construction, if $n \geq 0$ we have $\nu(r_n) \geq 1$ so $\nu(1/r_n) \leq -1$ so for $n \geq 1$ we have:

$$\nu(c_n) = \nu((1/r_{n-1}) - r_n) = \nu(1/r_{n-1}) \leq -1.$$

Assuming $r_n \neq 0$ gives $c_{n+1} \neq 0$. By equation (5.21) and induction,

$$-e_{n+1} = \nu(q_{n+1}) = \nu(c_{n+1}q_n + q_{n-1}) < \nu(q_n) = -e_n. \tag{5.24}$$

It follows that $\nu(q_{n+1} + r_{n+1}q_n) = \nu(q_{n+1}) = -e_{n+1}$. Thus $\nu(z - f_n/q_n) = e_n + e_{n+1}$ as claimed.

Now suppose that $f, q \in K[x]$ are relatively prime and that $\nu(z - f/q) > 2e_n$. Assume that $e = \deg(q) \leq e_n = \deg(q_n)$. We must show that $f/q = f_n/q_n$. Assume for the moment that $\deg(f) \leq \deg(q)$ and $\deg(f_n) \leq \deg(q_n)$. (We will remove these assumptions below.) Let

$$\hat{f} = x^{-e}f, \quad \hat{q} = x^{-e}q, \quad \hat{f}_n = x^{-e_n}f_n, \text{ and } \hat{q_n} = x^{-e_n}q_n.$$

Then $\hat{f}, \hat{q}, \hat{f}_n, \hat{q}_n \in K[x^{-1}]$ are polynomials with

$$\frac{f}{q} = \frac{\hat{f}}{\hat{q}}, \text{ and } \frac{f_n}{q_n} = \frac{\hat{f}_n}{\hat{q}_n}.$$

Thus

$$\frac{\hat{f}}{\hat{q}} \equiv \frac{\hat{f}_n}{\hat{q}_n} \pmod{x^{-2e_n - 1}}$$

in the ring $K[[x^{-1}]]$. It follows that

$$\hat{f}\hat{q}_n \equiv \hat{f}_n\hat{q} \pmod{x^{-2e_n-1}}.$$

However, by assumption, the left and right sides of this congruence have degrees $\leq 2e_n$ in x^{-1}, so they are in fact equal. It follows then that $fq_n = f_nq$. Now we can take this as an equation in $K[x]$. Since f and q are relatively prime, f divides f_n and q divides q_n. Since f_n and q_n are relatively prime, f_n divides f and q_n divides q. It follows that $f/q = f_n/q_n$.

Now suppose that $\deg(f) \geq \deg(q)$ (or that $\deg(f_n) \geq \deg(q_n)$). Since $\delta(z, f/q) \leq 1$ this implies that $[z] = [f/q] = [f_n/q_n] = c_0$. So we may subtract off this integral part, and apply the previous case to the fractional part. In other words, let $z' = z - c_0$, $f' = f - c_0q$, and $f'_n = f_n - c_0q_n$. Then $\deg(f') < \deg(q)$ and $\deg(f'_n) < \deg(q_n)$, while $\delta(z', f'/q) = \delta(z, f/q)$ and $\delta(z', f'_n/q_n) = \delta(z, f_n/q_n)$. We conclude that $\deg(q) > \deg(q_n)$ unless $f'/q = f'_n/q_n$, in which case, by adding back the integral part c_0, we have $f/q = f_n/q_n$. □

5.7.c Continued fractions for Laurent series and p-adic numbers

Continued fractions can be developed almost identically for the field F of Laurent series in x, $F = K((x)) = \{\sum_{i=k}^{\infty} a_i x^i : a_i \in K\}$. In this case the "integer part" is a polynomial in x^{-1}. Every statement in Section 5.7.b now holds with x^{-1} replaced by x. The terms $c_n = [z_n]$ will be polynomials in x^{-1} and the convergents f_n/q_n will be quotients of polynomials in x^{-1}. By multiplying numerator and denominator by an appropriate power of x, we can convert these into approximations by ordinary rational functions, and the series of approximations will generally differ from that in Section 5.7.b because the metrics on $K((x))$ and $K((x^{-1}))$ are different.

A similar situation exists with the p-adic numbers. We can define continued fraction expansions for $z = \sum_{i=k}^{\infty} a_i 2^i$ with $a_i \in \{0, 1, \cdots, p-1\}$ (and k possibly negative) by taking the "integer part", $[z]$, to be the part involving non-positive powers,

$$[z] = \sum_{i=k}^{0} a_i 2^i \quad \text{and} \quad \{z\} = \sum_{i=1}^{\infty} a_i 2^i$$

to be the fractional part. The appropriate metric comes from the usual p-adic valuation. However, the set of "integral parts", (polynomials in 2^{-1}) is not closed under addition or multiplication. This leads to continued fraction expansions that do not converge. For example, consider the 2-adic CF expansion for $-1/2 = 2^{-1} + 2^0 + 2^1 + \cdots$. Since $[-1/2] = 2^{-1} + 2^0$ and $\{-1/2\} = 2^1 + 2^2 + \cdots = -2$ we obtain the infinite expansion

$$-\frac{1}{2} = 2^0 + 2^{-1} + \frac{1}{2^0 + 2^{-1} + \frac{1}{2^0 + 2^{-1} + \frac{1}{\cdots}}}$$

In Section 19.3.a we find the best rational approximation to a p-adic number using the theory of *approximation lattices*.

5.8 Exercises

1. Let $F = \mathbb{Q}$ be the rational numbers and $q(x) = x - \frac{1}{2}$. Show that the power series expansion of $1/q(x)$ is not eventually periodic.

2. Let F be a field and suppose that k is a positive integer that is invertible in F. Let $a(x) = \sum_{i=0}^{\infty} a_i x^i \in F[[x]]$ be a power series such that a_0 is a kth power in F. Show that a is a kth power in $F[[x]]$.

3. Let F be a field that is not algebraically closed. Show that $F[[x]]$ does not contain the algebraic closure of F.

4. If $a, b \in \mathbb{Z}_N$, make the definition of ab precise and show that $\mathbb{Z}_N$ is a ring.

5. Show that $\mathbb{Z}_3$ does not contain $\sqrt{-1}$. Show that $\mathbb{Z}_5$ contains two elements whose squares are -1, and compute the first 6 terms of each.

6. Use Theorem 5.4.8 to give an alternate proof that there is an injective homomorphism

$$\{f/g : f, g \in \mathbb{Z}, \gcd(g, N) = 1\} \to \mathbb{Z}_N.$$

7. Complete the details of the proof of Theorem 5.4.8, showing that all the appropriate homomorphisms commute.

8. Generalize Theorem 5.4.8 to π-adic integers. What properties of the ring R are needed to make this work?

9. Take $R = \mathbb{Z}$ and $\pi = 5$. Show that the element $5\pi^0 + 4\pi^1 + 4\pi^2 + \cdots \in \widehat{R}_\pi$ is in the kernel of $\widehat{\varphi}_n$ for all n.

10. Prove that the ring $\widehat{R}$ in Theorem 5.6.6 is an integral domain.

11. Finish the proof of Theorem 5.6.10.

12. Let R be a finite ring and let I be an ideal of R. Prove that the completion of R at I is a quotient ring of R.

13. Prove that if the continued fraction expansion of $z \in \mathbb{R}$ is eventually periodic, then z is a root of a quadratic polynomial with rational coefficients.

14. Use Hensel's lemma to determine which integers $m \in \mathbb{Z}$ have a square root in $\mathbb{Z}_p$.

15. (Reciprocal Laurent series) Let K be a field and let

$$z = \frac{x^3}{x^2 - 1} = x + x^{-1} + x^{-3} + \cdots = \sum_{i=0}^{\infty} x^{1-2i}.$$

Show that the continued fraction expansion of z is $[x, x, -x]$. That is,

$$x + x^{-1} + x^{-3} + \cdots = x + \frac{1}{x + \frac{1}{-x}}.$$

16. (Reciprocal Laurent series) Let K be a field whose characteristic is not equal to 2. Let $z^2 = (1 - x^{-1})$. Show that this equation has two solutions z in the ring $K[[x^{-1}]]$ of power series in x^{-1}. Hint: set $z = a_0 + a_1 x^{-1} + \cdots$, solve for $a_0 = \pm 1$. For $a_0 = +1$ solve recursively for a_n to find

$$0 = 2a_0 a_n + 2a_1 a_{n-1} + \cdots + \begin{cases} a_{n/2}^2 & \text{if } n \text{ is even} \\ 2a_{(n-1)/2} a_{(n+1)/2} & \text{if } n \text{ is odd.} \end{cases}$$

Do the same for $a_0 = -1$.

17. (continued) Show that the continued fraction expansion for the above z is

$$(1 - x^{-1})^{1/2} = 1 + \cfrac{1}{-2x + 1/2 + \cfrac{1}{8x - 4 + \cfrac{1}{-2x + 1 + \cfrac{1}{8x - 4 + \cfrac{1}{-2x + 1 + \cdots}}}}}$$

Part II

Algebraically Defined Sequences

Chapter 6 Linear Feedback Shift Registers and Linear Recurrences

Besides being interesting fundamental mathematical objects in their own right, linearly recurrent sequences have proved to be useful in many applications, including pseudorandom number generation, error correcting codes, private key cryptosystems, radar ranging, code division multiple access communications, and many other areas. They provide a fast and simple method of generating statistically random sequences. Moreover, many of their properties can be analyzed using various algebraic structures. The primary algebraic tools used to analyze linearly recurrent sequences are polynomials, power series, and trace functions on finite fields. The results in this section are all classical, many of them having been known for over a hundred years. However we have organized this section in a slightly unusual way (from the modern perspective) in order to better illustrate how they are parallel to the FCSR and AFSR theory which will be described in later chapters.

There are many ways to describe the output sequence of an LFSR, each of which has its merits. In this chapter we discuss the matrix presentation (Section 6.2), the generating function presentation (Theorem 6.4.1), the algebraic presentation (Proposition 6.6.1), the trace representation (Theorem 6.6.4) and the sums of powers representation (Theorem 6.6.8).

6.1 Definitions

In this section we give the definitions and describe the basic properties of linear feedback shift registers and linearly recurrent sequences. Throughout this chapter we assume that R is a commutative ring (with identity denoted by 1).

Definition 6.1.1. *A (Fibonacci mode)* linear feedback shift register *of length m over R, with coefficients $q_1, q_2, \cdots, q_m \in R$ is a sequence generator (cf. Definition 5.1.2) whose state is an element*

$$\mathbf{s} = (a_0, a_1, \cdots, a_{m-1}) \in R^m = \Sigma,$$

whose output is $\mathbf{out}(\mathbf{s}) = a_0$, and whose state change operation τ is given by

$$(a_0, a_1, \cdots, a_{m-1}) \longrightarrow (a_1, a_2, \cdots, a_{m-1}, \sum_{i=1}^{m} q_i a_{m-i}).$$

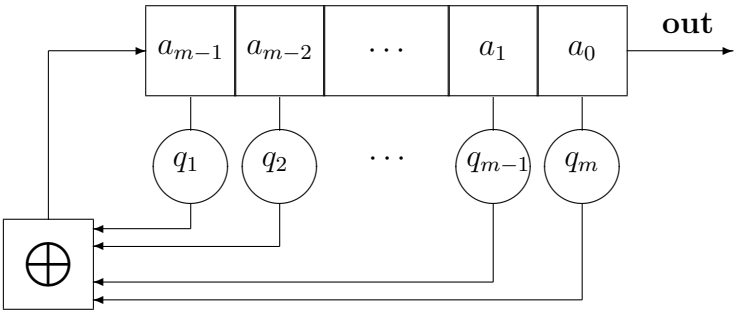


Figure 6.1: A Linear Feedback Shift Register of Length m.

(See also Section 10.1.) It is convenient to think of a linear feedback shift register (or LFSR) as a physical circuit, as pictured in Figure 6.1.

Note that, because we like to think of bits as flowing out to the right, the order of the components a_i in the diagram is the reverse of the order when we write the state as a vector. When thinking of LFSRs as physical devices as in the figure, we sometimes list the components of the state in descending order of their indices. We write $\boxed{a_{m-1}}\boxed{a_{m-2}}\boxed{\cdots}\boxed{a_1}\boxed{a_0}$ for the *machine state* to distinguish the two notations.

A LFSR defines a sequence generator $(R^m, R, \tau, \mathbf{out})$ in the sense of Section 5.1.c. As usual for a sequence generator, from any given initial state (or "initial loading") $(a_0, \cdots, a_{m-1})$, the LFSR generates an infinite *output sequence*

$$\mathbf{a} = a_0, a_1, \cdots, a_{m-1}, a_m, \cdots$$

This sequence may also be described as a linearly recurrent sequence (see Section 5.2.b).

Definition 6.1.2. *A sequence* $\mathbf{a} = a_0, a_1, \cdots$ *of elements of* R *is* linearly recurrent *if there exists a finite collection* $q_1, \cdots, q_m$ *of elements of* R *such that for all* $n \geq m$ *we have*

$$a_n = q_1 a_{n-1} + \cdots + q_m a_{n-m}. \tag{6.1}$$

Equation (6.1) is called the *recurrence relation*. The integer m is called the *degree* of the recurrence. The elements $q_1, \cdots, q_m$ are called the *coefficients* of the recurrence, to which we may associate the *connection polynomial*

$$q(x) = -1 + \sum_{i=1}^{m} q_i x^i \in R[x]. \tag{6.2}$$

For simplicity we let $q_0 = -1$, so that $q(x) = \sum_{i=0}^{m} q_i x^i$. A sequence $\mathbf{a} = a_0, a_1, \cdots$ satisfies a linear recurrence with coefficients $q_1, \cdots, q_m$ if and only if it may be realized as the output sequence of a LFSR with connection polynomial $q(x) = \sum_{i=0}^{m} q_i x^i$. In particular, LFSR sequences are eventually periodic.

The phrase "the LFSR with connection polynomial $q(x)$" is ambiguous because we can add initial cells (with no feedback taps) to an LFSR to obtain a new LFSR with the same connection polynomial $q(x)$ but of length $m > \deg(q)$. In this case the output sequence will consist of an initial "transient" $a_0, \cdots, a_{m-d-1}$ followed by the periodic part of the sequence. The LFSR will output strictly periodic sequences if and only if its length m equals the degree of its connection polynomial, that is, if $q_m \neq 0$. In order to be explicit, we will sometimes refer to "the LFSR with connection polynomial $q(x)$ whose length is the degree of $q(x)$".

Every periodic sequence $\mathbf{a}$ (of elements in a ring R) can be realized as the output of a LFSR (with entries in R): just take a LFSR whose length is a single period of $\mathbf{a}$, and feed the last cell back to the first cell. The number of cells in the shortest LFSR that can generate $\mathbf{a}$ is called the *linear complexity* or equivalent *linear span* of $\mathbf{a}$.

Example 6.1.3. As a first example, take $R = \mathbb{F}_2$ and consider the recurrence

$$a_4 = a_1 + a_0.$$

This recurrence corresponds to the connection polynomial $q(x) = x^4 + x + 1$ and to the LFSR pictured in Figure 6.2.

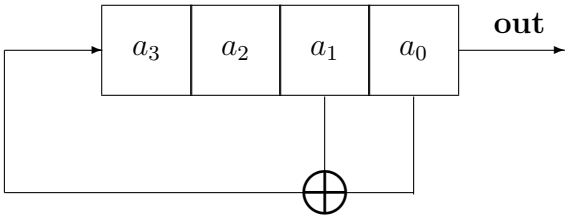


Figure 6.2: A Linear Feedback Shift Register of Length 4 over $\mathbb{F}_2$.

Note that we can construct an LFSR over $\mathbb{F}_2$ by simply including just the taps corresponding to coefficients $q_i = 1$ and omitting those corresponding to $q_i = 0$. The first 16 states and outputs from this LFSR are shown in Table 6.1 (where we write the machine states $[a_{m-1}, \cdots, a_0]$), starting in state $\boxed{1}\boxed{1}\boxed{1}\boxed{1}$. After this point the states repeat.

Example 6.1.4. Now let $R = \mathbb{F}_3$ and consider the recurrence

$$a_3 = 2a_2 + a_0.$$

	state	**out**		state	**out**		state	**out**		state	**out**
1	1111	1	5	1000	0	9	1100	0	13	1010	0
2	0111	1	6	0100	0	10	0110	0	14	1101	1
3	0011	1	7	0010	0	11	1011	1	15	1110	0
3	0001	1	8	1001	1	12	0101	1	16	1111	1

Table 6.1: States of the LFSR over $\mathbb{F}_2$ with $q(x) = x^4 + x^3 + 1$.

This recurrence corresponds to the connection polynomial $q(x) = x^3 + 2x - 1$ and to the LFSR pictured in Figure 6.3.

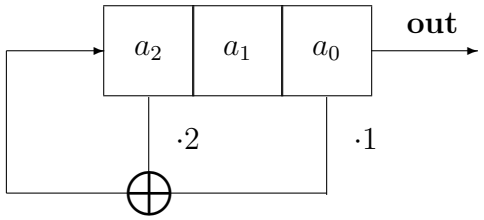


Figure 6.3: A Linear Feedback Shift Register of Length 3 over $\mathbb{F}_3$.

The first 14 states and output from this LFSR are shown in Table 6.2, starting in state $\boxed{1}\boxed{1}\boxed{1}$. After this point the states repeat.

	state	**out**		state	**out**		state	**out**		state	**out**
1	111	1	5	001	1	9	012	2	13	112	2
2	011	1	6	100	0	10	201	1	14	111	1
3	101	1	7	210	0	11	220	0	15		
3	010	0	8	121	1	12	122	2	16		

Table 6.2: States of the LFSR over $\mathbb{F}_3$with $q(x) = x^3 + 2x - 1$.

More generally, let $q(x) = \sum_{i=0}^{m} q_i x^i \in R[x]$ be any polynomial. We say the sequence **a** satisfies the linear recurrence defined by $q(x)$ if q_0 is invertible in R and if

$$q_0 a_n + q_1 a_{n-1} + \cdots + q_m a_{m-n} = 0$$

for all $n \geq m$. The linear recurrence defined by $q(x)$ is the same as the linear recurrence defined by $vq(x)$, for any invertible element $v \in R$.

As in Definition 5.3, if $q(x) = q_0 + q_1 x + \cdots + q_m x^m \in R[x]$ is a polynomial of degree m, then its *reciprocal polynomial* is

$$q^*(x) = q_0 x^m + q_1 x^{m-1} + \cdots + q_{m-1} x + q_m = x^m q(1/x). \tag{6.3}$$

Lemma 6.1.5. *The polynomials q and q^* have the same degree and the same order. The polynomial q is irreducible if and only if q^* is irreducible. If R is a field or a finite local ring (Section 4.1) then q is primitive if and only if q^* is primitive.*

Proof. The degrees are the same because q_0 and q_m are nonzero. If there exists $g \in R[x]$ such that $q(x)g(x) = x^N - 1$, then $q^*(x)g^*(x) = 1 - x^N$. If $q(x) = g(x)h(x)$, then $q^*(x) = g^*(x)h^*(x)$. If $\alpha \in S$ is a root of q in some extension field, then $\alpha^{-1} \in S$ is a root of $q^*(x)$. □

If $q(x)$ is the connection polynomial of a LFSR, then its reciprocal $q^*(x)$ is sometimes referred to as the *characteristic polynomial* of the LFSR. For some applications (such as Lemma 6.2.1) the polynomial q is better suited while for other applications (such as Theorem 6.4.1) the polynomial q^* is better suited. But we see later that in the setting of FCSRs and more generally AFSRs there is a meaningful analog of the connection polynomial but not of its reciprocal.

Suppose that $u \in R$ is invertible and that it is a root of the polynomial $q(x)$. Let $a \in R$. Multiplying the equation $\sum_{i=0}^{m} q_i u^i = 0$ by au^{-n} gives $\sum_{i=0}^{m} q_i (au^{-n+i}) = 0$ so we conclude the following fact, which is the basis for much of the analysis of linear recurrent sequences.

Basic fact 6.1.6. *Let $q(x) \in R[x]$ be a polynomial with coefficients in a commutative ring R. If $u \in R$ is a root of $q(x)$ and if u is invertible in R then for any $a \in R$, the sequence*

$$a, au^{-1}, au^{-2}, \cdots$$

satisfies the linear recurrence defined by $q(x)$.

6.2 Matrix description

6.2.a Companion matrix

Consider a linear feedback shift register of length m with connection polynomial $q(x) = \sum_{i=0}^{m} q_i x^i \in R[x]$ of degree m, with $q_0 = -1$. If we consider each state $\mathbf{s} = (a_0, a_1, \cdots, a_{m-1})$ to be a (column) vector in R^m then the state change (which shifts up) is given by matrix multiplication, $\mathbf{s}' = A\mathbf{s}$

where A is the *companion matrix*

$$A = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ & & \vdots & & \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ q_m & q_{m-1} & \cdots & q_2 & q_1 \end{pmatrix}. \tag{6.4}$$

The characteristic polynomial of A is $(-1)^{m+1}q^*(x)$ (see below). Recall that the *order* of a matrix A is the smallest positive integer T such that $A^T = I$. If such a T does not exist then we say that A does not have an order, or that its order is infinite. If $q \in R[x]$ is a polynomial with $q_0 = -1$, the *impulse response sequence* for q is the linearly recurring sequence with initial values $a_0 = 1$ and $a_i = 0$ for $1 \le i \le \deg(q) - 1$.

Theorem 6.2.1. *Let $q(x) = \sum_{i=0}^{m} q_i x^i \in R[x]$ with $q_0 = -1$. Let A be its companion matrix (6.4) and let q^* be its reciprocal polynomial (6.3). Then the following statements hold.*

1. $\det(xA - I) = (-1)^m q(x)$ *and* $\det(A - xI) = (-1)^{m+1}q^*(x)$. *The matrix A is invertible if and only if q_m is invertible in R. If the ring R is a field then the eigenvalues of A are the roots of q^*, which are the inverses of the roots of q.*
2. $q^*(A) = 0$, *that is,* $-A^m + q_1 A^{m-1} + \cdots + q_m I = 0$.
3. *If $h \in R[x]$, $h(A) = 0$, and $\deg(h) < \deg(q)$ then $h = 0$.*
4. *The set of polynomials $h \in R[x]$ such that $h(A) = 0$ forms a principal ideal; it is the ideal (q^*) generated by q^*.*
5. *The order of q^* (= the order of q) is equal to the order of the companion matrix A, and in particular one has a finite order if and only if the other does. If this order is finite then A is invertible. If the ring R is finite and if A is invertible then q, q^*, and A have finite order.*

Proof. **Part (1).** The determinants $\det(xA - I)$ and $\det(A - xI)$ are found by expanding by minors along the first column, and using induction. Since $\det(A) = (-1)^{m+1}q_m$, Cramer's rule gives a formula for A^{-1} provided q_m is invertible. The characteristic polynomial of A is q^*, so the eigenvalues of A are the roots of q^*, if R is a field. The roots of q^* are the reciprocals of the roots of q.

Part (2). To show that $q^*(A) : R^m \to R^m$ is the zero mapping, consider the action of $q^*(A)$ on a (column) vector $\mathbf{s} = (a_0, a_1, \cdots, a_{m-1})$. It is $q_m\mathbf{s} + q_{m-1}A\mathbf{s} + q_{m-2}A^2\mathbf{s} + \cdots + q_0 A^m\mathbf{s}$ or

$$q_m \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{m-1} \end{bmatrix} + q_{m-1} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} + \cdots + q_0 \begin{bmatrix} a_m \\ a_{m+1} \\ \vdots \\ a_{2m-1} \end{bmatrix}$$

where $q_0 = -1$ and $a_0, a_1, \cdots$ is the q-linearly recurrent sequence with initial values $a_0, a_1, \cdots, a_{m-1}$. Each row sums to 0 because the sequence is linearly recurrent with connection polynomial q.

Part (3). Suppose there exists a polynomial $h(x) = \sum_{i=0}^{r} h_i x^i$ with $r < m$ such that $h(A) = 0$. We may assume that $h_r \neq 0$. Let $\mathbf{s}$ be the machine state $\mathbf{s} = \boxed{1}\boxed{0}\boxed{\cdots}\boxed{0}$. Then the equation $h(A)(\mathbf{s}) = 0$ gives

$$h_0 \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix} + h_1 \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ * \end{bmatrix} + \cdots + h_r \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ * \\ \vdots \\ * \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \\ \vdots \\ 0 \end{bmatrix}.$$

This implies that $h_r = 0$ which is a contradiction.

Part (4). Let $h(x) = \sum_{i=1}^{r} h_i x^i$ be a polynomial such that $h(A) = 0$. We have already seen that if h is nonzero then $r = \deg(h) \geq m = \deg(q^*)$. By the Division Theorem (Theorem 2.4.2), we have $h(x) = g(x)q^*(x) + r(x)$ with $\deg(r) < m$. It follows that $r(A) = 0$, so $r(x) = 0$. Therefore h is a multiple of q^*. This proves that the ideal in part (4) is principal and is generated by q^*.

Part (5). To compare the orders of q^* and of A consider the mapping $R[x] \to \mathrm{Hom}(R^m, R^m)$ which associates to any polynomial $h(x) = \sum_{i=0}^{r} h_i x^i$ the homomorphism $h(A) = \sum_{i=0}^{r} h_i A^i$ (where $A^0 = I$). This homomorphism takes sums to sums and it takes products to the composition of homomorphisms. Moreover a polynomial $h \in R[x]$ is mapped to 0 if and only if it lies in the ideal (q^*) by part (4). Therefore this mapping passes to an injective homomorphism

$$\phi : R[x]/(q^*) \to \mathrm{Hom}(R^m, R^m),$$

and $\phi(x) = A$. Therefore $\phi(x^k) = A^k$ for all $k \leq \infty$ so the order of x in $R[x]/(q^*)$ coincides with the order of A. But the order of x in $R[x]/(q^*)$ is the order of q^*. □

6.2.b The period

Let

$$q(x) = \sum_{i=0}^{m} q_i x^i \in R[x]$$

be a polynomial of degree m and suppose that q_0 is invertible in R. Let q^* be the reciprocal polynomial. Let T be the order of q (which equals the order of q^*).

Corollary 6.2.2. *Let* $\mathbf{a}$ *be a periodic sequence of elements in* R *that satisfies the linear recurrence defined by* q*. Then the (minimal) period of* $\mathbf{a}$ *divides* T*. If* q_m *is invertible in* R *then the period of the impulse response sequence is exactly equal to* T*. If* R *is finite then* T *divides the order* $|R[x]/(q^*)^\times|$ *of the set of invertible elements in* $R[x]/(q^*)$*. If* R *is a finite field and if* q^* *is irreducible then* T *divides* $|R|^{\deg(q^*)} - 1$*. If* R *is a finite field then* q^* *is primitive if and only if* $T = |R|^{\deg(q^*)} - 1$*.*

Proof. Let A be the companion matrix (6.4) of q. Then $A^T = I$ by Theorem 6.2.1 Part (5). Let $\mathbf{s}_0 \in R^m$ be the vector $(a_0, a_1, \cdots, a_{m-1})$ representing the initial segment of the sequence $\mathbf{a}$. Then $A^T\mathbf{s}_0 = \mathbf{s}_0$ which implies that the sequence $\mathbf{a}$ has T as a period. But for any periodic sequence, every period is some multiple of the minimal period. Hence the minimal period of $\mathbf{a}$ divides T.

Next, suppose q_m is invertible and let r be the period of the impulse response sequence. Consider the first m states $\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_{m-1}$ of the shift register, corresponding to the impulse response sequence. Expressed as column vectors, they are

$$\mathbf{s_0} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ \mathbf{s_1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ a_m \end{bmatrix}, \ \mathbf{s_2} = \begin{bmatrix} 0 \\ \vdots \\ a_m \\ a_{m+1} \end{bmatrix}, \ \cdots, \ \mathbf{s_{m-1}} = \begin{bmatrix} 0 \\ a_m \\ \vdots \\ a_{2m-2} \end{bmatrix}$$

and $a_m = q_m$. Since q_m is invertible, these vectors form a basis of R^m. Now suppose the impulse response sequence has period $r < T$. Then $A^r\mathbf{s}_0 = \mathbf{s}_0$. So $A^r\mathbf{s}_1 = A^rA\mathbf{s}_0 = A\mathbf{s}_0 = \mathbf{s}_1$ and similarly $A^r\mathbf{s}_k = \mathbf{s}_k$ for each k. Since A is linear it follows that $A^rv = v$ for all $v \in R^m$. That is, $A^r = I$. Therefore the order of A divides r, which is a contradiction.

From the proof of Theorem 6.2.1 the mapping $\phi : R[x]/(q^*) \to \mathrm{Hom}(R^m, R^m)$ is one to one, and it maps the polynomial x to the companion matrix A. If the order T of A is finite then A is invertible (with inverse A^{T-1}), hence $x \in R[x]/(q^*)^\times$ is invertible. So the order of x (in $R[x]/(q^*)^\times$) divides the order of the group $R[x]/(q^*)^\times$. Finally, if R is a finite field and if q^* is irreducible then $R[x]/(q^*)$ is a field, whose multiplicative group contains $|R|^{\deg(q^*)} - 1$ elements. □

6.3 Initial loading

Let $q(x) \in R[x]$ be the connection polynomial as in equation (6.2) and set $q_0 = -1$. Define a mapping

$$\mu : \Sigma = R^m \to R[x]$$

from states of the shift register to polynomials by

$$(a_0, a_1, \cdots, a_{m-1}) \longrightarrow f(x) = \sum_{n=0}^{m-1} \left(\sum_{i=0}^{n} q_i a_{n-i} \right) x^n. \tag{6.5}$$

The importance of this polynomial is explained in Theorem 6.4.1 and Theorem 6.6.2.

Lemma 6.3.1. *The association $\mu : \Sigma = R^m \to R[x]$ is a one to one correspondence between states of the shift register and polynomials of degree $\leq m-1$. If $\mathbf{s} = (a_0, a_1, \cdots, a_{m-1})$ is a state and if $\mathbf{s}' = (a'_0, a'_1, \cdots, a'_{m-1})$ is the succeeding state, with corresponding polynomials $f(x)$ and $f'(x)$ respectively, then*

$$f(x) - xf'(x) = a_0 q(x). \tag{6.6}$$

Proof. For the first statement, it suffices to show that the numerator $f(x)$ determines a unique initial state $\mathbf{s}$ of the shift register (such that $\mu(\mathbf{s}) = f$). Suppose $f(x) = f_0 + f_1 x + \cdots + f_{m-1}x^{m-1}$. Then equation (6.5) gives $f_0 = q_0 a_0 = -a_0$ which determines a_0 uniquely. Then $f_1 = q_0 a_1 + q_1 a_0$, which, together with the knowledge of a_0, determines a_1 uniquely. Continuing in this way by induction we uniquely determine $a_2, a_3, \cdots, a_{m-1}$.

The second statement is a calculation. The leading term (corresponding to $n = m-1$) of the polynomial

$$f'(x) = \sum_{n=0}^{m-1} \sum_{i=0}^{n} q_i a'_{n-i} x^n$$

is

$$\begin{aligned} \sum_{i=0}^{m-1} q_i a'_{m-1-i} x^{m-1} &= \sum_{i=0}^{m-1} q_i a_{m-i} x^{m-1} + q_0 \sum_{j=1}^{m} q_j a_{m-j} x^{m-1} \\ &= -q_m a_0 x^{m-1}. \end{aligned}$$

Therefore

$$\begin{aligned} xf'(x) &= \sum_{n=0}^{m-2} \sum_{i=0}^{n} q_i a_{n-i+1} x^{n+1} - q_r a_0 x^m \\ &= \sum_{j=1}^{m-1} \sum_{i=0}^{j-1} q_i a_{j-i} x^j - q_m a_0 x^m \\ &= \sum_{j=0}^{m-1} \sum_{i=0}^{j} q_i a_{j-i} x^j - q_0 a_0 x_0 - \sum_{j=1}^{m-1} q_j a_0 x^j - q_m a_0 x^m \\ &= f(x) - a_0 q(x) \end{aligned}$$

□

6.4 Generating functions

Let $R[[x]]$ denote the ring of formal power series with coefficients in R (see Section 5.2). Recall that a polynomial (or even a formal power series)

$$h(x) = \sum_{i \geq 0} h_i x^i$$

is invertible in $R[[x]]$ if and only if $h_0 \in R$ is invertible in R, and in this case the inverse $h^{-1}(x) \in R[[x]]$ is uniquely defined. If $\mathbf{a} = a_0, a_1, \cdots$ is an infinite sequence of elements of R we define its *generating function* to be the formal power series

$$a(x) = \sum_{i=0}^{\infty} a_i x^i \in R[[x]]. \tag{6.7}$$

Conversely, to any formal power series $a(x)$ as in equation (6.7) we associate its *coefficient sequence* $\mathbf{seq}(a) = a_0, a_1, \cdots$. The following theorem describes the relation beween the shift register, its initial loading, and the generating function associated to its output sequence.

Theorem 6.4.1. *Let R be a commutative ring. Suppose a LFSR (over R) of length m with connection polynomial $q(x)$ of degree m has initial loading $(a_0, a_1, \cdots, a_{m-1})$. Let $f(x)$ be the polynomial in equation (6.5) that is determined by this initial loading. Then the output sequence $\mathbf{a}$ is the coefficient sequence of the power series expansion (Section 5.2.b) of the function $f(x)/q(x)$, that is,*

$$\frac{f(x)}{q(x)} = a_0 + a_1 x + a_2 x^2 + \cdots \tag{6.8}$$

or $\mathbf{a} = \mathbf{seq}(f/q)$. The output sequence $\mathbf{a}$ is eventually periodic if $q(x)$ divides $x^T - 1$ for some $T \geq 1$. In this case, the (eventual) period of $\mathbf{a}$ divides T. The output sequence $\mathbf{a}$ is strictly periodic if and only if $\deg(f) < \deg(q)$. If the ring R is finite and if $q_m \neq 0$ then every output sequence is strictly periodic.

Conversely, given any infinite eventually periodic sequence $\mathbf{b} = b_0, b_1, \cdots$ its generating function $b(x) = \sum_{i=0}^{\infty} b_i x^i$ may be uniquely expressed as the quotient of two polynomials $f(x)/q(x)$ such that $q_0 = -1$. In this case the sequence $\mathbf{b}$ may be generated by a LFSR of length $\max\{\deg(f)+1, \deg(q)\}$. The denominator $q(x)$ is the connection polynomial of the LFSR and the numerator $f(x)$ uniquely determines the initial loading by equation (6.5).

Proof. For any $n \geq m$ we have

$$\sum_{i=0}^{m} q_i a_{n-i} = 0.$$

So

$$\begin{aligned} q(x)a(x) &= \sum_{k=0}^{\infty}\sum_{i=0}^{m} q_i a_k x^{i+k} \\ &= \sum_{n=0}^{m-1}\left(\sum_{i=0}^{n} q_i a_{n-i}\right)x^n + \sum_{n=m}^{\infty}\left(\sum_{i=0}^{m} q_i a_{n-i}\right)x^n \\ &= f(x) + 0, \end{aligned}$$

which verifies (6.8). The remaining statements follow from Theorem 5.2.5 and Corollary 6.2.2.

Conversely, suppose we are given an eventually periodic sequence $\mathbf{b}$. By Theorem 5.2.5 its generating function is a rational function, say, $b(x) = f(x)/q(x)$, and the denominator $q(x)$ divides $x^T - 1$ (where T is the period of the periodic part). Hence the constant term q_0 is invertible in R. Multiplying numerator and denominator by $-q_0^{-1}$ we may assume that $q_0 = -1$. The preceding calculation shows that the sequence $\mathbf{b}$ satisfies the linear recurrence defined by $q(x)$. The remaining statements follow from Theorem 5.2.5 and Corollary 6.2.2. □

Corollary 6.4.2. *Suppose $a(x) = f(x)/q(x)$ is the generating function of an eventually periodic sequence $\mathbf{a}$ of elements of a field R, with $q_0 = -1$. If the fraction $f(x)/q(x)$ is in lowest terms, then the smallest shift register which can generate this sequence has size equal to* $\max\{\deg(f) + 1, \deg(q)\}$. *If the sequence $\mathbf{a}$ is strictly periodic then this number is* $\deg(q)$. *If $\mathbf{a} = c_1\mathbf{b_1} + c_2\mathbf{b_2}$ is the termwise linear combination of two eventually periodic sequences (with $c_1, c_2 \in R$) then the generating functions are likewise: $a(x) = c_1 b_1(x) + c_2 b_2(x)$.*

Thus, if $b_1(x) = f_1(x)/g_1(x)$ and $b_2(x) = f_2(x)/g_2(x)$ then

$$a(x) = \frac{c_1 g_2(x) f_1(x) + c_2 g_1(x) f_2(x)}{g_1(x) g_2(x)}$$

If this fraction is in lowest terms then the linear recurrence satisfied by $\mathbf{a}$ has degree $\deg(g_1) + \deg(g_2)$ but if the fraction can be further reduced then the degree of the linear recurrence satisfied by $\mathbf{a}$ will be smaller.

6.5 When the connection polynomial factors

Let R be a commutative ring and suppose that $q(x) = q_0 + q_1 x + \cdots + q_m x^m$, $q_m \neq 0$, is the connection polynomial of a LFSR U of length m, with $q_0 = -1$. Suppose also that the leading coefficient q_m is invertible in R and that the polynomial q factors

$$q(x) = c_0 r_1(x) r_2(x) \cdots r_k(x), \tag{6.9}$$

where $c_0 \in R$, the constant term of each r_i is -1, and the polynomials $r_i(x)$ are pairwise relatively prime in $R[x]$. (It follows that c_0 is $(-1)^k$.) Consider the linear feedback shift registers U_i of length $m_i = \deg(r_i)$ with connection polynomial r_i for $1 \leq i \leq k$.

Proposition 6.5.1. *For any initial loading of the shift register U, there are unique initial loadings of the shift registers U_i such that the output sequence of U is the term-wise sum of the output sequences of the shift registers U_i. Conversely, for any initial loadings of the shift registers U_i, the term-wise sum of the resulting sequences is the output sequence of U with some initial loading.*

Proof. Note that since $q_m \neq 0$, every initial loading of U produces strictly periodic output.

It suffices to consider the case of $k = 2$ factors, since the general case follows immediately by induction. The initial loading of U determines a polynomial $f(x)$ such that the generating function of the output sequence is the rational function $f(x)/q(x)$ and so that $\deg(f) < \deg(q)$. In the next paragraph we show that there is a unique "partial fraction" decomposition

$$\frac{f}{q} = \frac{h_1}{r_1} + \frac{h_2}{r_2} \tag{6.10}$$

with $\deg(h_i) < \deg(r_i)$. Viewing these rational functions as formal power series over R, this equation shows that the output of the shift register U is the sum of the outputs of the shift registers U_1 and U_2 with initial loadings determined by $h_1(x)$ and $h_2(x)$ respectively.

Since $(r_1) + (r_2) = R[x]$ (cf. Section 2.2.e) there exist polynomials $u_1(x), u_2(x)$ so that

$$f(x) = u_1(x)r_1(x) + u_2(x)r_2(x). \tag{6.11}$$

The leading coefficient of q is the product of the leading coefficients of r_1 and r_2 which are therefore both invertible in R. So we may use the division theorem for u_1/r_2 (Theorem 2.4.2) to write $u_1(x) = g_2(x)r_2(x) + h_2(x)$ for some (uniquely determined) polynomials $g_2, h_2 \in R[x]$ with $\deg(h_2) < \deg(r_2)$. Similarly, $u_2 = g_1r_1 + h_1$ with $\deg(h_1) < \deg(r_1)$. Then

$$f = (g_1 + g_2)r_1r_2 + h_2r_1 + h_1r_2.$$

The degrees of f, of h_2r_1, and of h_1r_2 are all less than $\deg(q) = \deg(r_1) + \deg(r_2)$. It follows that $g_1 + g_2 = 0$. This proves equation (6.10).

Conversely, suppose the outputs from U_i, $i = 1, 2$ have generating functions $f_i(x)/r_i(x)$. Then the term-wise sum of these sequences has generating function

$$\frac{f(x)}{q(x)} = \frac{f_1(x)r_2(x) + f_2(x)r_1(x)}{r_1(x)r_2(x)}.$$

Since $\deg(f) < \deg(q)$, this power series is an output sequence from U. □

It has been shown that the computational complexity of computing such partial fraction decompositions, even of multiple factors simultaneously, is quadratic in the degree of q [50, pp. 126–128]. It is not always the case that a polynomial q over a ring R can be factored as in equation (6.9), but see Theorem 2.2.14.

Repeated factors. In many cases (for example, if R is simultaneously a factorization domain and a GCD domain) Proposition 6.5.1 reduces the analysis of an LFSR with arbitrary connection polynomial to that of an LFSR with connection polynomial $q(x) = f(x)^t$ where $f(x)$ is an irreducible polynomial. We have seen in Section 4.1.a that the structure of the group of units in the ring $R[x]/(f^t)$ can be quite complicated. Fortunately, we do not need to fully understand the structure of this group of units in order to understand the behavior of a shift register with connection polynomial $q(x) = f(x)^t$.

If $f = \sum_{i=0}^{r} a_i x^i \in R[x]$ and if $n \geq 0$ is an integer, then define

$$f^{\{n\}}(x) = \sum_{i=0}^{r} a_i^n x^i.$$

If R is a field with t elements, and if n is a power of t then $f^{\{n\}} = f$. If R is a ring of characteristic p, where p is prime, and if n is a power of p then $(a+b)^n = a^n + b^n$ (because each binomial coefficient $\binom{n}{k}$ will be divisible by p), and hence

$$f(x)^n = f^{\{n\}}(x^n). \tag{6.12}$$

Proposition 6.5.2. *Suppose R is a ring of characteristic p where p is prime. Let $f \in R[x]$ be an irreducible polynomial with $f(0) = -1$. Let $q(x) = (-1)^{t+1} f(x)^t$, so that $q(0) = -1$. Let $n = p^k$ be the smallest power of p such that $n \geq t$. Then the output of the LFSR with connection polynomial $q(x)$ whose length is the degree of $q(x)$ is the interleaving of n LFSR sequences, each of which has connection polynomial $f^{\{n\}}(x)$.*

Note that if a polynomial $q(x)$ with $q(0) = -1$ is a unit times a power of an irreducible polynomial $f(x)$, then $f(0)$ is a unit. Thus $q(x)$ is also a unit times a power of $-f(0)^{-1} f(x)$, whose constant term is -1.

Proof. First we remark that if $h(x)$ and $u(x)$ are polynomials with $\deg(h) < \deg(u)$ then the coefficient sequence of the power series $g(x) = h(x^n)/u(x^n)$ is obtained from the coefficient sequence for $h(x)/u(x)$ by inserting $n-1$ zeroes between every pair of symbols.

Now let $g(x)$ be the generating function of the output sequence of the LFSR, so that

$$g(x) = \frac{v(x)}{(-1)^{t+1} f(x)^t} = \frac{v(x) f(x)^{n-t}}{(-1)^{t+1} f(x)^n} = \frac{h(x)}{f(x)^n}$$

where $h(x) = (-1)^{t+1} v(x) f(x)^{n-t}$ is a polynomial of degree $m < n \deg(f) = \deg(f^{\{n\}})$. Suppose $h(x) = \sum_{i=0}^{m} h_i x^i$. By collecting every nth term together we may write

$$h(x) = H_0(x^n) + xH_1(x^n) + \cdots + x^{n-1} H_{n-1}(x^n),$$

where $H_k(y) = h_k + h_{n+k} y + h_{2n+k} y^2 + \cdots$. Then $\deg(H_k) < \deg(f)$ and by equation (6.12) the generating function for the output sequence is

$$g(x) = \frac{H_0(x^n)}{f^{\{n\}}(x^n)} + x \frac{H_1(x^n)}{f^{\{n\}}(x^n)} + \cdots + x^{n-1} \frac{H_{n-1}(x^n)}{f^{\{n\}}(x^n)}.$$

This is an interleaving of n sequences, each of which has connection polynomial $(-1)^{t+1} f(y)$. □

Proposition 6.5.2 is most useful when R is finite and its characteristic is small relative to the power t. In Section 6.6 we also consider repeated factors when the characteristic of R is greater than t, making use of the following proposition.

Proposition 6.5.3. *Let R be a commutative ring, let*

$$f(x) = \sum_{i=0}^{d} f_i x^i \in R[x]$$

and let $u \in R$ be a root of $f(x)$. Assume u is invertible in R. Fix integers $0 \leq s < t$. Then the sequence

$$1, 1^s u^{-1}, 2^s u^{-2}, \cdots, m^s u^{-m}, \cdots \tag{6.13}$$

satisfies the linear recurrence defined by $q(x) = f(x)^t$.

Proof. Let $g(x)$ be the generating function for the sequence (6.13). We need to show that the power series

$$f(x)^t g(x) = f(x)^t \sum_{m=0}^{\infty} m^s u^{-m} x^m \tag{6.14}$$

is a polynomial of degree less than dt where d is the degree of f. That is, that all the terms of degree dt and greater vanish. We prove this by induction on t. Let us first compute $f(x)g(x)$:

$$\begin{aligned} f(x) \cdot \sum_{m=0}^{\infty} m^s u^{-m} x^m &= \sum_{i=0}^{d} \sum_{m=0}^{\infty} f_i m^s u^{-m} x^{i+m} \\ &= h(x) + \sum_{n=d}^{\infty} \sum_{i=0}^{d} f_i (n-i)^s u^{-n+i} x^n. \end{aligned}$$

where $n = i + m$ and where

$$h(x) = \sum_{n=0}^{d-1} \sum_{i=0}^{n} f_i (n-i)^s u^{-n+i} x^n$$

is a polynomial of degree $\leq d-1$. We can use the binomial theorem to obtain

$$\begin{aligned} f(x)g(x) &= h(x) + \sum_{n=d}^{\infty} \sum_{i=0}^{d} \sum_{r=0}^{s} \binom{s}{r} n^r (-i)^{s-r} f_i u^{-n+i} x^n \\ &= h(x) + \sum_{n=d}^{\infty} n^s u^{-n} x^n \cdot \left(\sum_{i=0}^{d} f_i u^i \right) \\ &\quad + \sum_{n=d}^{\infty} \sum_{i=0}^{d} \sum_{r=0}^{s-1} \binom{s}{r} (-i)^{s-r} f_i u^i n^r u^{-n} x^n. \end{aligned}$$

The middle term vanishes. Thus if $t = 1$, and hence $s = 0$, then we are done. Otherwise we are left with

$$\begin{aligned} f(x)g(x) &= h(x) + \sum_{r=0}^{s-1} \sum_{i=0}^{d} \binom{s}{r} (-i)^{s-r} f_i u^i \sum_{n=d}^{\infty} n^r u^{-n} x^n \\ &= h(x) + \sum_{r=0}^{s-1} A_r \sum_{n=d}^{\infty} n^r u^{-n} x^n \end{aligned}$$

where $A_r \in R$. By induction on t,

$$f(x)^{t-1} \cdot \sum_{n=0}^{\infty} n^r u^{-n} x^n$$

is a polynomial of degree less than $d(t-1)$ provided $r < t-1$. Thus

$$f(x)^t g(x) = f(x)^{t-1} h(x) + \sum_{r=0}^{s-1} A_r f(x)^{t-1} \sum_{n=d}^{\infty} n^r u^{-n} x^n$$

is a polynomial of degree at most $\max(d(t-1) + d - 1, d(t-1) - 1) = dt - 1 < dt$. □

6.6 Algebraic models and the ring $R[x]/(q)$

Throughout this section we fix a commutative ring R with identity and we consider a LFSR over R of length m with connection polynomial

$$q(x) = -1 + q_1 x + q_2 x^2 + \cdots + q_m x^m$$

with $q_0 = -1$. We also assume the leading coefficient q_m is nonzero and is invertible in R. Let $\Sigma = R^m$ denote the set of all possible states. We describe three different ways to represent the output of this LFSR using successive powers of an element (or elements) in some extension ring of R. In the language of Section 5.1.c, we are describing algebraic models for the sequence generator defined by an LFSR. The first theorem, Theorem 6.6.2 is the most general: it gives a complete mathematical model of the action of a LFSR over any ring R. However it is also the most abstract. The second, Theorem 6.6.4 (the familiar "Trace" description), and the third, Theorem 6.6.8 (the familiar sum of roots description) require some mild assumptions on the ring R. Let $R[x]$ be the ring of polynomials with coefficients in R.

6.6.a Abstract representation

The polynomial $q(x)$ determines an ideal (q) in the polynomial ring $R[x]$ and we consider the quotient ring $R[x]/(q)$. In this ring, x is invertible with inverse

$$x^{-1} = q_1 + q_2 x + \cdots + q_m x^{m-1} \tag{6.15}$$

as may be seen by multiplying both sides by x. Moreover,

$$x^m = \frac{1}{q_m}(1 - q_1 x - \cdots - q_{m-1} x^{m-1}) \tag{6.16}$$

because $q = 0$ in $R[x]/(q)$. It follows that any element $h \in R[x]/(q)$ may be uniquely represented as a polynomial $h(x)$ of degree less than m. If h_1, h_2 are two polynomials of degree $< m$ and if $a \in R$ then $h_1 + a h_2$ also has degree $< m$, so the sum of polynomials and scalar multiplication by elements of R agrees with the sum and scalar multiplication in the ring $R[x]/(q)$. However if $\deg h_1 + \deg h_2 \geq m$ then the product $h_1 h_2$ of polynomials will look very different from their product in $R[x]/(q)$.

If $h \in R[x]$ is a polynomial, then define

$$T(h) = h \pmod{q} \pmod{x} \in R$$

to be the following element: first, project h into the quotient $R[x]/(q)$ (by subtracting off the appropriate multiple of q), represent it by a polynomial $\widehat{h}$ of degree $\leq m-1$, then take the constant term, $\widehat{h}(0)$. In the same way, if $h \in R[x]/(q)$ we may refer to $h \pmod{q} \pmod{x} \in R$.

Proposition 6.6.1. *Let $q = -1 + q_1 x + \cdots + q_m x^m \in R[x]$ and assume that q_m is invertible. Then q is invertible in the ring $R[[x]]$ of power series. Let $h \in R[x]/(q)$ be a polynomial of degree $\leq m-1$ and let*

$$-\frac{h(x)}{q(x)} = a_0 + a_1 x + a_2 x^2 + \cdots$$

be the power series expansion of $-h(x)/q(x)$. *Then for all* $i \geq 1$,

$$a_i = x^{-i}h \pmod{q} \pmod{x}$$

where x^{-i} *is interpreted using equation (6.15).*

Proof. The polynomial $q(x)$ is invertible in $R[[x]]$ because $q_0 = -1$ is invertible in R. Therefore the quotient $-h(x)/q(x)$ can be uniquely expressed as a power series $a_0 + a_1x + a_2x^2 + \cdots$ with strictly periodic coefficients which, for $n \geq m$ satisfy the linear recurrence

$$a_n = q_1a_{n-1} + \cdots + q_ma_{n-m}.$$

Equating $q(x)(a_0 + a_1x + \cdots) = -h(x)$ gives

$$a_0 = h_0 = h \pmod{x}. \tag{6.17}$$

The power series $a_1+a_2x+a_3x^2+\cdots$ is also strictly periodic and satisfies the same linear recurrence so by Theorem 6.4.1 it is the power series expansion of a unique fraction $-h'(x)/q(x)$ where $\deg(h') < m$. Then

$$-\frac{h'}{q} = \frac{1}{x}(-\frac{h}{q} - a_0) \in R[[x]]$$

or $xh'(x) = h(x)+a_0q(x)$. But this is an equation among polynomials, so reducing modulo q gives $h' = x^{-1}h \pmod{q}$. Proceeding by induction we conclude that the power series $a_n + a_{n+1}x + a_{n+2}x^2 + \cdots$ (obtained by "shifting" the original power series by n steps) is the unique power series expansion of a fraction $h^{(n)}/q$ where $h^{(n)}$ is the polynomial representative of the element $x^{-n}h \pmod{q} \in R[x]/(q)$. Consequently, as in equation (6.17), the leading coefficient of this power series is

$$a_n = h_0^{(n)} = h^{(n)} \pmod{x} = x^{-n}h \pmod{q} \pmod{x} \qquad \square.$$

Returning to the shift register, define the mapping $\psi : R[x]/(q) \to \Sigma$ which associates to any $h \in R[x]/(q)$ the state

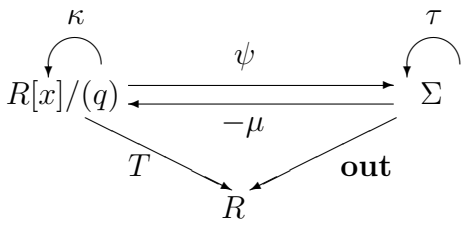
$$\psi(h) = (a_0, a_1, \cdots, a_{m-1})$$

of the shift register, where $a_i = x^{-i}h \pmod{q} \pmod{x}$. That is, the corresponding machine state is

$$\psi(h) = \begin{array}{|c|c|c|c|} \hline x^{-m+1}h \ (q)(x) & \cdots & x^{-1}h \ (q)(x) & h \ (q)(x) \\ \hline \end{array}$$

where $(q)(x)$ denotes $\pmod{q} \pmod{x}$. Recall the mapping $\mu : \Sigma \to R[x]$ of equation (6.5). The following is a concise statement of the discussion found in [159] Section 7.

Theorem 6.6.2. *The mappings $\psi : R[x]/(q) \to \Sigma$ and $-\mu : \Sigma \to R[x]/(q)$ are inverses of each other. In the language of* Section 5.1.c, *$\psi : (S, R, \kappa, T) \to (\Sigma, R, \tau, \mathbf{out})$ is a complete injective model and $-\mu : (\Sigma, R, \tau, \mathbf{out}) \to (S, R, \kappa, T)$ is a complete projective model. That is, the following diagram commutes.*



Here, $\tau : \Sigma \to \Sigma$ is the state change and $\kappa : R[x]/(q) \to R[x]/(q)$ is multiplication by x^{-1}. That is, $\kappa(h) = x^{-1}h$. As above, the mapping $T : R[x]/(q) \to R$ is given by $T(h) = h \pmod q \pmod x$. Then T is a homomorphism of R modules. That is, $T(h_1 + ah_2) = T(h_1) + aT(h_2)$ for any $a \in R$ and any $h_1, h_2 \in S$. However T is *not* a ring homomorphism. That is, $T(h_1h_2)$ does not equal $T(h_1)T(h_2)$ in general.

Proof. First we check that ψ commutes with the state change. That is, that $\tau(\psi(h)) = \psi(x^{-1}h)$. To do so let $h \in S$ and compute the successor to the state $\psi(h)$ which was described in the preceding diagram. The contents of each cell are shifted to the right and the contents of the leftmost cell is given by the linear recursion (6.1). That is,

$$\begin{aligned} a_m &= \sum_{i=1}^{m} q_i a_{m-i} \\ &= \sum_{i=1}^{m} q_i x^{i-m} h \pmod q \pmod x \\ &= x^{-m}(q(x) + 1)h \pmod q \pmod x \\ &= x^{-m} h \pmod q \pmod x. \end{aligned}$$

In other words, the successor state is $\psi(x^{-1}h)$ as claimed. We have that ψ commutes with the output functions since $\mathbf{out}(\psi(h)) = h \pmod q \pmod x$. Now consider the mapping $-\mu : \Sigma \to S$. It follows from equation (6.6) that for any state $\mathbf{s} \in \Sigma$ we have,

$$\mu(\mathbf{s}) \equiv x\mu(\mathbf{s}') \pmod q$$

where $\mathbf{s}' = \tau(\mathbf{s})$ is the succeeding state. This shows that μ is compatible with the state change, and the same holds for $-\mu$. To see that $-\mu$ commutes with the output functions, let $\mathbf{s} =$

$(a_0, a_1, \cdots, a_{m-1}) \in \Sigma$. Then $\mathbf{out}(\mathbf{s}) = a_0$. On the other hand, $\mu(\mathbf{s})$ is given by the polynomial $f(x)$ of equation (6.5), which has degree $\leq m-1$. So $f(x) \pmod q$ is given by this same polynomial, and its constant term is $f(x) \pmod q \pmod x = q_0 a_0 = -a_0 = -\mu(\mathbf{s})$.

Finally we verify that ψ and $-\mu$ are inverses. A state $\mathbf{s} \in \Sigma$ is determined by the m outputs $\mathbf{out}(\mathbf{s}), \mathbf{out}(\tau\mathbf{s}), \cdots, \mathbf{out}(\tau^{m-1}\mathbf{s})$ since these are just the initial cell contents of the shift register. However, for any state $\mathbf{s}$ we have shown that for any $n \geq 1$,

$$\begin{aligned} \mathbf{out}(\psi(-\mu(\tau^n \mathbf{s}))) &= T(\kappa^n(-\mu(\mathbf{s}))) \\ &= \mathbf{out}(\tau^n \mathbf{s}) \end{aligned}$$

from which it follows that $\psi(-\mu(\mathbf{s})) = \mathbf{s}$. □

Corollary 6.6.3. *The output sequence of the LFSR with the above initial loading $\psi(h)$ is given by $a_0, a_1, \cdots$ where*

$$a_i = x^{-i} h \pmod q \pmod x.$$

6.6.b Trace representation

In the next representation, we suppose that R is a commutative ring, that S is an extension ring (i.e. a ring which contains R) and that the connection polynomial $q(x)$ has a root $\alpha \in S$ which is invertible in S. We also suppose there exists an R-linear surjective mapping $T : S \to R$ (which in many cases will be a "Trace" mapping).

Theorem 6.6.4. *Fix a surjective R-linear mapping $T : S \to R$. Then, for any $A \in S$, the sequence*

$$T(A), T(\alpha^{-1}A), T(\alpha^{-2}A) \cdots \tag{6.18}$$

satisfies the linear recurrence defined by $q(x)$, with initial values depending on the choice of A.

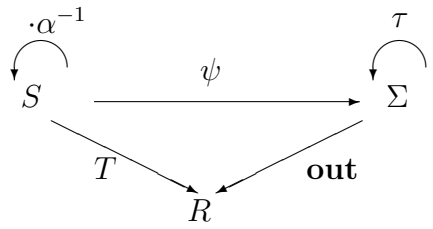
Proof. Set $a_n = T(A\alpha^{-n})$ and let $q(x) = q_0 + q_1 x + \cdots + q_m x^m$. Then

$$A\alpha^{-n}(q_0 + q_1\alpha + q_2\alpha^2 + \cdots + q_m\alpha^m) = 0$$

for any $n \geq 0$ and any $A \in R$. Applying T to both sides gives

$$\sum_{i=0}^{m} q_i T(A\alpha^{-n+i}) = \sum_{i=0}^{m} q_i a_{n-i} = 0$$ □

Hence the diagram



commutes, where $\Sigma = R^r$ is the set of states, $\tau : \Sigma \to \Sigma$ is the state change, the output sequence is given by equation (6.18), and $\psi : S \to \Sigma$ is given by $\psi(A) = (T(A), T(\alpha^{-1}A), \cdots, T(\alpha^{1-m}A))$. That is, the corresponding machine state is

$$\psi(A) = \boxed{T(\alpha^{1-m}A)} \boxed{\cdots} \boxed{T(\alpha^{-1}A)} \boxed{T(A)}$$

Thus $\psi : (S, R, \cdot\alpha^{-1}, T) \to (\Sigma, R, \tau, \mathbf{out})$ is an injective algebraic model for the shift register. The unit α^{-1} generates a cyclic subgroup of the group of units in S. So the deeper analysis of such an LFSR sequence relies on an understanding of the structure of the group of units in S. Perhaps the first question is to determine when such a model is complete.

Proposition 6.6.5. *Let R be a ring, let $S \supset R$ be an extension ring, let $T : S \to R$ be a surjective R-linear mapping, and let $q(x) \in R[x]$ be a polynomial of degree d. Suppose either*

1. *R is a finite local ring, $q(x)$ is a monic basic irreducible polynomial, and $S = GR(R, d)$ is the degree d Galois extension of R, so that by Theorem 4.4.1 q splits completely in S, or*
2. *R is a field, $q(x)$ is irreducible, and S is the splitting field of $q(x)$ (that is, S is a degree d extension of R and $q(x)$ splits into linear factors over S).*

Let α be a root of $q(x)$ in S. Then the above mapping $\psi : S \to \Sigma$ is a bijection (one to one and onto). Consequently every output sequence $a_0, a_1, \cdots$ of the LFSR is given by $a_n = T(\alpha^{-n}A)$ for some choice of $A \in S$.

Proof. First consider the case (1) that S is a Galois extension of R. By Corollary 4.4.2 α is invertible in S and the collection $\{1, \alpha, \alpha^2, \cdots, \alpha^{d-1}\}$ forms a basis for S over R. It follows that the collection $\{1, \alpha^{-1}, \alpha^{-2}, \cdots, \alpha^{1-d}\}$ also forms a basis for S over R. The ring S and the state space Σ contain the same number, $|R|^d$ of elements, so we need only show that the mapping ψ is one to one. Since ψ is R-linear, this amounts to showing that the kernel of ψ consists of $\{0\}$. By Theorem 4.4.6 the mapping T is nonsingular. However, if $\psi(a) = 0$ then $T(\alpha^{-i}a) = 0$ for $0 \le i \le d-1$ which implies that $T(ba) = 0$ for every $b \in S$. By Theorem 4.4.6 this implies that $a = 0$. The same argument also works in case (2), when R and S are fields. □

Repeated factors. The following theorem describes a "trace representation" for the output of a LFSR whose connection polynomial is a power of an irreducible polynomial. Suppose $R \subset S$ are rings and $f(x) \in R[x]$ is a polynomial with a root $\alpha \in S$ which is invertible in S. Let $T : S \to R$ be a surjective R-linear mapping.

Theorem 6.6.6. *Fix an integer $m \geq 1$ and set $g(x) = f(x)^m$. Then for any polynomial $P(y) = \sum_{k=0}^{m-1} P_k y^k \in S[y]$ of degree $\deg(P) \leq m-1$, the sequence*

$$a_n = T(P(n)\alpha^{-n}) = \sum_{r=0}^{m-1} n^r T(P_r \alpha^{-n}), n = 0, 1, \cdots, \tag{6.19}$$

satisfies the linear recurrence defined by $g(x)$.

Proof. In Proposition 6.5.3 it was shown that for any power $k \leq m-1$ the sequence $c_n = n^k \alpha^{-n} \in S$ satisfies the linear recurrence defined by $g(x)$. Hence the same is true for any linear combination

$$P(n)\alpha^{-n} = \sum_{k=0}^{m-1} P_k n^k \alpha^{-n}.$$

Since the coefficients of the recursion $g(x)$ are in R, the sequence $T(P(n)\alpha^{-n})$ also satisfies the linear recurrence defined by $g(x)$. □

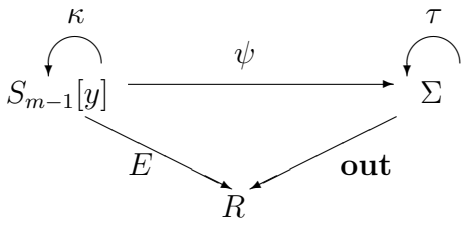
According to this theorem, we may construct the following mathematical "model" for the action of a shift register (of length md) with connection polynomial $g(x) = af(x)^m$. (In order that $g(x)$ be a connection polynomial we must have $g(0) = -1$, hence $f(0)$ must be invertible in R and then we can take $a = -f(0)^{-m}$.) Let $S_{m-1}[y]$ be the collection of all polynomials $P(y) \in S[y]$ of degree less than m. Define a state change operation $\kappa : S_{m-1}[y] \to S_{m-1}[y]$ by $\kappa(P)(y) = \alpha^{-1}P(y+1)$. (Note that κ does not change the degree of the polynomial P.) Let $E : S_{m-1}[y] \to R$ be the composition

$$S_{m-1}[y] \xrightarrow{\iota} S \xrightarrow{T} R$$

where $\iota(P) = P(0)$. Let us define a mapping $\psi : S_{m-1}[y] \to \Sigma$ to the set of states by $\psi(P) = (T(P(0)), T(P(1)\alpha^{-1}), \cdots, T(P(s)\alpha^{-s}))$. That is, the corresponding machine state is

$$\psi(P) = \begin{array}{|c|c|c|c|} \hline T(P(s)\alpha^{-s}) & \cdots & T(P(1)\alpha^{-1}) & T(P(0)) \\ \hline \end{array}$$

where $s = md - 1$. If $\tau : \Sigma \to \Sigma$ denotes the state change mapping for the shift register with connection polynomial $g(x)$ and length md, then the following diagram commutes,



Theorem 6.6.7. *Suppose that R is a finite field of characteristic $p \geq m$, that $f \in R[x]$ is irreducible of degree d, and that S is the degree d extension of R. Then the above model is complete. Hence, for any initial loading $(a_s, a_{s-1}, \cdots, a_1, a_0)$ of the shift register, there is a unique polynomial $P(y) \in S_{m-1}[y]$ so that equation (6.19) holds for all $n \geq 0$.*

Proof. The mapping ψ is R-linear, so it suffices to show that $\mathrm{Ker}(\psi) = 0$. Let $P(y) \in S[y]$ be a polynomial of some degree $\leq m-1 < p$. If $\psi(P) = 0$ then $T(P(n)\alpha^{-n}) = 0$ for all $0 \leq n \leq md-1$ from which it follows that $T(P(n)\alpha^{-n}) = 0$ for all $n \geq 0$. Since $\deg(P) < p$ there exists an integer t with $0 \leq t < p$ such that $P(t) \neq 0$. Let $a = P(t)\alpha^{-t} \in S$ and let $\beta = \alpha^p$. Then for all $k \geq 0$ we have

$$T(a\beta^{-k}) = T(P(t)\alpha^{-t-pk}) = T(P(t+pk)\alpha^{-t-pk}) = 0.$$

The mapping $b \mapsto b^p$ is a field automorphism of S which preserves R, from which it follows that $\beta = \alpha^p$ is also a root of an irreducible polynomial $f'(x) \in R[x]$ of degree d and therefore the collection $\{1, \beta^{-1}, \beta^{-2}, \cdots, \beta^{-d+1}\}$ forms a basis of S over R. Since S is a field, the element $a \neq 0$ is invertible. Therefore $T(s) = 0$ for all $s \in S$, which is a contradiction. □

Remarks. If $\mathrm{char}(R) = 0$ and if f splits into distinct linear factors over a degree d Galois extension S of R, then the same result holds for arbitrary m. The proof in [148], chapt. 13 works. When $\mathrm{char}(R)$ is arbitrary, Theorem 6.6.7 fails, but a general solution to the linear recurrence is described in [47] and [143]. These results are recalled in the next section, Section 6.6.c. Theorem 6.6.7 fails when R is a finite local ring because the element A constructed in the proof may be a zero divisor.

6.6.c Sums of powers representation

Consider a LFSR over a ring R with connection polynomial $q(x) \in R[x]$ whose degree is the length of the LFSR. Suppose S is a ring which contains R and suppose q factors into distinct linear factors over S. That is, there exist $\alpha_1, \alpha_2, \cdots, \alpha_m \in S$ such that

$$q(x) = q_m(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_m).$$

In this case we have the following result.

Theorem 6.6.8. *Suppose each root $\alpha_i \in S$ is invertible and also that the product*

$$\prod_{i<j}(\alpha_i^{-1} - \alpha_j^{-1})$$

is invertible in S. Then for any given initial loading $(a_0, a_1, \cdots, a_{m-1})$ of the shift register, there exist unique constants $A_0, A_1, \cdots, A_{m-1} \in S$ so that the output sequence is given by

$$a_n = \sum_{i=1}^{m} A_i \alpha_i^{-n}, \ n = 0, 1, \cdots. \tag{6.20}$$

Proof. Taking $n = 0, 1, \cdots, m-1$ in equation (6.20) gives a system of m linear (inhomogeneous) equations in the r unknowns $A_0, A_1, \cdots, A_{m-1}$ of Vandermonde type, meaning that the coefficient matrix is the following:

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_0^{-1} & \alpha_1^{-1} & \cdots & \alpha_{m-1}^{-1} \\ \alpha_0^{-2} & \alpha_1^{-2} & \cdots & \alpha_{m-1}^{-2} \\ & & \cdots & \\ \alpha_0^{2-r} & \alpha_1^{2-r} & \cdots & \alpha_{m-1}^{2-m} \end{pmatrix}$$

According to the theorem of Vandermonde, the determinant of this matrix is

$$\prod_{i<j}(\alpha_i^{-1} - \alpha_j^{-1}).$$

By assumption this determinant is invertible and hence the above system of equations has a unique solution. This gives elements $A_1, \cdots, A_m$ so that equation (6.20) is satisfied for $n = 0, 1, \cdots, m-1$.

It follows from Section 6.1.6 that the sequence $1, \alpha_i^{-1}, \alpha_i^{-2}, \cdots$ satisfies a linear recurrence with connection polynomial $q(x)$. Since the same recurrence is satisfied by the sequence of a_is, by induction on n equation (6.20) holds for all $n > 0$.

□

Repeated factors. Suppose R is a finite field of characteristic p, suppose $f \in R[x]$ is irreducible of degree d and let S be the degree d extension of R. Consider a LFSR over R having connection polynomial $q(x) = f(x)^m$ where $m \leq p$. Let $\alpha_1, \cdots, \alpha_d \in S$ be the roots of f. By taking $T = Tr_R^S$ in Theorem 6.6.7, we conclude the following. For any given initial loading of the shift register there exists polynomials $P_1(y), P_2(y), \cdots, P_d(y) \in S[y]$ such that the output sequence of the LFSR is:

$$a_n = \sum_{i=1}^{d} P_i(n) \alpha_i^{-n}.$$

The general solution of a linear recurrence with characteristic polynomial $g(x) = f(x)^m$ is due to [47]. Fix a field F of characteristic p (possibly $p = 0$). For each $h \geq 0$ define the sequence

$$\delta_k(n) = \binom{n+k}{k} \pmod{p}.$$

The sequences δ_k makes sense in any field of characteristic p, and it satisfies the linear recurrence defined by $(1-x)^m$ whenever $k \leq m-1$.

Theorem 6.6.9. *Let $f(x) \in F[x]$ be a polynomial with distinct roots in some algebraic closure $\bar{F}$ of F. Fix $m \geq 1$ and suppose $a_0, a_1, \cdots$ is a sequence of elements in k which satisfy the linear recurrence defined by $g(x) = f(x)^m$. Then there exist m unique sequences $\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \cdots, \mathbf{b}^{(m-1)}$, each of which satisfies the linear recurrence defined by $f(x)$, such that, for all $n \geq 0$,*

$$a_n = \sum_{i=0}^{m-1} \delta_i(n) b_n^{(i)}.$$

6.7 Families of recurring sequences and ideals

Until now, we have been studying individual sequences generated by a given LFSR (or, equivalently, individual sequences that satisfy a given linear recurrence). In this section we consider, for a fixed periodic sequence $\mathbf{a}$ (or for a given finite set A of periodic sequences) the set of linear recurrences that are satisfied by $\mathbf{a}$ (resp. by every $\mathbf{a} \in A$). This is the approach taken by Zierler [196].

6.7.a Families of recurring sequences over a finite field

Throughout this section, F denotes a finite field. Most of the results in this section may be found in Zierler's paper [196]. From Theorems 2.2.14 and 2.4.8 we have that the ring $F[x]$ of polynomials is a principal ideal domain and a Euclidean domain. Also, any polynomials $f_1, f_2 \in F[x]$ have a greatest common divisor $\gcd(f_1, f_2)$ and a least common multiple $\operatorname{lcm}(f_1, f_2)$ which are uniquely determined up to multiplication by an element of F, as described in Theorem 2.4.9.

In this section, we denote by $\mathcal{S} \subseteq F^\infty$ the collection of all (infinite) periodic sequences of elements of F. We identify the ring $F[[x]]$ of formal power series with F^∞ by associating to each series $a(x) = a_0 + a_1 x + \cdots$ its coefficient sequence $\mathbf{seq}(a) = (a_0, a_1, \cdots)$.

Let $q(x) = \sum_{i=0}^m q_i x^i \in F[x]$ be a polynomial of degree m. Consider the linear recurrence (Definition 6.1.2)

$$q_0 a_n + q_1 a_{n-1} + \cdots + q_m a_{n-m} = 0 \tag{6.21}$$

where $m = \deg(q)$. We refer to this as the linear recurrence defined by q. We say that $\mathbf{a}$ *satisfies the recurrence for n* if equation (6.21) holds. We say that $\mathbf{a}$ *satisfies the recurrence defined by*

q if it holds for every $n \geq m$. We say that $\mathbf{a}$ *eventually satisfies the recurrence*, or *satisfies the recurrence almost everywhere* (or a.e.) if it satisfies the equation for all n greater than or equal to some k.

Recall from Lemma 5.2.2 that q is invertible in $F[[x]]$ if and only if and if q_0 is nonzero. In this case if $u(x) \in F[x]$ is another polynomial, then the power series expansion

$$\frac{u}{q} = a(x) = \sum_{i=0}^{\infty} a_i x^i$$

is well defined. The sequence $\mathbf{a} = \mathbf{seq}(u/q) = (a_0, a_1, \cdots)$ of its coefficients satisfies the recurrence (6.21) for all $n \geq \max(m, \deg(u) + 1)$. The sequence $(a_0, a_1, \cdots)$ is (strictly) periodic if and only if $\deg(u) < m$. More generally, if $\mathbf{a}$ is a sequence of elements of F and q_0 is possibly zero, then $\mathbf{a}$ eventually satisfies the linear recurrence defined by q if and only if the formal power series $q(x)a(x) \in F[[x]]$ is in fact a polynomial. The sequence satisfies the linear recurrence defined by q if and only if this polynomial has degree less than m, and this holds if and only if $\mathbf{a}$ is (strictly) periodic.

Definition 6.7.1. *If $\mathbf{a} \in F^\infty$, denote by $I(\mathbf{a})$ the set of all $q = q(x) \in F[x]$ such that* $\mathbf{a}$ eventually *satisfies the linear recurrence defined by q. If $A \subset F^\infty$ is a set of sequences, denote by $I(A)$ the intersection of the $I(\mathbf{a})$ over all $\mathbf{a} \in A$. If $q \in F[x]$, then denote by $G(q)$ the set of all sequences* $\mathbf{a}$ *that eventually satisfy the linear recurrence (6.21) defined by q. Set $G_0(q) = G(q) \cap \mathcal{S}$, the set of sequences that satisfy the recurrence defined by q.*

Thus $q \in I(\mathbf{a})$ if and only if $\mathbf{a} \in G(q)$.

Suppose q_0 is nonzero. Then $G(q)$ is identified with the set of all fractions $u/q \in F[[x]]$ and $G_0(q)$ is identified with those fractions such that $\deg(u) < \deg(q)$. The sets $G(q)$ and $G_0(q)$ are vector spaces (with $G_0(q)$ finite dimensional) which are closed under the shift operation, for if $\mathbf{a}$ satisfies the linear recurrence defined by q then so does every left shift of $\mathbf{a}$.

Lemma 6.7.2. *Suppose that q_0 is nonzero. Then $q(x)$ divides some $x^T - 1$. It follows that every element of $G(q)$ is eventually periodic.*

Proof. We have $1 - x(-q_0^{-1} \sum_{i=1}^{m} q_i x^{i-1}) = q_0^{-1} q(x)$, so that x is a unit in $F[x]/(q(x))$. Since F is finite, the group of units in $F[x]/(q(x))$ is a finite group, so every unit has finite order. Thus for some $T > 0$, $q(x)$ divides $x^T - 1$.

To see the second statement, let $x^T - 1 = q(x)s(x)$. If $\mathbf{a} \in G(q)$, then

$$a(x) = \frac{u(x)}{q(x)} = \frac{u(x)s(x)}{x^T - 1} = u(x)s(x)(1 + x^T + x^{2T} + \cdots).$$

The latter product is periodic for indices greater than $\deg(u) + \deg(s)$ since it is a sum of the sequences corresponding to certain $x^j \sum_{i=0}^{\infty} x^{iT}$, each of which is periodic for indices greater than $\deg(u) + \deg(s)$. □

Suppose $\mathbf{a}$ is periodic and $q \in I(\mathbf{a})$. We can write $q(x) = x^k q'(x)$ with $q'(0) \neq 0$. Let $q'(x)$ divide $x^T - 1$, say $x^T - 1 = s(x)q'(x)$. We may assume that T is a multiple of the period of $\mathbf{a}$. Thus

$$a(x) = \frac{v(x)}{q'(x)} = \frac{s(x)v(x)}{x^T - 1}$$

for some polynomial v. Since the period is less than T, we must have $\deg(sv) < T$, so also $\deg(v) < \deg(q')$. Thus $q(x)a(x)$ is a polynomial of degree less than the degree of $q(x)$. That is, $\mathbf{a}$ satisfies the linear recurrence defined by q (everywhere, not just almost everywhere). If $\mathbf{a}$ is not eventually periodic, then by Lemma 6.7.2 $I(\mathbf{a}) = \{0\}$. That is, $\mathbf{a}$ satisfies no recurrence, even a.e.

For any $\mathbf{a}$, we claim that the set $I(\mathbf{a})$ is an ideal in $F[x]$. To see this, suppose that $q(x)a(x)$ and $r(x)a(x)$ are polynomials and $u(x)$ is any polynomial. Then $(q(x)+r(x))a(x) = q(x)a(x)+r(x)a(x)$ and $(u(x)q(x))a(x) = u(x)(q(x)a(x))$ are polynomials, so $I(\mathbf{a})$ is indeed an ideal.

Theorem 6.7.3. *For every eventually periodic sequence* $\mathbf{a}$ *there is a nonzero polynomial* $q(x)$ *of minimal degree, called the* minimal polynomial*, such that* $\mathbf{a} \in G(q)$*. It is uniquely determined up to multiplication by a nonzero element of* F *and its constant term is nonzero. Conversely, for any polynomial* $q \in F[x]$ *of degree greater than 1 whose constant term is nonzero, there exists a periodic sequence* $\mathbf{a}$ *whose minimal polynomial is* q*. For any finite collection* A *of sequences there exists a unique shift register of shortest length that generates every sequence in the set* A*.*

Proof. Since, given $\mathbf{a}$, the set $I(\mathbf{a})$ is an ideal in $F[x]$, and $F[x]$ is a principal ideal domain, there exists a polynomial $q \in F[x]$ of minimal degree, uniquely determined up to a scalar multiple, such that $I(\mathbf{a}) = (q)$. The polynomial q has a nonzero constant term, for if it were divisible by x then $\mathbf{a}$ would also satisfy the linear recurrence defined by $q(x)/x$. This has a smaller degree, which is a contradiction.

For the converse statement, let $q \in F[x]$ and let $\mathbf{a}$ be the *impulse response sequence*, that is, the unique sequence $(a_0, a_1, \cdots)$ in $G(q)$ with initial values $a_i = 0$ for $0 \leq i \leq \deg(q) - 2$ and $a_{\deg(q)-1} = 1$. Then $\mathbf{a}$ does not satisfy any linear recurrence of degree less than $\deg(q)$ (otherwise it would be the zero sequence), so q is its minimal polynomial. □

If $\mathbf{a}$ is a sequence, then we let $\mathbf{a}^r$ be its left shift by r steps.

Lemma 6.7.4. *If* $\mathbf{a} \in F^\infty$ *is a sequence, then* $I(\mathbf{a}) = I(\mathbf{a}^r + \mathbf{b})$ *for every integer* $r \geq 0$ *and every sequence* $\mathbf{b}$ *with only finitely many nonzero terms. Suppose that* $\mathbf{a}$ *is eventually periodic and let* d *be the degree of the minimal polynomial of* $\mathbf{a}$*. The vector space* $G_0(I(\mathbf{a}))$ *is closed under left shift. It has a basis consisting of* $\{\mathbf{a}^r : 0 \leq r < d\}$*. The vector space* $G(I(\mathbf{a}))$ *is closed*

under left shift and under arbitrary changes to finitely many positions. It has a basis consisting of $\{\mathbf{a}^r : 0 \le r < d\} \cup \{\mathbf{b} : \mathbf{b}$ *has one nonzero term*$\}$.

Proof. If $\mathbf{a}$ is not eventually periodic, then $I(\mathbf{a}^r + \mathbf{b})$ is the zero ideal so the first statement is trivial. Now suppose $\mathbf{a}$ is eventually periodic. Then $\mathbf{a}$ differs from a periodic sequence by a $\mathbf{b}$ with only finitely many nonzero terms, so we may assume that $\mathbf{a}$ is periodic. Let $a^{(r)}(x)$ be the generating function of $\mathbf{a}^r$. If $a(x) = a^{(0)}(x) = \sum_{i=0}^{\infty} a_i x^i$, then $a^{(1)}(x) = (a(x) - a_0)/x$. If $f \in I(\mathbf{a})$ then $f(x)a(x)$ is a polynomial. Therefore $f(x)(a(x) - a_0)$ is a polynomial and so also $f(x)(a(x) - a_0)/x$. Thus $f \in I(\mathbf{a}^{(1)})$. By inducton $I(\mathbf{a}) \subset I(\mathbf{a}^{(1)}) \subset I(\mathbf{a}^{(2)}) \subset \cdots$. But $\mathbf{a}$ is periodic so $\mathbf{a}^{(N)} = \mathbf{a}$ where N is its period, hence these containments are all equalities. Moreover, if $\mathbf{b}$ has only finitely many nonzero terms, then its generating function $b(x)$ is a polynomial. The generating function of $\mathbf{a}^r + \mathbf{b}$ is $a^{(r)}(x) + b(x)$, and $f(x)a^{(r)}(x)$ is a polynomial if and only if $f(x)(a^{(r)}(x) + b(x))$ is a polynomial. Thus $I(\mathbf{a}^{(r)} + \mathbf{b}) = I(\mathbf{a}^r) = I(\mathbf{a})$.

To prove the second statement, we first show that the sequences $\mathbf{a}, \mathbf{a}^{(1)}, \cdots, \mathbf{a}^{(d-1)}$ are linearly independent. Suppose there is a relation of linear dependence, say,

$$\sum_{i=0}^{d-1} c_i \mathbf{a}^{(i)} = 0$$

(for some choice of constants $c_i \in F$). This equation says that the polynomial $\sum_{i=0}^{d-1} c_i x^i$ is in $I(\mathbf{a})$. But this polynomial has degree at most $d - 1 < d$ which is a contradiction. Finally, the vector space $G_0(q)$ has dimension equal to d because the sequences in $G_0(q)$ are uniquely determined by the d initial values $a_0, a_1, \cdots, a_{d-1}$. Therefore the elements $\mathbf{a}, \mathbf{a}^{(1)}, \cdots, \mathbf{a}^{(d-1)}$ must be a basis for $G_0(q)$.

Similarly, the sequences $\{\mathbf{a}^{(r)} : 0 \le r < d\} \cup \{\mathbf{b} : \mathbf{b}$ has one nonzero term$\}$ are linearly independent since no nontrivial linear combination of the $\mathbf{a}^{(r)}$ has only finitely many nonzero terms. They span $G(I(\mathbf{a}))$ since every eventually periodic sequence differs from some periodic sequence in finitely many terms. □

If $I = (f)$ is an ideal in $F[x]$ then we denote by $G(I)$ the set of all sequences that satisfy the linear recurrences associated with every element of I. Then by Theorem 6.7.3, $G(I) = G(f)$. Similarly for G_0. Recall (from Theorem 2.4.9) that if (f_1) and (f_2) are ideals in $F[x]$ then $(f_1) + (f_2) = (\gcd(f_1, f_2))$ is the smallest ideal containing both (f_1) and (f_2), and that $(f_1) \cap (f_2) = (\mathrm{lcm}(f_1, f_2))$ is the largest ideal contained in both (f_1) and (f_2).

Theorem 6.7.5. *([196]) Let* $f, f_1, f_2 \in F[x]$ *be polynomials with nonzero constant terms and let* A, A_1, $A_2 \subset \mathcal{S}$ *be sets of eventually periodic sequences. Then the following statements hold.*

1. $G(f_1) \subset G(f_2)$ *if and only if* $G_0(f_1) \subset G_0(f_2)$ *if and only if* f_1 *divides* f_2.

2. *If* $\mathbf{a} \in G(f)$ *then the minimum polynomial of* $\mathbf{a}$ *divides* f.
3. $G(f_1) \cap G(f_2) = G(\gcd(f_1, f_2)) = G((f_1) + (f_2))$ *and* $G_0(f_1) \cap G_0(f_2) = G_0(\gcd(f_1, f_2)) = G_0((f_1) + (f_2))$.
4. $G(f_1) + G(f_2) = G(\mathrm{lcm}(f_1, f_2)) = G((f_1) \cap (f_2))$ *and* $G_0(f_1) + G_0(f_2) = G_0(\mathrm{lcm}(f_1, f_2)) = G_0((f_1) \cap (f_2))$.
5. $I(G(f)) = I(G_0(f)) = (f)$.
6. *If* $A_1 \subset A_2$ *then* $I(A_2) \subset I(A_1)$.
7. *Suppose every sequence in* A *is periodic. Then* $A \subset G_0(I(A))$. *If* A *is finite, then* $A = G_0(I(A))$ *if and only if also* A *is a finite dimensional vector space that is closed under left shift.*
8. *In general* $A \subset G(I(A))$. *Let* B *be the set of sequences with finitely many nonzero terms.* $A = G(I(A))$ *if and only if* A *is a vector space and contains a finite subspace* A_0 *which is closed under left shift such that* $A = A_0 + B$.
9. *If* $A_1, A_2 \subset \mathcal{S}$ *are vector spaces over* F *then* $I(A_1 + A_2) = I(A_1) \cap I(A_2)$.
10. *Suppose* $A_1, A_2 \subset \mathcal{S}$ *are vector spaces over* F, *and each* A_i *can be written as* $A_i = A_{i,0} + C$ *with* $A_{i,0}$ *a finite vector space closed under left shift and* C *an arbitrary vector space (this holds, for example, if both* A_1 *and* A_2 *are finite). Then* $I(A_1 \cap A_2) = I(A_1) + I(A_2)$.

Proof. For part (1), If $f_2(x) = u(x)f_1(x)$ for some polynomial $u(x)$ and if $\mathbf{a} \in G_0(f_1)$ then $f_1(x)a(x)$ is a polynomial of degree less than $\deg(f_1)$, so $u(x)f_1(x)a(x)$ is a polynomial of degree less than $\deg(f_1) + \deg(u) = \deg(f_2)$. Conversely, suppose $G_0(f_1) \subset G_0(f_2)$. Thus $\mathbf{seq}(1/f_1) \in G_0(f_1) \subseteq G_0(f_2)$, so that $f_2/f_1 = u$ is a polynomial. Thus $f_2 = uf_1$ so $f_1 | f_2$. A similar argument works in the non-periodic case by omitting the degree constraints.

For part (2), let q denote the minimum polynomial of $\mathbf{a}$. Then $\mathbf{a} \in G(f)$ implies that $f \in I(\mathbf{a}) = (q)$. This says that q divides f.

For part (3), if $g = \gcd(f_1, f_2)$ then $G_0(g) \subset G_0(f_i)$ $(i = 1, 2)$ by part (1), so $G_0(g) \subset G_0(f_1) \cap G_0(f_2)$. To verify the other inclusion, let $\mathbf{a} \in G_0(f_1) \cap G_0(f_2)$. There exist polynomials u_1, u_2 so that $g = u_1 f_1 + u_2 f_2$ from which it follows that

$$g(x)a(x) = u_1(x)f_1(x)a(x) + u_2(x)f_2(x)a(x)$$

is again a polynomial. Thus $\mathbf{a} \in G(g)$. But $\mathbf{a}$ is periodic, so $\mathbf{a} \in G_0(g)$. A similar argument works in the non-periodic case.

For part (4), we know from part (1) that

$$G_0(f_1) \subset G_0(\mathrm{lcm}(f_1, f_2)) \text{ and } G_0(f_2) \subset G_0(\mathrm{lcm}(f_1, f_2)).$$

But $G_0(\mathrm{lcm}(f_1, f_2))$ is a vector space so it contains the sum $G_0(f_1) + G_0(f_2)$. Therefore it will suffice to prove that they have the same dimension. Now

$$\dim(G_0(\mathrm{lcm}(f_1, f_2))) = \deg(f_1) + \deg(f_2) - \deg(\gcd(f_1, f_2))$$

since $\operatorname{lcm}(f_1, f_2) = f_1 f_2 / \gcd(f_1, f_2)$. On the other hand,

$$\dim(G_0(f_1) + G_0(f_2)) = \deg(f_1) + \deg(f_2) - \dim(G_0(f_1) \cap G_0(f_2)).$$

By part (3),

$$\dim(G_0(f_1) \cap G_0(f_2)) = \deg(\gcd(f_1, f_2)).$$

This argument does not work in the non-periodic case since the Gs are infinite dimensional. However, for any $f \in F[x]$ the set

$$G_k(f) = \{\mathbf{a} \in G(f) : \deg(f(x)a(x)) < \deg(f) + k\}$$

is a finite dimensional vector space and

$$G(f) = \cup_{k=0}^{\infty} G_k(f).$$

It is possible to generalize the proofs of parts (4) and (3) to G_k (some care is required in obtaining the degree bound in part (3)). We leave the details to the reader. This completes the proof of Part (4)

For part (5) we have $f \in I(G_0(f)) \subset I(G(f))$ so $(f) \subset I(G_0(f))$. On the other hand, the sequence $\mathbf{seq}(1/f)$ is in $G(f)$. Thus if $g \in I(G(f))$, then $g(x)/f(x) = u(x)$ is a polynomial, so $g = uf \in (f)$. Part (6) is straightforward.

Suppose every sequence in A is periodic. For any $\mathbf{a} \in A$ we have

$$\mathbf{a} \in G_0(I(\mathbf{a})) \subset G_0(I(A)),$$

from which the first statement in part (7) follows. If $A = G_0(I(A))$ then A is a vector space, closed under the shift operation because this is true of every $G_0(f)$. On the other hand, suppose that A is a vector space that is closed under the shift operation. We must show that $G_0(I(A)) \subset A$. For any $\mathbf{a} \in A$ the set $G_0(I(\mathbf{a}))$ has a basis consisting of certain shifts of $\mathbf{a}$, by Lemma 6.7.4. Since A is a vector space and it contains all the shifts of $\mathbf{a}$ we conclude that $G_0(I(\mathbf{a})) \subset A$. But $\mathbf{a}$ was an arbitrary element of A. By Part (4),

$$G_0(I(A)) = G_0(\bigcap_{\mathbf{a} \in A} I(a)) = \sum_{\mathbf{a} \in A} G_0(I(a)) \subset A.$$

This completes the proof of part (7).

For any $\mathbf{a} \in A$ we have

$$\mathbf{a} \in G(I(\mathbf{a})) \subset G(I(A)),$$

from which the first statement in Part (8) follows. Suppose $A = G(I(A))$ and let $A_0 = G_0(I(A))$. Then A is a vector space and A_0 is a subspace is closed under left shift. Every eventually periodic

sequence differs from a periodic sequence in finitely many terms, so $A = A_0 + B$. On the other hand, suppose that $A = A_0 + B$ is a vector space with A_0 a subspace closed under left shift. Then by part (7),

$$\begin{aligned} G(I(A)) &= G(I(A_0 + B)) = G(I(A_0)) \\ &= G_0(I(A_0)) + B = A_0 + B = A. \end{aligned}$$

This completes the proof of part (8).

To verify part (9) first note that by Part (6), $I(A_1 + A_2) \subset I(A_j)$ for $j = 1, 2$ hence

$$I(A_1 + A_2) \subset I(A_1) \cap I(A_2).$$

It remains to verify the reverse inclusion. There exist polynomials f_1, f_2 such that $I(A_1) = (f_1)$ and $I(A_2) = (f_2)$. Then

$$I(A_1) \cap I(A_2) = (f_1) \cap (f_2) = (h)$$

where $h = \operatorname{lcm}(f_1, f_2)$. So it suffices to prove that $h \in I(A_1 + A_2)$. That is, we must show: if $\mathbf{a} \in A_1 + A_2$, then $h(x)a(x)$ is a polynomial. Such an element $\mathbf{a}$ can be written as a sum $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$ where $\mathbf{a}_i \in A_i$. Since h is a multiple of f_1 we know that $h(x)a_1(x)$ is a polynomial. Similarly $h(x)a_2(x)$ is a polynomial. Therefore

$$h(x)a(x) = h(x)a_1(x) + h(x)a_2(x)$$

is also a polynomial, as claimed.

For part (10), let $I(A_j) = (f_j)$ for $j = 1, 2$ and let $g = \gcd(f_1, f_2)$ so that $(f_1) + (f_2) = (g)$. Then if C is finite so each A_i is finite,

$$\begin{aligned} I(A_1) + I(A_2) = (g) &= I(G_0(g)) && \text{by Part (5)} \\ &= I(G_0((f_1) + (f_2))) \\ &= I(G_0(f_1) \cap G_0(f_2)) && \text{by Part (3)} \\ &= I(G_0(I(A_1)) \cap G_0(I(A_2))) \\ &= I(A_1 \cap A_2) && \text{by Part (7).} \end{aligned}$$

If B is infinite, then an identical proof with G replaced by G_0 works. This completes the proof of Theorem 6.7.5. □

Most of the applications of this theorem appear in Chapters 13 and 15.

There is another point of view for the material in this section that is often useful. Note that F^∞ is a module over F using componentwise addition and scalar multiplication. Define $E : F^\infty \to F^\infty$ to be the shift operator, $E(a_0, a_1, \cdots) = (a_1, a_2, \cdots)$. Then E commutes with addition and scalar multiplication, so we can multiply an element of F^∞ by an arbitrary polynomial in $F[E]$. This makes F^∞ into a module over $F[E]$. If $\mathbf{a} \in F^\infty$, then we say that a polynomial $g(E)$ *annihilates* $\mathbf{a}$ if $g(E)\mathbf{a} = \mathbf{0}$, the all zero sequence.

Theorem 6.7.6. *Suppose $a(x) = f(x)/q(x)$ is the generating function of an eventually periodic sequence $\mathbf{a} = \mathbf{seq}(a)$ of elements of a field F. Let E be the shift operator on sequences over F. Then*

$$g(E) = E^{\max(\deg(f)-\deg(q)+1,0)}q^*(E) \tag{6.22}$$

is an annihilator of $\mathbf{a}$ (where q^* is the reversal of q). Conversely, if $g(E) = E^t q^*(E)$ is an annihilator of $\mathbf{a}$ for some natural number t and polynomial $q(E)$, then $q(x)a(x)$ is a polynomial of degree at most $t + \deg(q) + 1$.

Proof. Let $d = \deg(f)$ and $m = \deg(q)$. We have $q(x)a(x) = f(x)$. Equating the coefficients of terms on either side of this equation, it follows that for all $n \geq \max(d+1, m)$,

$$q_0 a_n + q_1 a_{n-1} + \cdots + q_m a_{n-m} = 0.$$

This equals the term with index $n - m$ in $q^*(E)\mathbf{a}$. Thus

$$E^{\max(\deg(f)-\deg(q)+1,0)}q^*(E)\mathbf{a} = \mathbf{0}.$$ □

6.7.b Families of Linearly Recurring Sequences over a Ring

In this subsection we consider the generalization to the case when the field F is replaced by an arbitrary finite ring R. Let $\mathcal{S} \subseteq R^\infty$ be the collection of all (infinite) periodic sequences of elements of R. As usual we identify the ring $R[[x]]$ of formal power series with R^∞ by associating to each series $a(x) = a_0 + a_1 x + \cdots$ its coefficient sequence $\mathbf{seq}(a(x)) = (a_0, a_1, \cdots)$.

If $q(x) = \sum_{i=0}^m q_i x^i$ is any nonzero polynomial in $R[x]$ we may again consider the linear recurrence (Definition 6.1.2)

$$q_0 a_n + q_1 a_{n-1} + \cdots + q_m a_{n-m} = 0. \tag{6.23}$$

If $\mathbf{a} = \mathbf{seq}(a(x))$ is a sequence, then $\mathbf{a}$ satisfies the recurrence (6.23) (for all $n \geq m$) if and only if $q(x)a(x)$ is a polynomial $f(x)$ of degree less than $m = \deg(q)$. Note that we can only conclude that $a(x) = f(x)/q(x)$ for some polynomial $f(x)$, if q_0 is a unit in R. Also, if q_0 is not a unit, then the recurrence (6.23) may not uniquely determine $\mathbf{a}$ from the initial values $a_0, a_1, \cdots, a_{m-1}$.

The definitions of $G(q)$, $G_0(q)$, and $I(\mathbf{a})$ are unchanged from Section 6.7. The set $I(\mathbf{a})$ is still an ideal, however it may not be a principal ideal: there may not be a unique minimal degree polynomial in $I(\mathbf{a})$. For example, if b is a common annihilator of all terms of $\mathbf{a}$, then the degree zero polynomial $q(x) = b$ is in $I(\mathbf{a})$. But there are many cases when $\mathbf{a}$ is also annihilated by polynomials some of whose coefficients are units. For a specific example, let $R = F[u,v]/(u,v)^2$, where F is a finite field. Thus R is a finite local ring with maximal ideal $\mathfrak{m} = (u, v)$ and quotient field F. The sequence $\mathbf{a} = (u, u, u, \cdots)$ is annihilated by the degree zero polynomials u and v and by the polynomial $x - 1$. These three polynomials are pairwise relatively prime. In fact a

polynomial $q(x) = \sum_{i=0}^{m} q_i x^i$ annihilates $\mathbf{a}$ if and only if $u(q_0 + q_1 + \cdots + q_m) = 0$. This holds if and only if

$$\sum_{i=0}^{m} q_i \in \mathfrak{m}.$$

For any such $q(x)$, we have

$$q(x) = \sum_{i=1}^{m} q_i x^i - \sum_{i=1}^{m} q_i + \sum_{i=0}^{m} q_i.$$

The polynomial

$$\sum_{i=1}^{m} q_i x^i - \sum_{i=1}^{m} q_i$$

is a multiple of $x - 1$, so

$$q(x) \in \mathfrak{m} \cdot R[x] + (x-1) = (u, v, x-1).$$

In particular, there is no reasonable notion of minimal polynomial for $\mathbf{a}$ since $I(\mathbf{a}) = (u, v, x-1)$ is not principle and has no distinguished generator.

For a given nonzero polynomial q we may consider $G(q)$ and $G_0(q)$. We may ask, for example, whether they contain nonzero sequences and whether they are finite. They are modules over R, although they may not be finitely generated. They are closed under left shift.

The easiest case is when q_0 is a unit in R. Then any sequence $\mathbf{a} = (a_0, a_1, \cdots)$ that satisfies the recurrence (6.23) for all but finitely many n is eventually periodic — each term is determined by the previous m terms and there are finitely many such m-tuples since R is finite. The sequence $\mathbf{a}$ is strictly periodic if and only if it satisfies the recurrence for all $n \geq m = \deg(q)$. This is equivalent to saying that $\deg(q(x)a(x)) < \deg(q(x))$.

When q_0 is not a unit, the picture is more complicated for a general ring. However, if R is a finite local ring we can say something about $G(q)$. To do so, we generalize $G(q)$. Let M be a module over R and let $\mathbf{a} = (a_0, a_1, \cdots) \in M^\infty$ be an infinite sequence of elements of M. Then we say that $\mathbf{a}$ satisfies the recurrence defined by q if equation (6.23) holds for all $n \geq m$. We let $G(q, M)$ be the set of sequences in M^∞ that satisfy the recurrence defined by q. Note that it is not necessary that m be the degree of q. That is, we may allow $q_m = 0$. In general $G(q, M)$ is a module over R, and may not be finitely generated. However, if R is a field, q is nonzero, and M is a finite dimensional vector space over R, then $G(q, M)$ is a finite dimensional vector space over R. Indeed, if we choose a basis for M so that we can write $M = R^t$ for some t, then $G(q, M) = G(q)^t$.

Theorem 6.7.7. *Let R be a finite local ring with maximal ideal $\mathfrak{m}$ and quotient field F. Let $q(x)$ be a nonzero polynomial with coefficients in R. If every coefficient of q is in $\mathfrak{m}$, then $G(q)$ is uncountably infinite. In particular, it contains sequences that are not eventually periodic. If at least one coefficient of q is a unit, then $G(q)$ is finite.*

Proof. Suppose that $\mathfrak{m}^k = (0)$ but that $\mathfrak{m}^{k-1} \neq (0)$. If every coefficient of $q(x)$ is in $\mathfrak{m}$, then $G(q)$ contains every sequence all of whose entries are in $\mathfrak{m}^{k-1}$ (and possibly other sequences). Since there are at least two possibilities for each position, $G(q)$ is uncountable. Since the set of eventually periodic sequences is countable, $G(q)$ contains sequences that are not eventually periodic.

Now suppose that at least one coefficient of q is a unit. Let μ_i be the reduction map from $\mathfrak{m}^i$ to $\mathfrak{m}^i/\mathfrak{m}^{i+1}$. We use the same name for the induced map on various types of objects defined over $\mathfrak{m}^i$ (polynomial or sequences, for example). Thus $\mu_0(q) \neq 0$. Each $\mathfrak{m}^i/\mathfrak{m}^{i+1}$ is a finite dimensional vector space over F, so $G(\mu_0(q), \mathfrak{m}^i/\mathfrak{m}^{i+1})$ is finite.

Note that if $\mathbf{a} \in G(q)$, then $\mu_0(\mathbf{a}) \in G(\mu_0(q))$. Since $\mu_0(q)$ is nonzero, $G(\mu_0(q), \mathfrak{m}^i/\mathfrak{m}^{i+1})$ is finite for each i. For each i we can choose elements

$$\mathbf{b}_{i,j} \in G(q, \mathfrak{m}^i), j = 1, \cdots, t_i$$

for some t_i, with

$$\mu_i(G(q, \mathfrak{m}^i)) = \{\mu_i(\mathbf{b}_{i,j}) : j = 1, \cdots, t_i\},$$

and the $\mu_i(\mathbf{b}_{i,j})$ distinct. Then an induction argument shows that we have

$$\mathbf{a} = \sum_{i=0}^{k-1} \mathbf{b}_{i,j_i} \tag{6.24}$$

for some $j_0, \cdots, j_{k-1}$. Every such sum is in $G(q)$, so $G(q)$ has $\prod_{i=0}^{k-1} t_i$ elements and thus is finite. □

Let $\mathbf{a} \in \mathcal{S}$ be a periodic sequence and let $\mathbf{a}^r$ be its shift by r steps. As is the case when R is a field, $I(\mathbf{a}) = I(\mathbf{a}^r)$ for every shift r and the module $G(I(\mathbf{a}))$ is closed under the shift operation.

Theorem 6.7.8. *Let $J, J_1, J_2 \in R[x]$ be ideals and let $A, A_1, A_2 \subset \mathcal{S}$ be finite sets of periodic sequences. Then the following statements hold.*

1. *If $J_2 \subseteq J_1$, then $G(J_1) \subset G(J_2)$.*
2. *$G(J_1) \cap G(J_2) = G(J_1 + J_2)$.*
3. *$G(J_1) + G(J_2) \subseteq G(J_1 \cap J_2)$ (but equality does not hold in general).*
4. *$J \subset I(G(J))$ (but equality does not hold in general).*
5. *If $A_1 \subset A_2$ then $I(A_2) \subset I(A_1)$.*
6. *$A \subset G(I(A))$ (but equality fails in general even if A is a module closed under shift).*
7. *If $A_1, A_2 \subset \mathcal{S}$ are modules over R then $I(A_1 + A_2) = I(A_1) \cap I(A_2)$.*
8. *If $A_1, A_2 \subset \mathcal{S}$ are modules, then $I(A_1 \cap A_2) \subset I(A_1) + I(A_2)$ (but equality does not hold in general, even if A_1 and A_2 are closed under translation).*

Proof. The proof of part (1) is similar to the proof of part (1) of Theorem 6.7.5.

For part (2), if $\mathbf{a} \in G(J_1) \cap G(J_2)$, then $\mathbf{a}$ satisfies the recurrence defined by every element of J_1 or J_2. Since recurrences can be added, $\mathbf{a}$ also satisfies the recurrence defined by every element of $J_1 + J_2$. Conversely, by part (1) $G(J_1 + J_2) \subseteq G(J_1)$ and $G(J_1 + J_2) \subseteq G(J_1)$, so $G(J_1 + J_2) \subseteq G(J_1) \cap G(J_2)$.

For part (3), we know from part (1) that $G(J_1) \subset G(J_1 \cap J_2)$ and $G(J_2) \subset G(J_1 \cap J_2)$. Thus $G(J_1) + G(J_2) \subset G(J_1 \cap J_2)$.

Part (4) holds since $f \in I(G(f)) \subset I(G(J))$ if $f \in J$.

Part (5) is elementary.

Part (6) holds since, for any $\mathbf{a} \in A$ we have $\mathbf{a} \in G(I(\mathbf{a})) \subset G(I(A))$.

To verify part (7) first note that by part (5), $I(A_1+A_2) \subset I(A_j)$ for $j = 1, 2$, hence $I(A_1+A_2) \subset I(A_1) \cap I(A_2)$. It remains to verify the reverse inclusion. Suppose that $q \in I(A_1) \cap I(A_2)$. Then every $\mathbf{a}_1 \in A_1$ and every $\mathbf{a}_2 \in A_2$ satisfies the recurrence defined by q. It follows that every $\mathbf{a}_1 + \mathbf{a}_2 \in A_1 + A_2$ satisfies the recurrence defined by q. Thus $I(A_1) \cap I(A_2) \subset I(A_1 + A_2)$, which proves part (7)

Part (8) follows from part (5) This completes the proof of Theorem 6.7.8. □

Counterexamples to equality in parts (3), (4), (6), and (8) can be found in the finite local ring $R = F[u, v]/(u, v)^2$ with maximal ideal $\mathfrak{m} = (u, v)$. For part (3), take $I_1 = (u) = uR[x]$ and $I_2 = (v) = vR[x]$. Then $G(I_1) = G(I_2) = \mathfrak{m}^\infty$, so $G(I_1) + G(I_2) = \mathfrak{m}^\infty$. But $I_1 \cap I_2 = (0)$, so $G(I_1 \cap I_2) = R^\infty$.

For part (4), take $I = (u)$. Thus $I(G(I)) = I(\mathfrak{m}^\infty) = (u, v)$.

For part (6), let $A = R \cdot (u, u, \cdots)$. Then A is a module closed under shift. We have $G(I(A)) = G(\mathfrak{m} \cdot R[x] + (1 - x)) = \{a^\infty : a \in \mathfrak{m}\}$.

For part (8), let $A_1 = R \cdot (u, u, \cdots)$ and $A_2 = R \cdot (v, v, \cdots)$. Then A_1 and A_2 are modules closed under shift. We have $I(A_1) = I(A_2) = I(A_1) + I(A_2) = \mathfrak{m} \cdot R[x] + (1 - x)$. But $I(A_1 \cap A_2) = I(0^\infty) = R[x]$.

We also mention that the approach using the shift operator E is valid for sequences over rings. Theorem 6.7.6 holds with F replaced by any ring.

6.8 Examples

6.8.a Shift registers over a field

Shift register sequences over a field F have been extensively utilized by the electrical engineering community, (particularly when $F = \mathbb{F}_2$). First let us consider the case of a LFSR whose connection polynomial $f(x) \in F[x]$ is irreducible. Let $d = \deg(f)$ and let K be the unique degree d extension

of F. Then $f(x)$ splits into linear factors over K. If $\alpha \in K$ is a root of $f(x)$ then for any initial loading of the shift register there exists a unique $A \in K$ such that the sequence

$$a_n = Tr_F^K(A\alpha^{-n})$$

is the output sequence of the shift register.

Suppose that $F = \mathbb{F}_q$ is the finite field with q elements. Then

$$\alpha^{q^d-1} = 1$$

so the period of this sequence is a divisor of $q^d - 1$. If $\alpha \in K$ is a primitive element (in which case f is a primitive polynomial) then this sequence has maximal period $T = q^d - 1$, and it is called an *m-sequence* over $\mathbb{F}_q$.

Returning to the general case, let us assume that F is a field. Because $F[x]$ is a unique factorization domain, it is possible to completely describe the output of a shift register over F with arbitrary connection polynomial. Let $g(x) \in F[x]$ be the connection polynomial of a shift register U where $g(0) = -1$. We assume that the length of the register is equal to $\deg(g)$. Then, for any initial loading, the output will be strictly periodic. The polynomial g factors uniquely (up to permutation of the factors),

$$g(x) = af_1(x)^{m_1} f_2(x)^{m_2} \cdots f_t(x)^{m_t},$$

where the f_i are distinct irreducible polynomials and $a \in F$ is chosen so that each f_i is a connection polynomial ($f_i(0) = -1$). The output of the shift register U is then the term-by-term sum of the outputs of the t shift registers U_i having connection polynomial $f_i^{m_i}$ (where $1 \leq i \leq t$) with initial loadings which are uniquely determined by the initial loading of U.

This reduces the analysis to that of a single shift register with connection polynomial $af(x)^m$. If $\deg(f) = d$ then the polynomial f splits into linear factors over its splitting field $K \supset F$. (If $F = \mathbb{F}_q$ then $K = \mathbb{F}_{q^d}$ where $d = \deg(f)$.) Let $\alpha \in K$ be a root of f. If $\mathrm{char}(F) = 0$ or if $m < \mathrm{char}(F)$ then there exists a unique polynomial $P(y) \in K[y]$ (which is determined by the initial loading $a_0, a_1, \cdots, a_{md-1}$) such that the output sequence is given by

$$a_n = Tr_F^K(P(n)\alpha^{-n})$$

for all $n \geq 0$. If $m \geq p = \mathrm{char}(F)$ then the output sequence is obtained as the interleaving of the outputs of n shift registers with connection polynomial $f(x)$, where $n = p^s$ is the smallest power of p such that $p^s \geq m$.

6.8.b Fibonacci numbers

The Fibonacci sequence $1, 1, 2, 3, 5, 8, \cdots$ satisfies the linear recurrence

$$a_n = a_{n-1} + a_{n-2} \tag{6.25}$$

over the ring $R = \mathbb{Z}$ of integers. The connection polynomial is $q(x) = -1 + x + x^2$ and indeed,

$$\frac{-1}{-1 + x + x^2} = 1 + x + 2x^2 + 3x^3 + 5x^4 + 8x^5 + \cdots$$

as predicted by Theorem 6.4.1. We consider the three exponential representations for this sequence.

According to Corollary 6.6.3, sequences satisfying this linear recurrence may be expressed as $hx^{-n} \pmod q \pmod x$ where $h, x^{-1} \in \mathbb{Z}[x]/(q)$. In the ring $\mathbb{Z}[x]/(q)$ we have $x^{-1} = 1 + x$ (since $x(1 + x) = x + x^2 = 1 + q \equiv 1 \pmod q$). It turns out that, in order to obtain the Fibonacci sequence, we must take $h = x$. That is, $a_n = x^{-n+1} \pmod q \pmod x$. This follows immediately from the formula $x^{-n} = a_{n+1} + a_n x$ which is verified by induction.

The roots of the connection polynomial are:

$$x = \frac{-1 \pm \sqrt{5}}{2},$$

the *golden mean* and its Galois conjugate. One may consider these roots to lie in the real numbers $\mathbb{R}$, or in the smaller extension field $\mathbb{Q}[\sqrt{5}]$ or even the extension ring $\mathbb{Z}[\sqrt{5}, 1/2]$. In any case, according to Theorem 6.6.8 there are unique constants A, B such that

$$a_n = \frac{A}{2}(-1 + \sqrt{5})^{-n} + \frac{B}{2}(-1 + \sqrt{5})^{-n},$$

and in fact we may take $A = 1/\sqrt{5}$ and $B = -1/\sqrt{5}$.

Choose any $\mathbb{Q}$-linear mapping $T : \mathbb{Q}(\sqrt{5}) \to \mathbb{Q}$, for example, the trace Tr assigns to $A \in \mathbb{Q}(\sqrt{5})$ the trace of the action of A on $\mathbb{Q}(\sqrt{5})$ when it is considered as a 2- dimensional vector space over $\mathbb{Q}$, (a basis of which may be taken to be $\{1, \sqrt{5}\}$). Then, as in Theorem 6.6.4, if $\alpha = (1 + \sqrt{5})/2$, then $T(A\alpha^{-n})$ will satisfy the linear recurrence (6.25), the choice of A being determined by the initial conditions. Taking $A = 1/\sqrt{5}$ will give the Fibonacci sequence: $a_n = \mathrm{Tr}(\alpha^{-n}/\sqrt{5})$.

6.9 Exercises

1. Phase taps: Consider an LFSR with primitive connection polynomial $q(x) = -1 + q_1 x + \cdots + q_m x^m$. Let α be a root of this polynomial, so that the output from the rightmost cell is $a_n = Tr(A\alpha^{-n})$ where $A \in \mathbb{F}_q$ corresponds to the initial loading.

Rather than take the output from the rightmost cell, suppose we tap several cells, add the results, and output the sum. For example, we could tap the cell counting i from the right and also the cell counting j from the right (where $j > i$), thus outputting

$$b_n = a_{n+i} + a_{n+j}$$

as in Figure 6.4.

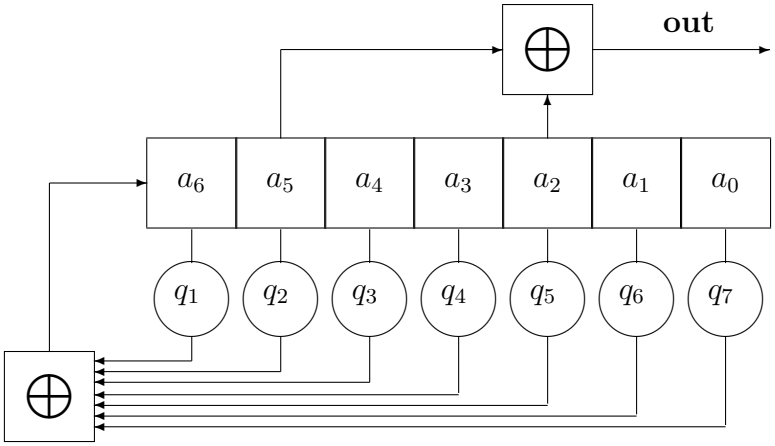


Figure 6.4: Phase taps: $b_n = a_{n+2} + a_{n+5}$

Show that the new sequence $b_0, b_1, \cdots$ is simply a shift (or "phase shift") of the sequence a by t steps (so $b_n = a_{n+t}$) where the shift t is determined by the equation

$$\alpha^t = 1 + \alpha^{j-i}.$$

2. If p, q are fixed integers, the linearly recurring sequence

$$a_n = pa_{n-1} + qa_{n-2}$$

is called a *generalized Lucas sequence*. The recurrence is called a *second order linear homogeneous recurrence*. Starting with $a_0 = 0$ and $a_1 = 1$ express the n-th element of this sequence as a sum of powers of roots of the characteristic polynomial. What happens if $p^2 + 4q = 0$?

3. A *perfect matching* of a finite graph G is a subset S of the edges so that every vertex lies on exactly one element of S. Let G_n denote the box-graph made out of n consecutive vertical struts, as in Figure 6.5. Prove that the number of perfect matchings of the graph G_n is the Fibonacci number F_{n+1} (so that the above graph G_5 has 8 perfect matchings).

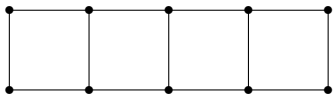


Figure 6.5: The graph G_5.

4. The squares of the Fibonacci numbers satisfy a linear recursion of degree 3. Find this recursion and express its generating function as a rational function.

5. Let R be a ring and let $q(x) = \sum_{i=1}^{m} q_i x^i - 1$ be a polynomial over R. Let A be the companion matrix of $q(x)$, as defined in equation (6.4). Prove that $\det(xA - I) = (-1)^r q(x)$. If the ring R is a field then prove that each root of the connection polynomial $q(x)$ is the inverse of an eigenvalue of A.

6. A sequence $a_0, a_1, \cdots$ of elements in a ring R satisfies an *mth order inhomogeneous linear recurrence* if there exist $q_1, q_2, \cdots, q_m \in R$ and $0 \neq c \in R$ so that for all $n \geq m$, the following holds:

$$a_n = q_1 a_{n-1} + \cdots + q_m a_{n-m} + c. \tag{6.26}$$

Suppose that $\sum_{i=1}^{m} q_i - 1$ divides c in R. Show that there is a $b \in R$ so that the sequence $a'_0, a'_1, \cdots$ with $a'_n = a_n - b$ satisfies the associated (homogeneous) linear recursion (6.26) with c replaced by 0.

7. If the sequence $a_0, a_1, \cdots$ satisfies the inhomogeneous linear recursion (6.26) of degree m, show that it also satisfies the (homogeneous) linear recursion with the following connection polynomial of degree $m + 1$,

$$Q(x) = -1 + \sum_{i=1}^{m+1} (q_i - q_{i-1}) x^i$$

where $q_0 = -1$ and $q_{m+1} = 0$.

8. Let N be a large fixed integer, possibly a power of 2. Fix A, B with $0 \leq A, B \leq N - 1$. The *linear congruence method* for generating (pseudo) random numbers modulo N uses a *seed* integer x_0 and the recurrence

$$x_n = A x_{n-1} + B \pmod{N}. \tag{6.27}$$

Prove: If N is prime and A is a primitive root modulo N then there is a single value of x_0 for which the output sequence is constant; for any other seed value the sequence will visit all the remaining $N - 1$ possible values before repeating.

9. Suppose $N = p^e$ $(e \geq 2)$ is a power of a prime p, and the sequence $x_0, x_1, \cdots$ is periodic and is generated by the linear congruence method (6.27). Show: if $A - 1$ is not divisible by p then the period of this sequence cannot exceed $N - N/p$. (In fact it is possible to achieve the full period, $N = p^e$, cf. [111], by taking $B \neq 0, 1$, taking $A - 1$ to be divisible by p and, if $p = 2$ then $A - 1$ should be divisible by 4.)

10. Let R be a finite local ring with maximal ideal $\mathfrak{m}$ with $\mathfrak{m}^k = (0)$ and quotient field $F = R/\mathfrak{m}$. Let $q(x) \in R[x]$ be a polynomial with exactly one coefficient that is a unit, say q_j.

a. Prove that if $q_i = 0$ when $i \neq j$, then $G(q) = \{\mathbf{0}\}$.
b. Suppose $j > 0$. Let $\mathbf{a} \in G(q)$. Prove that for every $l \geq 0$ we have $a_i \in \mathfrak{m}^l$ if $i \geq jl$.
c. Conclude that $a_i = 0$ if $i \geq jk$, and that $\mathbf{a} = \mathbf{0}$ if $q_i = 0$ for all $i < j$.
d. Show that if some $q_i \neq$ with $i < j$, then there is a nonzero sequence in $G(q)$.

Chapter 7 Feedback with Carry Shift Registers and Multiply with Carry Sequences

A *feedback with carry shift register* is a feedback shift register with a small amount of auxiliary memory. In its simplest form, the cells of the register consist of bits (0 or 1) while the memory contains a nonnegative integer. The contents (0 or 1) of the tapped cells of the shift register are added *as integers* to the current contents of the memory to form a sum σ. The parity bit, $\sigma \pmod 2$ of σ is fed back into the first cell, and the higher order bits, $\lfloor \sigma/2 \rfloor$ are retained for the new value of the memory. See Figure 7.1. There are many parallels between LFSR sequences and FCSR sequences, some of which we list in Table 7.1.

LFSR over $\mathbb{Z}/(2)$	FCSR over $\mathbb{Z}/(2)$
taps $q_1, \cdots, q_m \in \{0, 1\}$	taps $q_1, \cdots, q_m \in \{0, 1\}$
connection polynomial $q(x) = -1 + q_1 x + q_2 x^2 + \cdots + q_m x^m$	connection integer $q = -1 + q_1 2 + q_2 2^2 + \cdots + q_m 2^m$
output sequence $a_0, a_1, \cdots$ linearly recurring sequence	output sequence $a_0, a_1, \cdots$ multiply with carry sequence
generating function $a(x) = a_0 + a_1 x + a_2 x^2 + \cdots$	2-adic number $\alpha = a_0 + a_1 2 + a_2 2^2 + \cdots$
rational function: $a(x) = h(x)/q(x)$	rational number: $\alpha = h/q$
sum of two LFSR sequences	sum-with-carry of two FCSR sequences
exponential represesentation $a_i = Tr(A\alpha^i)$	exponential representation $a_i = A2^{-i} \pmod q \pmod 2$
maximal length m-sequence	maximal length ℓ-sequence
Berlekamp-Massey algorithm	rational approximation algorithm

Table 7.1: Comparison of LFSRs and FSRs.

The output sequences generated by an FCSR are examples of *multiply with carry* sequences. They enjoy many of the useful statistical properties of linearly recurrent sequences. As with linearly recurrent sequences, several algebraic structures are available for the analysis of multiply with carry sequences, including ordinary integer arithmetic, N-adic numbers, and an analog of the trace function. Multiply with carry sequences have been applied in such areas as pseudorandom number generation, cryptanalysis, and arithmetic codes.

Feedback-with-carry shift registers were first described in [102], [58] and [106] (whose publication was delayed for years by an unscrupulous editor). These devices were suggested as a method for high speed hardware generation of binary sequences with enormous periods, as might be used in a stream cipher or digital communication system. Around the same time the add-with-carry (AWC) generator was described [130], [132] as a method for the generation of (large) pseudo-random numbers for use in simulations and Monte Carlo integration. The AWC generator may be considered to be a single stage feedback-with-carry generator. It is a remarkable coincidence that the same basic idea occurred at roughly the same time to these two different research groups. In [131], [110], [29] the multistage generalization of the AWC generator was developed, where it was called the multiply-with-carry generator. Both lines of research continued to develop independently as their goals were somewhat different, but the article [62] places both points of view into a common setting.

7.1 Definitions

In this section we give the definitions and describe the basic properties of feedback with carry shift registers over the integers modulo N. Elements of $\mathbb{Z}/(N)$ will always be represented as integers between 0 and $N-1$. If u is an integer then $u \pmod N$ is the remainder after dividing u by N; the whole number quotient is denoted $u \text{ (div } N) = \lfloor u/N \rfloor$.

Definition 7.1.1. *Fix an integer $N > 1$. Let $S = \{0, 1, \cdots, N-1\}$. Let $q_1, q_2, \cdots, q_m \in S$. An* N-ary feedback with carry shift register of length m *with* multipliers *or* taps $q_1, \cdots, q_m$ *is a discrete state machine whose state is a collection*

$$(a_0, \cdots, a_{m-1}; z) \quad \textit{where } a_i \in S \quad \textit{and } z \in \mathbb{Z}$$

and whose state change operation is described as follows:

7.1.1.a *Compute the integer sum*

$$\sigma = \sum_{i=1}^{m} q_i a_{m-i} + z.$$

7.1.1.b *Replace $(a_0, a_1, \cdots, a_{m-1}; z)$ by $(a_1, a_2, \cdots, a_{m-1}, \sigma \pmod N; \sigma \text{ (div } N))$.*

Writing a_i' and z' for the new values of a_i and z, the change of state may thus be expressed by the equations $a_i' = a_{i+1}$ for $0 \le i \le m-2$ and

$$a_{m-1}' + Nz' = \sigma = q_1 a_{m-1} + q_2 a_{m-2} + \cdots + q_m a_0 + z.$$

It is convenient to think of a feedback with carry shift register (or FCSR) as a physical circuit, as pictured in Figure 7.1. Note the similarity to Figure 6.1. The contents of the tapped cells of the shift register are added *as integers* to the current contents of the memory to form a sum, σ. The residue, $\sigma \pmod N$ of σ is fed back into the first cell, and the higher order bits $\sigma \text{ (div } N) = \lfloor \sigma/N \rfloor$ are retained for the new value of the memory.

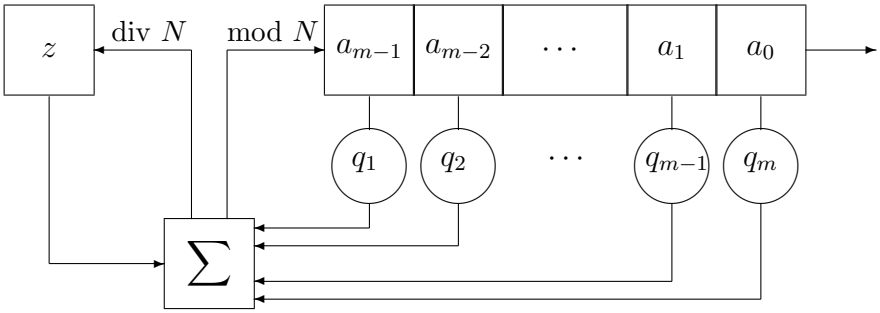


Figure 7.1: A Feedback with Carry Shift Register of Length m.

.

As with LFSRs, we like to think of bits as flowing out to the right, so the order of the components a_i in the diagram is the reverse of the order when we write the state as a vector. When thinking of FCSRs as physical devices as in the figure, we sometimes list the components of the state in descending order of their indices. We write $\boxed{z}\boxed{a_{m-1}}\boxed{a_{m-2}}\boxed{\cdots}\boxed{a_1}\boxed{a_0}$ for the *machine state* to distinguish the two notations.

The output sequence from an FCSR can also be described in more mathematical language using linear recurrences that incorporate a carry.

Definition 7.1.2. *Let $N > 1$ be an integer. A sequence $\mathbf{a} = a_0, a_1, \cdots$ of elements in $S = \{0, 1, \cdots, N-1\}$ is* linearly recurrent with carry mod N *or is a* multiply with carry sequence *if there exists a finite collection $q_1, \cdots, q_m \in S$ and an infinite sequence z_{m-1}, z_m, z_{m+1}, $\cdots$ of integers, or "memory values" such that for all $n \geq m$ we have*

$$a_n + Nz_n = q_1 a_{n-1} + \cdots + q_m a_{n-m} + z_{n-1}. \tag{7.1}$$

Equation (7.1) is called a recurrence relation with carry *and the integer m is called the* length *or* span *or* degree *of the recurrence. The integers $q_1, \cdots, q_m$ are called the* coefficients *of the recurrence. Set $q_0 = -1$. The* connection integer *of the recurrence is the integer*

$$q = -1 + \sum_{i=1}^{m} q_i N^i = \sum_{i=0}^{m} q_i N^i. \tag{7.2}$$

At first glance it may appear that the sequence z_i of memory values might grow without bound. But (cf. Proposition 7.6.1) if the sequence $\mathbf{a}$ is eventually periodic and satisfies the recurrence (7.1), then the memory values z_i will eventually enter the range $0 \le z_i \le w$ where $w = \sum_{i=1}^{m} q_i$, and will remain within this range thereafter, eventually becoming periodic as well.

Thus, for any initial set of values $a_0, \cdots, a_{m-1} \in \{0, \cdots, N-1\}$ and $z = z_{m-1} \in \mathbb{Z}$, the FCSR in Figure 7.1 generates a unique multiply with carry sequence mod N. In Section 7.7 we consider a slightly more general situation in which the multipliers q_i $(0 \le i \le m)$ are allowed to take on arbitrary positive or negative integer values.

7.2 Analysis of FCSRs

In this section we establish the relationship between N-adic numbers and the structure of an FCSR. Suppose we fix an m-stage FCSR whose connection integer is $q = -1 + q_1 N + q_2 N^2 + \cdots + q_m N^m$, and whose initial state is $(a_0, a_1, \cdots, a_{m-1}; z_{m-1})$. (See Figure 7.1.) The register will generate an infinite, eventually periodic sequence $\mathbf{a} = a_0, a_1, a_2, \cdots$ of integers modulo N, to which we associate the N-adic integer $a \in \mathbb{Z}_N$ such that $\mathbf{a} = \mathbf{seq}_N(a)$:

$$a = a_0 + a_1 N + a_2 N^2 + a_3 N^3 + \cdots \in \mathbb{Z}_N \tag{7.3}$$

which we call the *N-adic value* of the FCSR (with given initial state). Define

$$f = \sum_{i=0}^{m-1} \sum_{j=0}^{i} q_j a_{i-j} N^i - z_{m-1} N^m, \tag{7.4}$$

where we have set $q_0 = -1$ so that $q = \sum_{i=0}^{m} q_i N^i$.

Theorem 7.2.1. [106] *The output, $\mathbf{a}$, of an FCSR with connection integer $q > 0$, initial memory z_{m-1}, and initial loading $a_0, a_1, \cdots, a_{m-1}$, is the coefficient sequence of the N-adic representation of the rational number $a = f/q$. In other words, $\mathbf{a} = \mathbf{seq}_N(f/q)$ or*

$$a = \sum_{i=0}^{\infty} a_i N^i = \frac{f}{q} \in \mathbb{Z}_N. \tag{7.5}$$

Theorems 7.2.1 and 5.4.4 and Proposition 7.6.1 give the following:

Corollary 7.2.2. *If $\mathbf{a} = a_0, a_1, a_2, \cdots$ is an eventually periodic N-ary sequence then the associated N-adic number $a = \sum a_i N^i$ is a quotient of two integers, $a = f/q$ and the denominator q is the connection integer of a FCSR which generates the sequence $\mathbf{a}$. The sequence $\mathbf{a}$ is strictly periodic if and only if $-q \le f \le 0$. In this case the values of the memory may be taken to lie in the range*

$$0 \le z < \sum_{i=1}^{m} q_i.$$

Proof of Theorem 7.2.1. The computations in this proof parallel those in Theorem 6.4.1. Let us consider the transition from one state of the shift register to the next. Suppose that the state of the register is $(a_{n-m}, a_{n-m+1}, \cdots, a_{n-1}; z_{n-1})$. The next state is determined by calculating

$$\sigma_n = z_{n-1} + \sum_{i=1}^{m} q_i a_{n-i}, \tag{7.6}$$

writing the new memory contents as $z_n = \lfloor \sigma_n / N \rfloor = \sigma_n \pmod{N}$, and writing the new contents of the leftmost cell as $a_n = \sigma_n \pmod{N}$. These equations may be combined into the expression

$$\sigma_n = N z_n + a_n.$$

It follows that

$$a_n = \sum_{i=1}^{m} q_i a_{n-i} + (z_{n-1} - N z_n), \tag{7.7}$$

provided that $n \geq m$. Suppose the initial loading of the register consists of memory z_{m-1} and with register bit values $a_0, a_1, \cdots, a_{m-2}, a_{m-1}$. Now substitute equation (7.7) into the expression (7.5) for a to obtain

$$\begin{aligned} a &= a_0 + a_1 N + \cdots + a_{m-1} N^{m-1} + \sum_{n=m}^{\infty} a_n N^n \\ &= s + \sum_{n=m}^{\infty} \left(\sum_{i=1}^{m} q_i a_{n-i} \right) N^n + \sum_{n=m}^{\infty} (z_{n-1} - N z_n) N^n, \end{aligned} \tag{7.8}$$

where

$$s = a_0 + a_1 N + \cdots + a_{m-1} N^{m-1}$$

is the integer represented by the initial loading of the register. The second summation in equation (7.8) cancels except for the first term, z_{m-1}, leaving

$$\begin{aligned} a &= s + z_{m-1} N^m + \sum_{n=m}^{\infty} \sum_{i=1}^{m} q_i N^i a_{n-i} N^{n-i} \\ &= s + z_{m-1} N^m + \sum_{i=1}^{m} q_i N^i \Big(\sum_{n=m}^{\infty} a_{n-i} N^{n-i} \Big) \\ &= s + z_{m-1} N^m + \sum_{i=1}^{m} q_i N^i (a - (a_0 N^0 + a_1 N^1 + \cdots + a_{m-i-1} N^{m-i-1})) \\ &= s + z_{m-1} N^m + a \sum_{i=1}^{m} q_i N^i - \sum_{i=1}^{m-1} \sum_{j=0}^{m-i-1} q_i N^i a_j N^j \end{aligned}$$

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

$$\begin{aligned} a &= \frac{s + z_{m-1}N^m - \sum_{i=1}^{m-1}\sum_{j=0}^{m-i-1} q_i N^i a_j N^j}{1 - \sum_{i=1}^{m} q_i N^i} \\ &= \frac{\sum_{i=0}^{m-1}\sum_{j=0}^{m-i-1} q_i a_j N^{i+j} - z_{m-1}N^m}{q} \end{aligned} \tag{7.9}$$

since $q_0 = -1$. The double summation is over all pairs of integers $0 \le i, j \le m-1$ with $i+j \le m-1$. Setting $k = i + j$ gives

$$a = \frac{\sum_{k=0}^{m-1}\sum_{i=0}^{k} q_i a_{k-i} N^k - z_{m-1}N^m}{q} = \frac{f}{q}. \tag{7.10}$$

This completes the proof of Theorem 7.2.1. □

Corollary 7.2.3. [106] *Changing the initial memory by b changes the value of a by $-bN^m/q$. If $a = f/q < 0$ then the initial memory $z_{m-1} \ge 0$ is nonnegative.*

Remarks. There are three easy initial loadings which guarantee a strictly periodic output:

1. let $z_{m-1} = 1$ and all the $a_j = 0$ (so $f = -N^m$);
2. let $1 \le z_{m-1} \le q_m$, $a_0 = 1$, and all the other $a_j = 0$, (so $f = q - (q_m + z_{m-1})N^m$);
3. let $z_{m-1} = 0$, $a_j = 0$ if $0 \le j \le m-2$, and a_{m-1} be arbitrary (so $f = -a_{m-1}N^{m-1}$).

More generally, the sequence is strictly periodic if and only if

$$\frac{\sum_{k=0}^{m-1}\sum_{i=0}^{k} q_i a_{k-i} N^k}{N^m} \le z_{m-1} \le \frac{q + \sum_{k=0}^{m-1}\sum_{i=0}^{k} q_i a_{k-i} N^k}{N^m},$$

a condition that is readily checked for a given initial state $(a_0, a_1, \cdots, a_{m-1}; z_{m-1})$.

If $-q < f < 0$ and f is relatively prime to q, then by Corollary 5.4.5, the period of the sequence is $T = \mathrm{ord}_q(N)$. If f and q have a common factor, then the period is a divisor of $\mathrm{ord}_q(N)$. If $f \ge 0$ or $f \le -q$ then the sequence has a transient prefix before it drops into a periodic state.

Let $w = \sum_{i=1}^{m} q_i$. In the sequel we will use the following estimates.

Lemma 7.2.4. *Suppose $q = -1 + q_1 N + \cdots + q_m N^m$ with $q_m \ne 0$. Let $z \in \mathbb{Z}$ be an integer. Let $a_i \in \{0, 1\}$ for $0 \le i \le m-1$ and set $s = \sum_{i=0}^{m-1} a_i N^i$. As in equations (7.4) and (7.9) define*

$$f = \sum_{i=1}^{m-1}\sum_{j=0}^{m-i-1} q_i a_j N^{i+j} - s - zN^m. \tag{7.11}$$

Then the double sum is bounded:

$$\sum_{i=1}^{m-1} \sum_{j=0}^{m-i-1} q_i a_j N^{i+j} \leq (w-1)N^m.$$

Furthermore,

1. *if $f > 0$ then $z \leq w - q_m - 1$ and*

$$(-z-1)N^m < f \leq (w - z - q_m)N^m; \tag{7.12}$$

2. *if $f < 0$ then $z \geq 0$ and*

$$\max(0, (z - w + q_m)N^m) \leq |f| < (z+1)N^m. \tag{7.13}$$

Proof. The proof is a direct calculation:

$$\sum_{i=1}^{m-1} q_i N^i \sum_{j=0}^{m-i-1} a_j N^j \leq \sum_{i=1}^{m-1} q_i N^i (N^{m-i} - 1) \leq \sum_{i=1}^{m-1} q_i N^m = (w - q_m)N^m$$

since q_m does not appear in the sum. The other estimates follow from this and the fact that $0 \leq s < N^m$. □

7.3 Initial loading

In this section we answer the reverse question: Suppose we are given a fraction $a = f/q$ (with q an odd positive integer), how do we determine a FCSR and initial loading whose output sequence coincides with the N-adic expansion of a? Set $m = \lfloor \log_N(q+1) \rfloor$. Write $q = \sum_{i=0}^{m} q_i N^i$ with $q_0 = -1$ and $q_i \in \{0, 1\}$ for $i > 0$. Consider an FCSR with m stages and with connection integer q. The initial memory z_{m-1} and initial loading $a_0, a_1, \ldots, a_{m-1}$ are related to f and q by equation (7.10) which may be solved using the following procedure:

Procedure 7.3.1.

7.3.1.a *Compute $a_0 + a_1 N + \cdots + a_{m-1}N^{m-1} = f/q \pmod{N^m}$. (These are the first m symbols in the N-adic expansion for f/q.)*

7.3.1.b *Compute*

$$y = \sum_{i=0}^{m-1} \sum_{j=0}^{i} q_j a_{i-j} N^i.$$

7.3.1.c *Compute* $z = (y - p)/N^m$.

Use $a_0, \cdots, a_{m-1}$ as the initial loading and z as the initial memory in a FCSR with connection integer q. This FCSR will output the N-adic expansion of f/q.

Step 7.3.1.a can be carried out in time $\mathcal{O}(m^2)$. We do so by iteratively finding the a_i. Assume integers are represented in base N. We must find $a_0, a_1, \cdots, a_{m-1}$ so that

$$f \equiv q \sum_{i=0}^{m-1} a_i N^i \pmod{N^m},$$

since q is relatively prime to N. At the jth stage we consider this congruence modulo N^{j+1}. For a_0, the 0th stage gives $f \equiv q_0 a_0 = -a_0 \pmod{N}$, which uniquely determines a_0. More generally, assume that we have computed $a_0, \cdots, a_{j-1}$ and $\gamma = q \sum_{i=0}^{j-1} a_i N^i$. Then we want a_j so that

$$\begin{aligned} f &\equiv q \sum_{i=0}^{j} a_i N^i \pmod{N^{j+1}} \\ &\equiv \gamma + q_0 a_j \pmod{N^{j+1}} \\ &= \gamma - a_j N^j, \end{aligned}$$

which again uniquely determines $a_j \equiv (\gamma - f)/N^j \pmod{N}$. We then add $q a_j N^j$ to γ. Addition, subtraction, and the multiplication $q a_j$ can be carried out in time complexity $\mathcal{O}(m)$, so the overall time complexity of step 7.3.1.a is $\mathcal{O}(m^2)$.

Step 7.3.1.b can be carried out by multiplying the polynomials

$$q(x) = \sum_{i=0}^{m-1} q_i x^i \quad \text{and} \quad a(x) = \sum_{i=0}^{m-1} a_i x^i,$$

and evaluating the m lowest degree terms at $x = N$. Step 7.3.1.c can be carried out in time $\mathcal{O}(m)$.

Degenerate initial loadings.

An initial loading is *degenerate* if the N-adic number $a = f/q$ corresponding to the output sequence is an ordinary integer. In this case, after a transient prefix, the FCSR outputs all 0s (if $a \geq 0$) or all $(N-1)$s (if $a < 0$). There are only two possible degenerate final states: (a) $z = 0$ and all $a_i = 0$ and (b) $z = \sum_{i=1}^{m} q_i - 1$ and all $a_i = N - 1$. How long can the prefix be?

Theorem 7.3.2. *Consider an m-stage FCSR with $q_m \neq 0$. Suppose the initial loading is degenerate. If the initial memory $z > 0$ is positive then the output will stabilize to all $(N-1)$s after no more than*

$$\lceil \log_N(1 + z) \rceil$$

steps . If the initial memory $z < 0$ and the register has at least two nonzero q_js, then the output will stabilize to all 0s within

$$\lceil \log_N(|z| + w - q_m) \rceil$$

steps, where $w = \sum_{i=1}^{m} q_i$. If the initial memory $z < 0$ and $q_j = 0$ for $1 \leq j \leq m-1$, then the output will stabilize to all 0's within

$$\lceil \log_N(|z|N^m/(N^m - 1)) \rceil$$

steps. If the initial memory $z = 0$ then the only degenerate initial loading is the trivial one consisting of all 0's.

Proof. Suppose the value $a = f/q$ of the FCSR is an integer. If $a = 0$, then we must have $z = 0$ and all $a_j = 0$, so assume $a \neq 0$. First we consider the case that the initial memory is $z > 0$. Let us consider the possibilities $a > 0$ and $a < 0$ separately.

If $a > 0$ is an integer, eventually the FCSR will output all 0s. However, this is not possible: there must be a last state $\mathbf{s}_1$ after which the memory remains zero (otherwise, once the register has become all zero, the memory would eventually feed a nonzero value back into the register), and there must be a last state $\mathbf{s}_2$ in which some $a_j \neq 0$. Then $j = 0$. If $\mathbf{s}_1$ precedes $\mathbf{s}_2$ or $\mathbf{s}_1 = \mathbf{s}_2$ then the next state after $\mathbf{s}_2$ has $a_{m-1} \neq 0$, a contradiction. If $\mathbf{s}_2$ precedes $\mathbf{s}_1$, and $z = N^k b$ with $N \not| b$, then after k steps a_{m-1} will equal $b \pmod N \neq 0$, also a contradiction.

If $a < 0$ is an integer, then eventually the FCSR will output all $(N-1)$s. Since $z \geq 0$ and $f < 0$ we have by equation (7.9),

$$|f| \leq s + zN^m \leq (1+z)N^m,$$

where $s = a_0 + a_1 N + \ldots + a_{m-1}N^{m-1}$ as in equation (7.8). If $q \neq N^m - 1$ then $q > N^m$ and we conclude that $|a| < 1 + z$. The output, which is the N-adic expansion of the integer a becomes all 1's within $\lceil \log_N(1+z) \rceil$ steps by equation (5.9). A special argument, which we omit, must be made when $q = N^m - 1$.

Now consider the case of initial memory $z < 0$. We claim that the output string cannot degenerate to all $(N-1)$s: if so, then $a < 0$ so $f < 0$. However, by equation (7.9) the only negative contribution to f is from s, and $s < q$. So if $f < 0$ then $|a| \leq f/q < 1$ which therefore cannot be an integer (other than 0).

Finally, if the initial memory $z < 0$ and if the output stream degenerates to all 0's then $a > 0$ and $f > 0$. By. Lemma 7.2.4

$$f \leq (|z| + w - q_m)N^m. \tag{7.14}$$

If $q \neq N^m - 1$, then we have $q > N^m$, so $a < |z| + w - q_m$. So in this case, the output sequence is the N-ary expansion of a, which takes $\lceil \log_N(|z| + w - q_m) \rceil$ bits, after which we have all 0's. In the case of a single tap $q = N^m - 1$, so $a = |z|N^m/(N^m - 1)$. □

7.4 Representation of FCSR sequences

One of the most powerful techniques for the analysis of shift register sequences is its exponential representation. Suppose N is prime and $\mathbf{a} = (a_0, a_1, a_2, \cdots)$ is a periodic sequence of elements of $\mathbb{F}_N = \{0, 1, \cdots, N-1\}$ obtained from a linear feedback shift register of length m, with connection polynomial $q(X)$. If $q(X)$ is irreducible and if $\gamma \in \mathbb{F}_{N^m}$ is a root of $q(X)$ in the finite field with N^m elements, then as in Section 6.6.bfor all $i = 0, 1, 2, \cdots$ we have,

$$a_i = T(A\gamma^i)$$

for some $A \in \mathbb{F}_{N^m}$ (which corresponds to the choice of initial loading of the shift register). Here, $T : \mathbb{F}_{N^m} \to \mathbb{F}_N$ denotes any nontrivial $\mathbb{F}_N$ linear function such as the trace function. In this section we derive a similar representation for periodic sequences of bits obtained from feedback shift registers with memory.

Now let N be any integer greater than 2. Let $0 < N < q$ with N and q relatively prime. If $z \in \mathbb{Z}$ or $z \in \mathbb{Z}/(q)$, then we use the notation $z \pmod q \pmod N$ to mean that first the number z should be reduced modulo q to give an integer between 0 and $q-1$, and then that number should be reduced modulo N to give an element of $\mathbb{Z}/(N)$. (Notice that there is no group homomorphism $\mathbb{Z}/(q) \to \mathbb{Z}/(N)$ if q is odd, so the notation $z \pmod q \pmod N$ needs a precise definition.) Since N and q are relatively prime there is an element $\gamma \in \mathbb{Z}/(q)$ so that $\gamma N \equiv 1 \pmod q$ and we write $\gamma = N^{-1} \pmod q$. Similarly if $h \in \mathbb{Z}$, we write $N^{-i}h \pmod q \pmod N$ to mean that first h is reduced modulo q, then it is multiplied by $\gamma^i \pmod q$; the result is represented as an integer between 0 and $q-1$ and then that integer is reduced modulo N.

We will need the following fact which is an analog of Proposition 6.6.1

Proposition 7.4.1. [106] *Fix $N \geq 2$. Let $q = -1 + q_1 N + \cdots + q_m N^m$ where $0 \leq q_i < N$. Then q is invertible in the ring $\mathbb{Z}_N$. Let h be an integer with $0 \leq h \leq q$. Let*

$$-\frac{h}{q} = b_0 + b_1 N + b_2 N^2 + \cdots \tag{7.15}$$

be the N-adic expansion of the fraction $-h/q$. Then for all $i \geq 1$,

$$b_i = N^{-i}h \pmod q \pmod N.$$

Proof. First observe that, given h, q satisfying equation (7.15) we have

1. $b_0 = h \pmod N$ and
2. $h + qb_0$ is divisible by N *as an integer.*

To see this, write $h = h_0 + h_1 N + \cdots + h_{m-1}N^{m-1}$ and cross multiply in equation (7.15) to get

$$-(h_0 + h_1 N + \cdots + h_{m-1}N^{m-1}) = (-1 + q_0 N + \cdots + q_m N^M)(b_0 + b_1 N + \cdots)$$

so $h_0 = b_0$ which proves (1). For (2), multiply equation (7.15) by q to obtain

$$h + qb_0 = -Nq(b_1 + b_2 N + b_3 N^2 + \cdots). \tag{7.16}$$

So $h + qb_0$ is divisible by N in the N-adic numbers, which implies that it is divisible by N in the integers.

It follows from equation (7.16) that the N-adic number $b_1 + b_2 N + \cdots$ is equal to the fraction $-h'/q$ where

$$h' = (h + qb_0)/N. \tag{7.17}$$

So applying item (1) again we conclude that $b_1 = h' \pmod N$.

Now reduce equation (7.17) modulo q to conclude that $h' = N^{-1}h \pmod q$. In summary, $b_1 = N^{-1}h \pmod q \pmod N$. Moreover, we may continue in this way by induction, replacing h with h' to obtain $b_n = h^{(n)} \pmod N$ where $h^{(n)} = N^{-n}h \pmod q$. □

Together with Theorem 7.2.1, this proposition immediately gives an "exponential" representation for the output sequence of an FCSR, as follows.

Theorem 7.4.2. [106] *Let* $\mathbf{a} = (a_0, a_1, a_2, \cdots)$ *be the output sequence of an FCSR with entries in* $\mathbb{Z}/(N)$ *and with connection integer* q. *Then* $\mathbf{a}$ *is eventually periodic and it coincides with the coefficient sequence* $\mathbf{seq}_N(-h/q)$ *of the* N*-adic expansion of the fraction* $-h/q$, *where* h *is determined by the initial loading via equation (7.4). If* $\mathbf{a}$ *is strictly periodic then for all* i,

$$a_i = N^{-i}h \pmod q \pmod N.$$

Although Peterson and Weldon [159] consider only the case where q is prime and N is a primitive element modulo q, their proof of Theorem 15.5 (p. 458) may be used in this situation to give another proof of Theorem 7.4.2. The proof presented here is useful because, as we see later, it extends to AFSRs.

Corollary 7.4.3. [106] *An FCSR sequence with connection integer* q *is eventually periodic with period dividing the multiplicative order of* N *modulo* q. *If* q *is the least connection integer, then the period equals the multiplicative order of* N *modulo* q.

Proof. Suppose $\mathbf{a} = a_0, a_1, \cdots$ is eventually periodic, and

$$a = \sum_{i=0}^{\infty} a_i N^i = \frac{f}{q}.$$

Then **a** is periodic from some point m, and $a = c + N^m g/q$ for some integers c and g with the coefficient sequence of g/q periodic. The period of g/q equals the eventual period of **a**, so it suffices to consider the case when **a** is strictly periodic. Then by Theorem 7.4.2 the period of **a** is a divisor of the order of N modulo q. The equality in the second case follows from Theorem 5.4.4. □

7.5 Example: $q = 37$

Let us consider the 2-adic FCSR with connection integer $q = 37 = 32 + 4 + 2 - 1$. Then we have a 5 stage shift register with feedback connections on the first, second, and fifth cells, counting from the left. The element $\gamma = 2^{-1} \in \mathbb{Z}/(37)$ is $\gamma = 19$. In Table 7.2 we consider the initial state $\boxed{0}\boxed{1}\boxed{0}\boxed{0}\boxed{1}\boxed{1}$ so that the output sequence is given by

$$a_n = \gamma^n \pmod{37} \pmod 2$$

for $n = 0, 1, 2, \cdots$, i.e. with the constant $A = 1$, in the notation of the preceding section. The index

mem	register	a_0	f	n
0	10011	1	1	0
1	01001	1	19	1
1	10100	0	28	2
1	01010	0	14	3
1	00101	1	7	4
1	00010	0	22	5
0	10001	1	11	6
1	01000	0	24	7
1	00100	0	12	8
0	10010	0	6	9
0	11001	1	3	10
1	11100	0	20	11
1	11110	0	10	12
1	11111	1	5	13
2	01111	1	21	14
2	00111	1	29	15
1	10011	1	33	16
1	11001	1	35	17

mem	register	a_0	f	n
2	01100	0	36	18
1	10110	0	18	19
1	01011	1	9	20
1	10101	1	23	21
1	11010	0	30	22
1	11101	1	15	23
2	01110	0	26	24
1	10111	1	13	25
1	11011	1	25	26
2	01101	1	31	27
2	00110	0	34	28
1	00011	1	17	29
1	00001	1	27	30
1	00000	0	32	31
0	10000	0	16	32
0	11000	0	8	33
1	01100	0	4	34
1	00110	0	2	35

Table 7.2: The states of a 2-adic FCSR with $q = 37$.

n is recorded as the last column of Table 7.2. The column "mem" indicates the integer value of

the memory, and a_0 represents the output bit (i.e. the rightmost bit in the register). Each state S of the shift register corresponds to a rational number $r(S) = -f/37$ and the numerator f is recorded also in the table. The table therefore lists all the strictly periodic states of the FCSR.

7.6 Memory requirements

Unlike LFSRs, it is not immediately apparent that an FCSR is a finite state device. It is a priori possible that the memory grows unboundedly through an infinite execution of the FCSR (and we shall see exactly this phenomenon for a more general class of sequence generators in Theorem 8.3.1). In this section we show the memory of an FCSR remains bounded.

Let us consider an m-stage FCSR with positive connection integer $q = -1 + q_1 N + \ldots + q_m N^m$. Let $w = \sum_{i=1}^{m} q_i$. A state $(a_0, a_1, \cdots, a_{m-1}; z_m)$ is *periodic* if, left to run, the FCSR will eventually return to that same state.

Assume that the state of the FCSR after $n - m$ iterations is $(a_{n-m}, a_{n-m+1}, \cdots, a_{n-1}; z_n)$. Let

$$\sigma_n = q_1 a_{n-1} + q_2 a_{n-2} + \cdots + q_m a_{n-m} + z_n.$$

Proposition 7.6.1. [106] *If an FCSR is in a periodic state then the memory is in the range $0 \leq z < w$ (which may therefore be accomplished by using no more than $\lfloor \log_2(w-1) \rfloor + 1$ bits of memory). If the initial memory $z_n \geq w$, then it will monotonically decrease and will arrive in the range $0 \leq z < w$ within $\lfloor \log_N(z_n - w) \rfloor + m$ steps. If the initial memory $z_n < 0$, then it will monotonically increase and will arrive in the range $0 \leq z < w$ within $\lceil \log_N |z_n| \rceil + m$ steps.*

Proof. First, observe that if the initial memory value z_n lies in the range $0 \leq z_n < w$ then the same will be true for all later values of the memory, since in this case $\sigma_n \leq (N-1)w + z_n < Nw$.

By the same argument, if the initial memory value is $z_n = w$, then later values of memory will be no greater than w, but in this case within m steps the memory will drop below w (and will remain so thereafter) for the following reason. If the memory does not decrease (i.e. if $z_{n+1} = w$), then this means that an $N - 1$ appeared in *all* the tapped cells, that $\sigma_n = Nw$ and that $a_n = \sigma_n \pmod N = 0$ was fed into the register. The value of σ will fall below Nw when this 0 reaches the first tapped cell (if not before), at which time we will have $z = \lfloor \sigma/N \rfloor < w$.

Moreover, if we initialize a FCSR with a larger memory value, $z_n > w$, then with each step, the excess $e_n = z_n - w$ will become reduced by a factor of $1/N$. That is,

$$e_{n+1} \leq \left\lfloor \frac{e_{n-1}}{N} \right\rfloor.$$

So after $\lfloor \log_N(z_N - w) \rfloor + 1$ steps, the memory will be no more than w.

Now we consider the case of negative initial memory, $z_n < 0$. It is possible that $\sigma_n \geq 0$, in which case the next memory value will be $z_{n+1} \geq 0$ (where it will remain thereafter). So let us

suppose that $\sigma_n < 0$. Then $|z_{n+1}| \le (|\sigma_n| + N - 1)/N \le (|z_n| + N - 1)/N$. Iterating this formula, we find that after $k = \lceil \log_N |z_n| \rceil$ steps, either the memory z has become nonnegative, or else

$$|z| \le \frac{z_n}{N^k} + \frac{1}{N^k} + \frac{1}{N^{k-1}} + \cdots + \frac{1}{N} < \frac{N}{N-1}$$

in which case the memory must be $z = -1$. There is a single situation in which the memory can remain at -1 forever: if there are no feedback taps on the shift register (so $q = -1$). In this case the memory will feed $(N-1)$'s into the shift register forever. However, we assumed $q > 0$ to rule out this possibility. If $q > 0$, then as soon as a nonzero feedback occurs, the memory will become nonnegative, where it will remain thereafter. □

7.7 Random number generation using MWC

A pseudo random number generator (RNG) for high speed simulation and Monte Carlo integration should have several properties: (1) it should have enormous period, (2) it should exhibit uniform distribution of d-tuples (for a large range of d), (3) it should exhibit a good structure (usually a lattice structure) in high dimensions, and (4) it should be efficiently computable (preferably with a base b which is a power of 2). Generators with these properties are surprisingly rare. In the words of R. Coveyou, "The generation of random numbers is much too important to be left to chance."

Marsaglia and Zaman [132] showed that their add-with-carry (AWC) generators satisfy condition (1). An AWC amounts to an FCSR for which all coefficients q_i are 0 or 1, and exactly two of them are 1. By giving up on (4) and using an appropriate base b, they achieve good distribution properties of d-tuples for values d which are less than the "lag." Tezuka et al [183] showed that these generators fail the spectral test([30], [111]) for large d; cf. [29]. The MWC generator was proposed as a modification of the AWC generator which satisfies both conditions (1) and (4). They are essentially FCSR generators for which N is a power of 2 and the q_i are allowed to be negative. However the distributional properties (2) of MWC sequences are not optimal, and in fact they are rather difficult to determine. See [29], where estimates on the distribution of d-tuples are derived (using some sophisticated techniques from number theory).

In this section we summarize results from [62] in which the MWC and FCSR generators are modified slightly so as to allow a coefficient q_0 other than -1 and to allow for the possibility of negative multipliers q_i. In this way it is possible to construct pseudo-random number generators that exhibit properties (1), (2), and (4). It is expected that for large d the distribution of d-tuples will suffer from the same shortcomings as those described in [183] and [29].

7.7.a MWC generators

Throughout the rest of this section we fix an integer $N \geq 2$; the output of the generator will be a sequence of elements in $\mathbb{Z}/(N)$ which we also identify with the set $S = \{0, 1, \cdots, N-1\}$. Fix integer multipliers $q_0, q_1, \cdots, q_m$ such that q_0 is relatively prime to N. (In many applications N will be a power of 2 in which case this condition simply means that q_0 is odd. The q_i are allowed to be negative.) A state of the generator consists of cell values $a_0, a_1, \cdots, a_{m-1} \in \mathbb{Z}/(N)$ and a memory integer $z \in \mathbb{Z}$.

The change of state proceeds as with the FCSR: compute the integer sum $\sigma = \sum_{i=1}^{m} q_i a_{m-i} + z$. The new state $(a'_0, a'_1, \cdots, a'_{m-1}; z')$ is defined by $a'_i = a_{i+1}$ for $0 \leq i \leq m-2$ while $a'_{m-1} \in S$ and $z' \in \mathbb{Z}$ are the unique numbers such that

$$-q_0 a'_{m-1} + z'N = \sigma = \sum_{i=1}^{m} q_i a_{m-i} + z. \tag{7.18}$$

These values are found as follows. Calculate, once and for all,

$$A = -q_0^{-1} \pmod{N}$$

and realize this as an integer between 1 and $N-1$. Then

$$a'_{m-1} = (A\sigma) \pmod{N} \quad \text{and} \quad z' = \sigma \text{ (div } N) = \lfloor \sigma/N \rfloor .$$

This amounts to inserting the multiplier A in the "mod N" feedback line of Figure 7.1. Since $\mathbb{Z} \to \mathbb{Z}/(N)$ is a ring homomorphism, we may also write $a'_{m-1} = A(\sigma \pmod{N})$.

The connection integer $q = \sum_{i=0}^{m} q_i N^i$ can now take on any value that is relatively prime to N. As originally defined in [132], [131] and [29], the coefficient q_0 was equal to -1. If the base N is chosen to be a power of 2 then these generators admit efficient implementations, however the connection integer will be constrained to be of the form $q = Nc - 1$ for some integer c. In this case, N is never a primitive root modulo q so the generator will never have maximal period. A similar criticism applied to the subtract-with-borrow generator. The introduction of a nontrivial value for q_0 (as first described in [107]) comes at the cost of one more multiplication per round but it has the benefit that the connection integer q may be chosen so that N is primitive modulo q and this leads to properties (1), (2), and (4) listed above.

7.7.b Periodic states

For a given state $(a_0, a_1, \cdots, a_{m-1}; z)$ of the generator as in equation (7.4) set

$$f = \sum_{i=0}^{m-1} \sum_{j=0}^{i} q_j a_{i-j} N^i - zN^m.$$

Theorem 7.7.1. [106] *The output sequence $a_0, a_1, \cdots$ of the generator satisfies*

$$\frac{f}{q} = a_0 + a_1 N + a_2 N^2 + \cdots \in \mathbb{Z}_N.$$

This sequence is strictly periodic if and only if $-q \le f \le 0$. In this case,

$$a_i = AN^{-i} f \pmod{q} \pmod{N}.$$

The proofs are exactly the same as the proofs of Theorem 7.2.1, Corollary 7.2.2, and Theorem 7.4.2. For convenience we restate Corollary 16.2.1 in this setting.

Theorem 7.7.2. [106] *If N is a primitive root modulo q, then from any initial periodic state the generator will visit all $q-1$ periodic states before returning to the initial state. In this case the output sequence is an ℓ-sequence. (See Chapter 16.) Fix $d \ge 1$ and let $\mathbf{z} = (z_1, z_2, \cdots, z_d)$ with $0 \le z_i \le N$. Then the number $M(\mathbf{z})$ of occurrences of the d-tuple $\mathbf{z}$ which begin in any fixed period of the output sequence can vary at most by 1, so $M(\mathbf{z})$ is either*

$$\left\lfloor \frac{q-1}{N^d} \right\rfloor \quad \text{or} \quad \left\lfloor \frac{q-1}{N^d} \right\rfloor + 1.$$

(See Section *16.2.)*

7.7.c Memory requirements

The memory of the MWC generator is well behaved, but the introduction of (possible) negative multipliers changes the memory analysis slightly. First we need to eliminate two classes of multipliers. Let us say the MWC generator is *extremal* if either (a) $q_0 < 0$ and all remaining $q_i \le 0$ or (b) $q_0 > 0$ and all remaining $q_i \ge 0$. Define

$$w^+ = \sum_{\substack{q_i > 0 \\ 0 \le i \le m}} q_i \quad \text{and} \quad w^- = \sum_{\substack{q_i < 0 \\ 0 \le i \le m}} q_i.$$

Theorem 7.7.3. [106] *Suppose the MWC generator is not extremal. If it is in a (strictly) periodic state then the memory z lies in the range*

$$w^- < z < w^+.$$

If $z \ge w^+$ then it will drop monotonically and exponentially fast until it lies within this range and it will remain within this range thereafter. If $z \le w^-$ then it will rise monotonically and exponentially fast until it lies within this range and it will remain within this range thereafter. If the generator is extremal then z will move montonically until it lies within the range $w^- \le z \le w^+$ and it will remain within this range thereafter.

Proof. Let us assume $q_0 < 0$. (The proof for $q_0 > 0$ is completely parallel.) Since $0 \le a_i \le N-1$, from equation (7.18) we have

$$z' = \frac{1}{N}\left\lfloor \sum_{i=0}^{m} q_i a_{m-i} + z \right\rfloor \le \left(\frac{N-1}{N}\right) w^+ + \frac{z}{N}, \tag{7.19}$$

where we have written a_m in place of a'_{m-1}. If $z < w^+$ this gives $z' < w^+$. If $z = w^+$ this gives $z' \le w^+$. If $z > w^+$ this gives

$$z' - z \le (w^+ - z)\left(\frac{N-1}{N}\right) < 0,$$

hence the memory decreases monotonically. Moreover, if $z > 0$ then $z' - w^+ \le (z - w^+)/N$, which is to say that $z - w^+$ decreases exponentially.

There are no strictly periodic states with $z = w^+$ unless the generator is extremal. For if $z = z' = w^+$, then equation (7.18) gives

$$(N-1)w^+ = \sum_{i=0}^{m} q_i a_{m-i}.$$

where we have written $a_m = a'_{m-1}$. The right side of this equation achieves its maximum value, $(N-1)w^+$, when $a'_{m-1} = 0$ and

$$a_i = \begin{cases} 0 & \text{whenever } q_i < 0 \\ N-1 & \text{whenever } q_i > 0. \end{cases}$$

Eventually this $0 = a'_{m-1}$ will get shifted into one of the positions where $q_i > 0$ and then the value of z will drop below w^+. (If $q_i \le 0$ for all i, then the generator is extremal, $w^+ = 0$, and the degenerate "bottom" (all-zero) state satisfies $z = w^+$.) In summary, if the generator is not extremal and if the memory starts out at any positive value, it will drop until $z < w^+$ and will remain there forever.

To obtain the lower bound on z, equation (7.18) gives

$$z' = \frac{1}{N}\left\lfloor \sum_{i=o}^{m} q_i a_{m-i} + z \right\rfloor \ge \frac{N-1}{N} w^- + \frac{z}{N}.$$

If $z > w^-$, then $z' > w^-$. If $z = w^-$ then $z' \ge w^-$. If $z < w^-$, then

$$z' - w^- \ge \frac{N-1}{N}(z - w^-)$$

so the value of $z - w^-$ will increase monotonically (and exponentially) until it is nonnegative. Let us examine the possible periodic states with $z = z' = w^-$. For such a state, equation (7.18) gives

$$(N-1)w^- = \sum_{i=0}^{m} q_i a_{m-i}.$$

The right side of this equation achieves its minimum value, $(N-1)w^-$, when $a'_{m-1} = N-1$ and

$$a_i = \begin{cases} N-1 & \text{whenever } q_i < 0 \\ 0 & \text{whenever } q_i > 0. \end{cases}$$

If some coefficient q_i is positive (which is to say, if the generator is not extremal) then this $N-1 = a'_{m-1}$ will eventually be shifted into the ith position, and the value of z will rise above w^-. However, if the generator is extremal (that is, if $q_i \leq 0$ for $1 \leq i \leq m$) then this argument fails and indeed, the degenerate "top" state satisfies $z = w^-$ and $a_i = N-1$ for all i. In summary, if the generator is not extremal and if the memory starts out at some negative value, then it will rise until $z > w^-$ and it will remain there forever. □

7.7.d Finding good multipliers

Let $q > 2$ be a prime number and let $N = 2^s$ with $s \geq 1$. Then N is a primitive root modulo q if and only if 2 is a primitive root modulo q and s is relatively prime to $q-1$. Moreover, 2 is *not* primitive modulo q if and only if

$$2^{(q-1)/p} \equiv 1 \pmod{q}$$

for some prime factor p of $q-1$. If 2 is a primitive root modulo q then $q \equiv 3$ or $5 \pmod 8$. These facts make it fairly easy (using a computer algebra package) to find large primes q for which 2 is a primitive root. In [62] several examples are provided with $N = 2^s$ ($s = 25, 31, 32$ etc.) that involve only two or three multipliers but which have periods of the order of 10^{300} to 10^{600}.

7.8 Exercises

1. The Fibonacci sequence mod N can be generated using an LFSR over $\mathbb{Z}/(N)$, with connection polynomial $-1 + x + x^2$ and with initial loading $1, 1$. What are the possible periods for this sequence? What is the period of the Fibonacci sequence modulo 7?

2. Now consider the analogous FCSR over $\mathbb{Z}/(N)$, that is, the FCSR with connection integer $-1 + N + N^2$. What is the period of this "Fibonacci FCSR" sequence for $N = 7$? What are the possible periods of the Fibonacci FCSR sequence modulo N?

3. Let a be an integer and let b be a positive integer. Show that a is divisible by b in the N-adic numbers if and only if a is divisible by b in $\mathbf{Z}$.

4. Let $p \geq 2$ and u be integers. Show that the $(-p)$-adic expansion of u/q is periodic if and only if

$$-q/(p+1) \leq u \leq pq/(p+1).$$

What does this say about the number of periodic (-2)-adic sequences with a connection integer q?

Chapter 8 Algebraic Feedback Shift Registers

This is the most general, but also the most abstract chapter on sequence generators in the book. In this chapter we describe the theory of algebraic shift registers (AFSRs). This is a "shift register" setting for the generation of pseudo-random sequences, which includes as special cases the theory of linear feedback shift register (LFSR) sequences, feedback with carry shift register (FCSR) sequences, function field sequences, and many others. The general framework presented here first appeared in [107]. It "explains" the similarities between these different kinds of sequences. Many of the interesting properties of these sequences simply reflect general properties of any AFSR sequence. However, because this chapter is fairly abstract, the rest of the book has been written so as to be largely independent of this chapter, even at the expense of sometimes having to repeat, in a special case, a theorem or proof which is fully described in this chapter.

8.1 Definitions

The ingredients used to define an algebraic shift register are the following.

1. An integral domain R (see Definition 2.2.2).
2. An element $\pi \in R$.
3. A complete set of representatives $S \subset R$ for the quotient ring $R/(\pi)$. (This means that the composition $S \to R \to R/(\pi)$ is a one to one correspondence, see Section 5.5.)

For any $u \in R$ denote its image in $R/(\pi)$ by $\tilde{u} = u \pmod{\pi}$. Having chosen S, every element $a \in R$ has a unique expression $a = a_0 + b\pi$ where $a_0 \in S$. Then a_0 is the representative (in S) of $\tilde{a}$, and $a - a_0$ is divisible by π. Abusing notation slightly (by confusing $a_0 \in S$ with its image in $R/(\pi)$) we write

$$a_0 \equiv a \pmod{\pi} \quad \text{and} \quad b = a \,(\text{div } \pi) = \frac{a - a_0}{\pi}. \tag{8.1}$$

Definition 8.1.1. *Let $q_0, q_1, \cdots, q_m \in R$, and assume that q_0 is invertible* (mod π). *An algebraic feedback shift register (or AFSR) over (R, π, S) of length m with* multipliers *or* taps $q_0, q_1, \cdots, q_m$ *is a discrete state machine whose* states *are collections*

$$(a_0, a_1, \cdots, a_{m-1}; z) \quad \textit{where } a_i \in S \quad \textit{and } z \in R$$

consisting of cell contents a_i *and* memory z. *The state changes according to the following rules:*

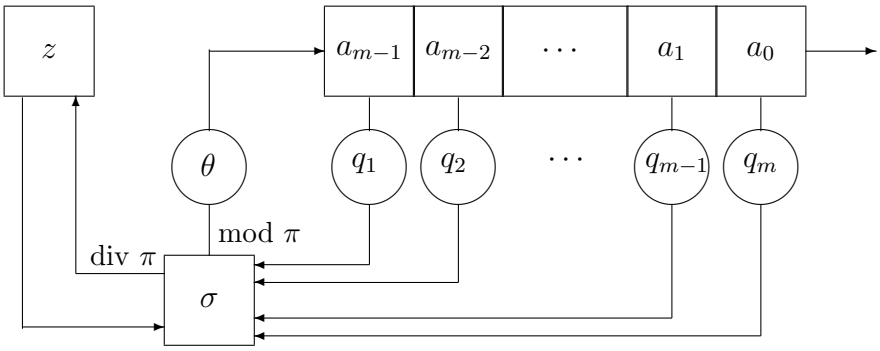


Figure 8.1: Diagram of an AFSR.

1. *Compute*
$$\sigma = \sum_{i=1}^{m} q_i a_{m-i} + z.$$
2. *Find* $a_m \in S$ *such that* $-q_0 a_m \equiv \sigma \pmod{\pi}$. *That is,* $\tilde{a}_m = -\tilde{q}_0^{-1}\tilde{\sigma}$.
3. *Replace* $(a_0, \cdots, a_{m-1})$ *by* $(a_1, \cdots, a_m)$ *and replace* z *by* $\sigma \text{ (div } \pi) = (\sigma + q_0 a_m)/\pi$.

We remark that a_m in step (2) depends only on the values of q_i, a_j, z modulo π. It is the unique lift to S of $-\tilde{q}_0^{-1}\tilde{\sigma} \in R/(\pi)$.

The register outputs an infinite sequence $a_0, a_1, \cdots$ of elements in S. The state change rules may be summarized by saying that this sequence satisfies a *linear recurrence with carry*,

$$-q_0 a_n + \pi z_n = q_1 a_{n-1} + \cdots + q_m a_{n-m} + z_{n-1} \tag{8.2}$$

for all $n \geq m$. Here, $z_i \in R$ represent the sequence of memory values, with $z = z_{m-1}$ being the initial value. At each stage the right side of this equation (denoted σ above) is determined by the memory and the previous m cell values. Then there is a unique $a_n \in S$ and a unique $z_n \in R$ so that equation (8.2) is satisfied. In practice one usually considers $q_i \in S$. However the analysis of the AFSR does not depend on this assumption. (See also Section 8.6.h). A diagram of an AFSR is given in Figure 8.1 where $\theta = -\tilde{q}_0^{-1}$. The element

$$q = \sum_{i=0}^{m} q_i \pi^i \in R \tag{8.3}$$

plays a central role[1]in the analysis of AFSRs and it is referred to as the *connection element*. The invertibility of $\tilde{q}_0$ is equivalent (cf. Section 5.5) to coprimality of π and q.

[1]If the q_i are restricted to lie in a complete set of representatives (such as S) for the elements of $R/(\pi)$, then the connection element q determines the multipliers $q_0, q_1, \cdots, q_{m-1}$. However if the q_i are not restricted in this

As with LFSRs and FCSRs, when thinking of AFSRs as physical devices as in the figure, we sometimes list the components of the state in descending order of their indices. We again write $\boxed{z}\boxed{a_{m-1}}\boxed{a_{m-2}}\boxed{\cdots}\boxed{a_1}\boxed{a_0}$ for the *machine state* to distinguish the two notations.

An LFSR (Chapter 6) over a field (or even an integral domain) K is an AFSR where $R = K[x]$ is the ring of all polynomials with coefficients in K, $\pi = x$, and $S = K$ is the set of polynomials of degree 0, which may also be identified with the quotient $R/(\pi) = K[x]/(x)$. If we initialize the memory to zero, then it remains zero throughout the infinite execution since there is no "carry" when multiplying and adding polynomials. In this case the connection element q coincides with the classical connection polynomial of equation (6.2).

An FCSR (Chapter 7) is an AFSR with $R = \mathbb{Z}$, $\pi = N$, and $S = \{0, 1, \cdots, N-1\}$. Three other cases of AFSRs will be described in this book: when $R = \mathbb{Z}[\pi]$ and $\pi = p^{1/d}$ (in Chapters 9 and 17), when R is the ring of integers in an algebraic number field and $\pi \in R$ is arbitrary (in Section 8.3.a), and when $R = K[x]$ and $\pi \in R$ is arbitrary and more generally when $R = K[x_1, \cdots, x_n]/I$ where I is an in ideal and $R/(\pi)$ is finite (in Chapter 15).

Note that the behavior (even the period) of an AFSR depends highly on the choice of S. (See Section 8.6.h.)

8.2 Properties of AFSRs

Throughout this section, we consider an AFSR over (R, π, S) and we assume, as in Section 5.5 that

$$\bigcap_{i=1}^{\infty} (\pi^i) = 0. \tag{8.4}$$

Recall from that section that R_π consists of all formal power series $\sum_{i=0}^{\infty} a_i \pi^i$ with $a_i \in S$. Let $q_0, q_1, \cdots, q_m \in R$ be a set of multipliers for an AFSR with connection element $q = q_0 + q_1\pi + \cdots + q_m\pi^m$ and assume that q_0 is invertible $\pmod{\pi}$. Hence q is relatively prime to π and by Proposition 5.5.4 it is invertible in R_π.

Fix an initial state (or initial loading) $(a_0, a_1, \cdots, a_{m-1}; z)$ of the register. The (infinite) output sequence $\mathbf{a} = a_0, a_1, \cdots$ may be identified with the π-adic number

$$a = a_0 + a_1\pi + a_2\pi^2 + \cdots \in R_\pi$$

which we refer to as the π-*adic value* of the register (with its given initial loading $(a_0, \cdots, a_{m-1}; z)$). We write $\mathbf{a} = \mathbf{seq}_\pi(a)$. We now show that the collection of π-adic numbers obtained in this way is exactly the set of fractions u/q with $u \in R$.

---

way then it is possible for two different AFSRs to correspond to the same connection element q. In fact, any $q \in R$ that is relatively prime to π is the connection element of a 1-stage AFSR obtained by taking $q = q_0 + \pi q_1$ where $q_0 = q \pmod{\pi} \in S$.

Lemma 8.2.1. *The equation*

$$u = \sum_{n=0}^{m-1} \sum_{i=0}^{n} q_i a_{n-i} \pi^n - z\pi^m \in R \tag{8.5}$$

determines a one to one correspondence between elements $u \in R$ and states of the AFSR.

Proof. It suffices to show that, given the multipliers q_i, every $u \in R$ has a unique representation as in equation (8.5) with $a_i \in S$. Suppose we are given $u \in R$. Reducing the right side of equation (8.5) modulo π gives $u \pmod{\pi} \equiv q_0 a_0 \pmod{\pi}$. Since q_0 is invertible mod π there is a unique solution for $a_0 \pmod{\pi}$, and this has a unique lift $a_0 \in S$. Similarly, by reducing $\pmod{\pi^2}$ we find

$$\frac{u - q_0 a_0}{\pi} \pmod{\pi} \equiv (q_0 a_1 + q_1 a_0) \pmod{\pi},$$

from which the knowledge of a_0 uniquely determines $a_1 \pmod{\pi} \in R/(\pi)$. Lifting this to S gives a unique value for a_1. Continuing in this way by induction, one arrives at unique values $a_0, a_1, \cdots a_{m-1}$ such that

$$-u + \sum_{n=0}^{m-1} \sum_{i=0}^{n} q_i a_{n-i} \equiv 0 \pmod{\pi^m}.$$

Therefore this difference is divisible by π^m. Hence it can be uniquely written as $z\pi^m$. □

The operation of the AFSR therefore defines a map $\tau : R \to R$ which models the state of the shift register in that the following diagram commutes.

$$\begin{array}{ccc} R & \longrightarrow & \text{states} \\ {\scriptstyle \tau}\downarrow & & \downarrow{\scriptstyle \text{state change}} \\ R & \longrightarrow & \text{states} \end{array} \tag{8.6}$$

In Corollary 8.2.3 we find a formula for τ.

Theorem 8.2.2. (Fundamental Theorem on AFSRs [107]) *The output of an AFSR with connection element q, initial memory value z_{m-1}, and initial loading $a_0, a_1, \cdots, a_{m-1}$, is the coefficient sequence $\mathbf{seq}_\pi(u/q)$ of the π-adic representation of the element $a = u/q \in R_\pi$, where u is given by (8.5).*

Proof. The proof is essentially the same as the proof of Theorem 7.2.1 (for the case of FCSRs), or of Theorem 6.4.1 (for the case of LFSRs) which ultimately goes back to [53]. Let

$$a = \sum_{i=0}^{\infty} a_i \pi^i \in R_\pi, \tag{8.7}$$

be the π-adic number corresponding to the output sequence of the AFSR. If $z_{m-1}, z_m, \cdots$ denotes the sequence of memory values then by equation (8.2)

$$-q_0 a_n = \sum_{i=1}^{m} q_i a_{n-i} + (z_{n-1} - \pi z_n), \tag{8.8}$$

provided that $n \geq m$. Now substitute equation (8.8) into the expression (8.7) for a to obtain

$$\begin{aligned} q_0 a &= q_0(a_0 + a_1\pi + \cdots + a_{m-1}\pi^{m-1} + \sum_{n=m}^{\infty} a_n \pi^n) \\ &= q_0 v - \sum_{n=m}^{\infty} \left(\sum_{i=1}^{m} q_i a_{n-i} \right) \pi^n - \sum_{n=m}^{\infty} (z_{n-1} - \pi z_n)\pi^n, \end{aligned} \tag{8.9}$$

where $v = a_0 + a_1\pi + \cdots + a_{m-1}\pi^{m-1}$ is the element represented by the initial loading of the register. The second summation in equation (8.9) cancels except for the first term, $-z_{m-1}$, leaving

$$\begin{aligned} q_0 a &= q_0 v - z_{m-1}\pi^m - \sum_{n=m}^{\infty} \sum_{i=1}^{m} q_i \pi^i a_{n-i} \pi^{n-i} \\ &= q_0 v - z_{m-1}\pi^m - \sum_{i=1}^{m} q_i \pi^i \left(\sum_{n=m}^{\infty} a_{n-i} \pi^{n-i} \right) \\ &= q_0 v - z_{m-1}\pi^m - \sum_{i=1}^{m} q_i \pi^i (a - (a_0\pi^0 + a_1\pi^1 + \cdots + a_{m-i-1}\pi^{m-i-1})) \\ &= q_0 v - z_{m-1}\pi^m - a\sum_{i=1}^{m} q_i \pi^i + \sum_{i=1}^{m-1} \sum_{j=0}^{m-i-1} q_i \pi^i a_j \pi^j \end{aligned}$$

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

$$\begin{aligned} a &= \frac{q_0 v - z_{m-1}\pi^m + \sum_{i=1}^{m-1} \sum_{j=0}^{m-i-1} q_i \pi^i a_j \pi^j}{q_0 + \sum_{i=1}^{m} q_i \pi^i} \\ &= \frac{\sum_{n=0}^{m-1} (\sum_{i=0}^{n} q_i a_{n-i})\pi^n - z_{m-1}\pi^m}{q}. \qquad \square \end{aligned} \tag{8.10}$$

We emphasize that the sequence $\mathbf{seq}_\pi(u/q)$ depends on the choice of S. How properties of $\mathbf{seq}_\pi(u/q)$ reflect choices of S is not well understood.

As discussed above, if $q_i \in R$ are arbitrarily chosen, then every $q \in R$ that is relatively prime to π can be realized as the connection element of some AFSR. However if the q_i are restricted to lie in a complete set of representatives (for example, the set S) for $R/(\pi)$, then the set of possible connection elements q that can be realized by a finite AFSR does not necessarily account for all elements in R that are relatively prime to π.

For example, in the basic N-ary FCSR model we always assume $q_0 = -1$. But there exist $q \in \mathbb{Z}$, relatively prime to N, that are not of this form. However, in the AFSR model of an N-ary FCSR we allow other values for q_0. Having done so, even if we require $q_i \in S$ for $i \geq 0$, it is easy to see that every $q \in \mathbb{Z}$ (with $\gcd(q, N) = 1$) has the property that either q or $-q$ can be realized as a connection integer for such an AFSR.

For $u \in R$, denote its reductions in $R/(\pi)$ and $R/(q)$ by $\tilde{u} \in R/(\pi)$ and $\bar{u} \in R/(q)$.

Corollary 8.2.3. *The map $\tau : R \to R$, which models the state change of the AFSR, is given by*

$$u = \pi\tau(u) + a_0 q \quad \text{or} \quad \tau(u) = \frac{u - a_0 q}{\pi}. \tag{8.11}$$

Here, $a_0 \in S$ is uniquely determined by Lemma 8.2.1; it may also be abstractly described as the unique lift to S of the element $\tilde{u}/\tilde{q} = \tilde{u}/\tilde{q}_0 \in R/(\pi)$. The mapping τ preserves the ideal $(q) \subset R$ and the induced mapping, also denoted by $\tau : R/(q) \to R/(q)$, is given by multiplication by $\bar{\pi}^{-1}$.

Proof. Let $u \in R$. If the π-adic expansion of u/q is

$$\frac{u}{q} = a_0 + a_1\pi + a_2\pi^2 + \cdots$$

then $a_0 \in S$ is the unique lift to S of $\tilde{u}/\tilde{q} \in R/(\pi)$. By Theorem 8.2.2 (and the definition of the state change operation $\tau : R \to R$), if $u' = \tau(u)$ denotes the succeeding element, then the π-adic expansion of u'/q is

$$\frac{u'}{q} = a_1 + a_2\pi + a_3\pi^2 + \cdots.$$

So

$$\pi\frac{u'}{q} + a_0 = \frac{u}{q} \quad \text{or} \quad u = \pi u' + a_0 q.$$

Therefore $u - a_0 q$ is divisible (in R) by π and $u' = (u - a_0 q)/\pi$, which proves equation (8.11). Reducing modulo q gives $\bar{u}' = \bar{\pi}^{-1}\bar{u}$. □

8.3 Memory requirements

For any hardware or software implementation of an AFSR, the memory must remain "bounded". That is, it should take on at most finitely many values throughout an infinite execution. There are

many examples of AFSRs for which the memory grows without bound, but in many cases explicit bounds are known. In an LFSR the memory is always 0. Memory requirements of an FCSR were analyzed in Section 7.6. There are two additional types of rings R for which it is possible to guarantee that the memory takes on finitely many values. The first is when the field of fractions (see Section 2.2.h) of R is an algebraic number field (a finite extension of the rational numbers, see Section 3.4). 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, so its field of fractions is a function field (see Section 5.2.d). In this case, when adding two elements, there is no "carry" from lower degree terms to higher degree terms, so the degree of the memory always remains bounded independent of the size of the AFSR.

8.3.a AFSRs over number fields

In this section we assume the ring R is an *order* in an algebraic number field F, cf. Section 3.4.c. (In this case, R is automatically an integral domain and F is its field of fractions.) Fix $\pi \in R$ and assume that $R/(\pi)$ is finite. Fix a complete set $S \subset R$ of representatives for $R/(\pi)$. Consider an AFSR based on a choice of elements $q_0, q_1, \cdots, q_m \in R$ where q_0 and π are relatively prime. A special class of such AFSRs is considered in detail in Chapter 9.

The field F 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. Let e_1 denote the number of real embeddings and let $2e_2$ denote the number of complex embeddings. Then $e_1+2e_2 = [F : \mathbb{Q}]$, the degree of the extension $F/\mathbb{Q}$ [17, p. 95].

Having fixed an embedding, denote by $|x|$ the norm of a complex number x. If z is the memory of the AFSR, we want to consider the growth of $|z|$ over an infinite execution.

Theorem 8.3.1. [107] *Let $\pi \in R \subset F$ be as above. If there is an embedding of F in the complex numbers such that $|\pi| < 1$, then there is an initial state of this AFSR such that the memory grows without bound.*

Proof. Suppose $a_0, a_1, \cdots$ is the output sequence, and $z_{m-1}, z_m, \cdots$ is the sequence of memory values as in Definition 8.1.1. Recalling equation (8.2),

$$\pi z_n = z_{n-1} + \sum_{i=0}^{m} q_i a_{n-i}.$$

It follows that

$$|z_n| \geq \frac{|z_{n-1}| - \sum_{i=0}^{m} |q_i a_{n-i}|}{|\pi|}$$

$$\geq \frac{|z_{n-1}| - (m+1)BC}{|\pi|}$$

where $B = \max\{|q_i| : 0 \leq i \leq m\}$ and $C = \max\{|c| : c \in S\}$. Suppose $|\pi| < 1$. Let

$$w_n = |z_n| - \frac{(m+1)BC}{1-|\pi|}.$$

Then

$$w_n \geq \frac{w_{n-1}}{|\pi|}.$$

Thus the $|z_n|$ are increasing as long as some w_n is positive or equivalently, as long as

$$|z_n| \geq \frac{(m+1)BC}{1-|\pi|}.$$

Thus there are initial states from which the memory increases without bound. In particular the memory takes infinitely many values in an infinite execution. □

To guarantee that the memory takes finitely many values, it is not sufficient to take a single embedding of F in $\mathbb{C}$ and show that the complex norm of the memory is bounded. Indeed, there may be infinitely many elements with complex norm bounded by a constant. For example, suppose that $F = \mathbb{Q}[\sqrt{2}]$. Then $u = \sqrt{2} - 1$ is positive, real, and less than one. But its powers u^k, $k \in \mathbb{N}$ are all positive, real, and distinct, with $0 < |u^k| < 1$.

Theorem 8.3.2. [107] *Let $\pi \in R \subset F$ be as above. Suppose that $|\pi| > 1$ for every embedding of F in the complex numbers. Then the memory in any infinite execution of the AFSR takes on only finitely many values. The output is therefore eventually periodic.*

Proof. First suppose that for a particular embedding of F we have $|\pi| > 1$. By the same reasoning as in the proof of Theorem 8.3.1 we see that

$$|z_n| < \frac{|z_{n-1}| + (m+1)BC}{|\pi|}.$$

If

$$|z_{n-1}| \leq \frac{(m+1)BC}{|\pi| - 1}, \tag{8.12}$$

then the same inequality (8.12) holds for $|z_n|$. If inequality (8.12) does not hold, then

$$w_n < \frac{w_{n-1}}{|\pi|},$$

where

$$w_n = |z_n| - \frac{(m+1)BC}{|\pi| - 1}.$$

We conclude that

$$|z_n| \leq D = \max\left(|z_{m-1}|, \frac{(m+1)BC}{|\pi| - 1}\right) \text{ for all } n \geq m-1. \tag{8.13}$$

As we have observed, a bound on all $|z_n|$ does not necessarily mean that the memory takes on only finitely many values: there may be infinitely many elements $y \in R$ such that $|y| < D$. To guarantee that the memory takes on finitely many values, we need to use the stronger assumption that $|\pi| > 1$ for every embedding of F into the complex numbers.

Let $e = e_1 + 2e_2$. Suppose $\sigma_1, \cdots, \sigma_{e_1+e_2}$ 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. The real embeddings are treated as functions into $\mathbb{R}$, while the complex embeddings are treated as functions to $\mathbb{C} = \mathbb{R}^2$. The preceding argument implies that there exists a bound $D' > 0$ so that $|\sigma_i(z_n)| < D'$ for all i and for all n. Consider the map $\psi : F \to \mathbb{R}^e$ defined by

$$\psi(x) = (\sigma_1(x), \cdots, \sigma_{e_1+e_2}(x)).$$

Then

$$||\psi(z_n)|| = (\sigma_1(x)^2, \cdots, \sigma_{e_1+e_2}(x)^2)^{1/2} < (e_1 + e_2)^{1/2} D'$$

for all $n \geq m-1$. The image of R under ψ is a full integer lattice of rank e, and ψ is injective [17, p. 95-99]. By Theorem 2.2.30, any set of points in $\psi(R)$ is finite if it is bounded in Euclidean norm. The theorem follows. □

8.3.b AFSRs over rational function fields

Fix a ring K which is an integral domain and take $R = K[x]$ to be the ring of polynomials with coefficients in K. It is also an integral domain (Theorem 2.2.14). If F denotes the fraction field of K then the fraction field of $K[x]$ is the field of rational functions, $F(x)$, see Section 3.5.b.

Fix $\pi \in K[x]$ a polynomial of any degree. Let $S \subset K[x]$ be a complete set of representatives for $K[x]/(\pi)$. Let $q_0, q_1, \cdots, q_m \in K[x]$ be polynomials and assume that q_0 is relatively prime to π. We consider the AFSR based on $(R = K[x], \pi, S)$ with multipliers given by $q_0, \cdots, q_m$. Sequences generated by such an AFSR are studied in more detail in Chapter 15.

Proposition 8.3.3. *Let* $U = \max\{deg(u) : u \in S\}$, $Q = \max\{deg(q_i) : 0 \leq i \leq m\}$ *and let* $d = \deg(\pi)$. *For any initial loading of the AFSR such that the degree of the initial memory* z_{m-1} *is* e, *the degree* $\deg(z_n)$ *of the memory* z_n *satisfies*

$$0 \leq \deg(z_n) \leq \max(U + Q - d, e - (n-m)d). \tag{8.14}$$

Thus

$$0 \leq \deg(z_n) \leq U + Q - d$$

provided n is sufficiently large. If the ring K is finite, then the output of the AFSR is eventually periodic.

Proof. From equation (8.2) we immediately obtain

$$\deg(z_n) \leq \max(U + Q, \deg(z_{n-1})) - d = \max(U + Q - d, \deg(z_{n-1}) - d)$$

which implies equation (8.14). If K has finitely many elements then there are only finitely many polynomials z_n with this property, so only finitely many states of the AFSR can be reached from any given initial loading. The proposition follows. □

We remark that the bound on the degree of the memory is independent of the length of the AFSR, and that if the q_i are restricted to lie in S then we may use U instead of Q, giving a bound on $\deg(z_n)$ which is also independent of the choice of multipliers.

8.3.c AFSRs over global function fields

Let $p \in \mathbb{Z}$ be prime, $h > 0 \in \mathbb{Z}$, and $r = p^h$. Let I be an ideal and π be an element in $\mathbb{F}_r[x_1, \cdots, x_n]$. Assume that $R = \mathbb{F}_r[x_1, \cdots, x_n]/I$ has transcendence degree 1 over $\mathbb{F}_r$ (meaning that the fraction field of R is a finite degree extension of the field of rational functions $\mathbb{F}_r(x)$, see Section 3.5.b). Assume also that $K = R/(\pi)$ is finite. Then $R/(\pi)$ is a vector space over $\mathbb{F}_r$ so its cardinality is a power of r, say r^e. If $n = 1$ then we are in the case of the previous section, and if $n = 1$, $\pi = x_1$, and $I = (0)$, then the AFSRs we obtain are exactly the LFSRs.

The general behavior of these AFSRs seems quite complex, so we want conditions under which they are well behaved enough to analyze. We need a "well structured" complete set of representatives S for R modulo π.

Hypothesis H1: The set S is closed under addition and contains $\mathbb{F}_r$.

Such sets S exist in abundance. For example, we can take S to be the $\mathbb{F}_r$-span of a set in R that contains 1 and maps one-to-one and onto a basis of $R/(\pi)$ over $\mathbb{F}_r$. Any S that satisfies H1 is closed under multiplication by $\mathbb{F}_p$, but possibly not under multiplication by any larger field. In general we can represent any element of R as a π-adic element with coefficients in S, but in order that we get good randomness properties we need to be able to represent the elements of R finitely.

Hypothesis H2: Every $v \in R$ is a finite linear combination $v = v_0 + v_1\pi + \cdots + v_\ell\pi^\ell$ with $v_i \in S$.

Since π is not a zero divisor, the above representation for v is unique if $v_\ell \neq 0$. Indeed, suppose

$$\sum_{i=0}^{m} u_i \pi^i = \sum_{i=0}^{\ell} v_i \pi^i \tag{8.15}$$

for some $u_i, v_i \in S$ with $u_m \neq 0 \neq v_\ell$ and $(u_0, u_1, \cdots) \neq (v_0, v_1, \cdots)$, and let ℓ be the minimal integer so that such a pair of representations exists. Reading equation (8.15) modulo π we see that $u_0 = v_0$. By subtraction and the fact that π is not a zero divisor we have

$$\sum_{i=0}^{m-1} u_{i+1} \pi^i = \sum_{i=0}^{\ell-1} v_{i+1} \pi^i$$

and $(u_1, u_2, \cdots) \neq (v_1, v_2, \cdots)$. This contradicts the minimality of ℓ.

The smallest ℓ in Hypothesis H2 with $v_\ell \neq 0$ is called the *π-degree of v*, $\ell = \deg_\pi(v)$.

Lemma 8.3.4. *If Hypotheses H1 and H2 hold, then for all $u, v \in R$,*

$$\deg_\pi(u+v) \leq \max(\deg_\pi(u), \deg_\pi(v)).$$

Let $a = \max\{\deg(st) : s, t \in S\}$. Then for all $u, v \in R$,

$$\deg_\pi(uv) \leq \deg_\pi(u) + \deg_\pi(v) + a.$$

Some randomness properties of AFSR sequences (see Theorem 15.2.2) require one further hypothesis. Let V_q denote the set of elements u of R such that $\mathbf{seq}_\pi(v/q)$ is (strictly) periodic.

Hypothesis H3: There exists k so that the elements of V_q are distinct modulo π^k.

Let $\mathbf{a} = a_0, a_1, \cdots$ be the output from an AFSR based on R, π, S with connection element $q = -q_0 + q_1\pi + \cdots + q_m\pi^m$ with $q_i \in S$ and with q_0 invertible modulo π. By Theorem 8.2.2 there exists $u \in R$ so $a_0 + a_1\pi + \cdots = u/q$. We call this a *rational representation* of $\mathbf{a}$. Any left shift of $\mathbf{a}$ can be generated by the same AFSR (with a different initial state), so also has a rational representation with denominator q.

Theorem 8.3.5. [96] *Suppose that S satisfies Hypotheses H1 and H2. Then the memory takes on only finitely many values in any infinite execution of an AFSR. Consequently every output sequence $\mathbf{a}$ is eventually periodic.*

Proof. Suppose that at some point the AFSR is in state

$$(a_j, a_{j+1}, \cdots, a_{j+m-1}; z)$$

with $z = \sum_{i=0}^{\ell} z_i \pi^i$ and $z_0, \cdots, z_\ell \in S$. Also, suppose that the maximal π-degree of the product of two elements of S is d. Then the carry z' of the next state satisfies

$$\pi z' = z + \sum_{i=0}^{m-1} q_i a_{j+m-i}.$$

The right hand side is divisible by π and its π-degree is at most $\max\{\ell, d\}$, so the π-degree of z' is at most $\max\{\ell - 1, d - 1\}$. Thus the π-degree of the carry decreases monotonically until it is less than d, and then remains less than d forever. This proves that only finitely many states of the AFSR are visited in any infinite execution, so eventually some state is repeated and from then on, the output is periodic. □

8.4 Periodicity

We now return to the general setting of AFSRs. We wish to consider the following question: Given an AFSR based on (R, π, S), for which initial loadings is the output sequence strictly periodic? This turns out to be a surprisingly subtle question in general. Throughout this section let us fix an AFSR based on (R, π, S) with multipliers $q_0, q_1, \cdots, q_m \in R$ where q_0 is relatively prime to π. Let $q = \sum_{i=0}^{m} q_i \pi^i$ be the connection element. Assume that $R/(\pi)$ and $R/(q)$ are finite. For $u \in R$ write $\tilde{u} \in R/(\pi)$ and $\bar{u} \in R/(q)$ for the image of u in these quotients.

The output sequence of the AFSR is $\mathbf{seq}_\pi(u/q)$ where $u \in R$ is given by equation (8.5). So the above question concerns periodicity of $\mathbf{seq}_\pi(u/q)$. We remark again that the operation of the AFSR and the expansion $\mathbf{seq}_\pi(u/q)$ both depend on the choice of the set $S \subset R$ of representatives for $R/(\pi)$.

Lemma 8.4.1. *Let $u \in R$. Consider any strictly periodic tail in the sequence $\mathbf{a} = \mathbf{seq}_\pi(u/q)$. Then this tail is $\mathbf{seq}_\pi(h/q)$ with $h \in R$ and with the same denominator q.*

Proof. By Lemma 8.2.1, the sequence $\mathbf{a}$ is the output sequence of an AFSR with connection element q, and some initial state. Then any strictly periodic tail of $\mathbf{a}$ is the output from some other state. By Theorem 8.2.2, this periodic tail is the π-adic expansion for some h/q with $h \in R$. □

Theorem 8.4.2. *Suppose (R, π, S) satisfies equation (8.4). Let $q \in R$ as above and suppose that for every $u \in R$ the sequence $\mathbf{seq}_\pi(u/q)$ is eventually periodic. Then there exists a complete set of representatives $\Delta \subset R$ for the elements of $R/(q)$ such that for every $u \in \Delta$ the sequence $\mathbf{seq}_\pi(u/q)$ is strictly peridic.*

Proof. We need to show that for any $h \in R/(q)$ there exists a representative $v \in R$ so that $v \pmod q = h$ and so that the π-adic expansion of v/q is strictly periodic. To see this, choose any lift, $u \in R$ of $h \in R/(q)$. By assumption, the π-adic expansion of u/q is eventually periodic. The periodic part occurs after some power, say, π^A. There exists $N > A$ so that $\pi^N \equiv 1 \pmod q$. (Just take N to be a sufficiently high multiple of the multiplicative order of $\bar{\pi} \in R/(q)$, which is finite by assumption.) Write the π-adic expansion of u/q as a polynomial of degree $(N-1)$ in π plus a tail,

$$\frac{u}{q} = a + \pi^N b,$$

where $a \in R$ and where the π-adic expansion of $b \in R_\pi$ is strictly periodic. By Lemma 8.4.1 we can write $b = v/q$ for some $v \in R$. This gives $u = aq + \pi^N v$. Reducing this equation modulo q we see that $u \equiv v \pmod q$ and that the π-adic expansion of v is strictly periodic. □

8.5 Exponential representation and period of AFSR sequences

The goal of this section is to describe AFSR sequences in a manner analogous to the trace representation and other exponential representations (Section 6.6) of an LFSR sequence. We have only succeeded in doing so under mild additional hypotheses.

Throughout this section, fix an AFSR based on (R, π, S) as in Section 8.1. Let $q_0, q_1, \cdots, q_m \in R$ be its collection of multipliers (where q_0 is relatively prime to π) and let $q = \sum_{i=0}^{m} q_i \pi^i \in R$ be the connection element. If $u \in R$, then denote by $\tilde{u}$ its image in $R/(\pi)$ and denote by $\bar{u}$ its image in $R/(q)$. Set $\gamma = \bar{\pi}^{-1} \in R/(q)$.

As in Theorem 8.4.2 we suppose that for any $u \in R$ the sequence $\mathbf{seq}_\pi(u/q)$ is eventually periodic (or, equivalently, that for any initial loading, the memory of the AFSR takes on finitely many values over its infinite execution). Let $V_q \subset R$ be the set of elements $u \in R$ such that $\mathbf{seq}_\pi(v/q)$ is strictly periodic. The change of state mapping (see Corollary 8.2.3) $\tau : R \to R$ preserves V_q. In Theorem 8.4.2 it was shown that there exists a complete set of representatives $\Delta \subset R$ for the elements of $R/(q)$ such that $\Delta \subset V_q$. Now we wish to make a further assumption, that such a set Δ can be chosen so that it is preserved by the change of state mapping τ.

Theorem 8.5.1. *Suppose there exists a set $\Delta \subset R$ so that*

1. *$\Delta \subset V_q$,*
2. *Δ is preserved by the change of state mapping τ, and*
3. *no two elements of Δ are congruent modulo q.*

Then for all $u \in \Delta$ the π adic expansion

$$\frac{u}{q} = a_0 + a_1\pi + \cdots$$

is given by

$$\tilde{a}_i = \tilde{q}_0^{-1}(\bar{u}\gamma^i \pmod q) \pmod \pi. \tag{8.16}$$

Note that now Δ is not necessarily a complete set of representatives modulo q.

The notation on the right side of this equation should be interpreted as follows. The element $\bar{u}\gamma^i \in R/(q)$ should be lifted to a complete set of representatives T for $R/(q)$ containing Δ, then reduced modulo π to give an element of $R/(\pi)$. Then it should be multiplied by $\tilde{q}_0^{-1} \in R/(\pi)$. The resulting element $\tilde{a}_i$ has a unique lift to S and this is a_i. There is another way to view this equation. Define the mapping $\Phi : R/(q) \to R/(\pi)$ to be the composition around the following diagram:

$$\begin{array}{ccccc} \Delta & \subset & T & \subset & R \\ & & \cong\downarrow & & \downarrow \\ & & R/(q) & \xrightarrow[\Phi]{} & R/(\pi) \end{array}$$

The mapping Φ is not a homomorphism. However, by abuse of notation we might refer to it as "reduction modulo π". Then the theorem says that $a_i \in R$ is the unique lift to S of the element

$$\tilde{a}_i = \tilde{q}_0^{-1}\Phi(\bar{u}\bar{\pi}^{-i}).$$

Proof. Let $u \in \Delta$. If the π-adic expansion of u/q is $\sum_{i=0}^{\infty} a_i\pi^i$, then by (8.11)

$$u = \pi\tau(u) + a_0 q \in \Delta. \tag{8.17}$$

Reading this equation modulo π gives

$$a_0 \equiv \tilde{q}^{-1}u \pmod \pi \equiv \tilde{q}_0^{-1}u \pmod \pi$$

which is the case $i = 0$ of equation (8.16). Reading equation (8.11) modulo q gives $\tau(\bar{u}) = \bar{u}\gamma$. So equation (8.16) follows by induction. □

Let n denote the multiplicative order of π modulo q. That is, n is the smallest nonnegative integer such that $\bar{\pi}^n = 1 \in R/(q)$. As in Section 3.2.c, if $u \in R$ define the *uth cyclotomic coset* modulo q relative to π to be the (finite) collection of elements $C_u(\pi) = \{\bar{u}\bar{\pi}^i : i = 0, 1, \cdots\}$. It is a subset of $R/(q)$ on which the group

$$\left\{\bar{\pi}^i : 0 \le i \le n-1\right\} \cong \mathbb{Z}/(n)$$

acts transitively by multiplication. It follows that the cardinality $c_u(\pi)$ of $C_u(\pi)$ is a divisor of n. It is the smallest nonnegative integer c such that $\bar{u}\bar{\pi}^c = \bar{u} \in R/(q)$.

Corollary 8.5.2. *Suppose that Δ satisfies hypotheses (1), (2), and (3) of Theorem 8.5.1. Let $u \in \Delta$. Then the period of the sequence $\mathbf{a} = \mathbf{seq}_\pi(u/q)$ is equal to $c_u(\pi)$. If $\bar{u} \in R/(q)$ is not a zero divisor, then $c_u(\pi) = n$.*

Proof. The period of the sequence of coefficients equals the period of the sequence of numerators $(u, u' = \tau(u), \cdots)$ as in the proof of Theorem 8.5.1. This coincides with the period of the sequence of images $(\bar{u}, \bar{u}\gamma, \cdots)$ in $R/(q)$. But this collection is exactly $C_u(\pi)$. Moreover, if $\bar{u}$ is not a zero divisor then the equation $\bar{u}\bar{\pi}^c = \bar{u}$ implies that $\bar{\pi}^c = 1$, and the least such c is in fact n, the order of $\bar{\pi}$. □

Corollary 8.5.3. *Suppose that no two elements of V_q are congruent modulo q. Then $\Delta = V_q$ satisfies hypotheses (1), (2), and (3) of Theorem 8.5.1. Therefore, if $u \in V_q$, then the sequence $\mathbf{a} = \mathbf{seq}_\pi(u/q)$ is:*

$$a_i = \tilde{q}_0^{-1}(\bar{u}\gamma^i \pmod q) \pmod \pi. \qquad \square$$

The description of V_q as the set of numerators u of q such that u/q has a periodic π-adic expansion is not usually very useful. Sometimes, however, we are able to more precisely identify this set. For example, if $R = K[x]$, $\pi = x$, and $S = K$, (the case when we have an LFSR), then $V_q = \{u : \deg(u) < \deg(q)\}$ and we can take $\Delta = V_q$. In this case the representation (8.16) was derived in Corollary 6.6.3. If $R = \mathbb{Z}$, $\pi > 0$, and $S = \{a : 0 \le a < q\}$, then $V_q = \{u : -q \le u \le 0\}$, and we can take $\Delta = \{u : -q < u \le 0\}$. The situation is more complex in other cases and we do not in general know analogous descriptions for V_q.

It is not always the case that a set Δ satisfying hypotheses (1), (2), and (3) of Theorem 8.5.1 exists. See for example Exercise 3 of this chapter. Even if such a Δ exists, it is not immediately clear how to characterize such a set because a given $h \in R/(q)$ might have several lifts $u \in V_q$. For example, suppose $\pi^2 = 2$, $R = \mathbb{Z}[\pi]$, 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 $u \equiv v \pmod q$ both give rise to periodic sequences. The congruence class modulo q does not uniquely determine a periodic sequence.

Suppose $\Delta \subset R$ is a complete set of representatives of $R/(q)$ consisting of elements u such that the π-adic expansion of u/q is strictly periodic. Even if Δ fails to satisfy conditions (2) and (3) of Theorem 8.5.1 it is still possible to say something about the period of this sequence.

Proposition 8.5.4. *If $\Delta \subset R$ is a complete set of representatives of $R/(q)$, then for any $u \in \Delta$ the period of $\mathbf{seq}_\pi(u/q)$ is a multiple of $c_u(\pi)$.*

Proof. If the sequence of coefficients is periodic of period T then by summing the geometric series we find that $u/q = v/(1-\pi^T) \in R_\pi$, or

$$u(1-\pi^T) = vq.$$

Therefore $\bar{u}(1-\bar{\pi}^T) = 0 \in R/(q)$ from which it follows that T is a multiple of the order of the coset of u. □

In particular, if u and q are relatively prime, then the period is a multiple of n, 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 8.5.4?

Proposition 8.5.5. *Let $q, u \in R$ with q relatively prime to π. There is a complete set of representatives $S \subset R$ for $R/(\pi)$ such that $\mathbf{seq}_\pi(u/q)$ is strictly periodic (with coefficients in S) and its period equals the cardinality $c_u(\pi)$ of the coset of u modulo q.*

Proof. Let $t = c_u(\pi)$. 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. □

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 8.5.5 is precisely the order of π modulo q. Examples where this fails (and the hypotheses of Theorem 8.5.1 fail) are given in Section 8.6.h.

8.5.a Function Fields

Let us return to the situation of Section 8.3.c: Let $r \in \mathbb{Z}$ be a power of a prime. Let I be an ideal and π, q be coprime elements in $\mathbb{F}_r[x_1, \cdots, x_n]$. Let $R = \mathbb{F}_r[x_1, \cdots, x_n]/I$. Assume that $K = R/(\pi)$ is finite. Let $S \subset R$ be a complete set of representatives for $R/(\pi)$. Suppose that S satisfies Hypotheses H1 and H2 of Section 8.3.c.

Theorem 8.5.6. *Under these hypotheses, no two elements of V_q are congruent modulo q. Let $u \in R$ and let $\mathbf{a} = \mathbf{seq}_\pi(u/q)$ be the coefficient sequence for the π-adic expansion of u/q. Then $\mathbf{a}$ is eventually periodic and the eventual period of $\mathbf{a}$ is a divisor of the multiplicative order of π modulo q. If $R/(q)$ is an integral domain, then the period equals the multiplicative order of π modulo q. In general, for a given q there is at least one periodic sequence with connection element q whose period is the multiplicative order of π modulo q.*

Proof. By hypothesis H1, the π-adic sum of two periodic sequences is the term-wise sum, so is also periodic. Thus, to prove that no two elements of V_q are congruent modulo q it suffices to show that no nonzero element of V_q is divisible by q. Suppose to the contrary that $uq \in V_q$. Then $u = uq/q$ has a periodic π-adic expansion. But this contradicts the fact that any element of R has a unique π-adic expansion since by Hypothesis H2 u has a finite π-adic expansion.

The eventual periodicity of $\mathbf{a}$ follows from Theorem 8.3.5. To describe the eventual period, it suffices to consider strictly periodic sequences since q is also a connection element of every shift of $\mathbf{a}$. Now consider the series of numerators $u_0, u_1, \cdots$ of the rational representations of the π-adic elements associated with the shifts of $\mathbf{a}$. The period is the least t such that $u_t = u_0$. By the preceding paragraph, this is equivalent to $u_t \equiv u_0 \pmod{q}$. For every i,

$$\frac{u_{i-1}}{q} = a_{i-1} + \pi \frac{u_i}{q},$$

and so

$$u_{i-1} = qa_{i-1} + \pi u_i.$$

Therefore

$$u_i \equiv \pi^{-1} u_{i-1} \equiv \pi^{-i} u_0 \pmod{q}$$

by induction. Thus

$$u_t \equiv u_0 \pmod{q} \text{ if and only if } \pi^t u_0 \equiv u_0 \pmod{q},$$

which is equivalent to: $(\pi^t - 1)u_0 \equiv 0 \pmod{q}$.

If $R/(q)$ is an integral domain, then this says that t is the multiplicative order of π modulo q. If $R/(q)$ is not an integral domain, then it implies that t is a divisor of the multiplicative order of π modulo q, see Section 2.2.a.

Finally, consider the coefficient sequence of the π-adic expansion of $1/q$. This sequence may not be periodic, but it is eventually periodic. Thus for some j its shift by j positions is periodic. This shift has a rational representation u/q, and by the above argument, $u \equiv \pi^{-j} \pmod{q}$. In particular, u is invertible modulo q, so

$$(\pi^i - 1)u \equiv 0 \pmod{q} \text{ if and only if } \pi^i \equiv 1 \pmod{q}.$$

Thus in this case the period equals the multiplicative order of π modulo q. □

8.6 Examples

In this section we illustrate the behavior of AFSRs by several examples. In each case we give a table of some of the successive states of the AFSR. These tables are written so that we shift toward the left at each state transition.

8.6.a $R = \mathbb{Z}$, $p = \pi = 2$

Suppose that $R = \mathbb{Z}$ so $F = \mathbb{Q}$. Let $\pi = p = 2$, so $K = \mathbb{F}_2$, and let $S = \{0, 1\}$. If $q_0 = -1$, then this is the setting that gives rise to FCSRs. Suppose that

$$q = \pi^4 + \pi^3 + \pi^2 - 1 = 27.$$

Then an AFSR with connection element q has four stages, with coefficients 1, 1, 1, and 0. If we start the register in the initial state $(0, 0, 0, 1)$, and with initial memory 0, then the sequence of states of the register is given in Table 8.1. The output sequence thus has period 18, and one period is

$$\mathbf{a} = 000101101111010010 \cdots.$$

Note that the memory size never exceeds two bits, so in effect we have a six stage binary feedback register with period 18. Also note that

$$\mathrm{ord}_q(2) = 18,$$

and $2^{-1} \pmod{q} = 14$. The exponential representation of this sequence is

$$a_i = (4 \cdot 14^i \pmod{27}) \pmod{2}.$$

register	mem	i	register	mem	i
0001	0	0	1110	2	9
0010	0	1	1101	2	10
0101	0	2	1010	2	11
1011	0	3	0100	2	12
0110	1	4	1001	1	13
1101	1	5	0010	1	14
1011	1	6	0100	1	15
0111	1	7	1000	1	16
1111	1	8	0000	1	17

Table 8.1: The states of an AFSR with $R = \mathbb{Z}$, $p = \pi = 2$, and $q = 27$.

Finally, since the period is 18 and one period gives the binary representation of 80672, the rational representation of the sequence is

$$\frac{-80672}{2^{18} - 1} = \frac{-80672}{262143} = \frac{-1}{27}.$$

8.6.b $R = \mathbb{Z}[\pi]$ with $\pi = \sqrt{2}$

Suppose that $R = \mathbb{Z}[\pi]$ with $\pi^2 = p = 2$, so $F = \mathbb{Q}[\pi]$ is a real quadratic number field and $K = \mathbb{F}_2$. Let $S = \{0, 1\}$. This is an example of a d-FCSR, $d = 2$ [58], which are analyzed in detail in Chapters 9 and 17. Every element in R can be written in the form $a + b\pi$ with a and b integers. The embeddings of F in $\mathbb{C}$ are determined by mapping π to either the positive or negative real square root of 2, each of which has complex norm greater than 1. Thus every AFSR in this setting has eventually periodic output. Let

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

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 $(1, 1, 1; 0)$, then the sequence of states of the register is given in Table 8.2.

The output sequence has period 16. One period is:

$$\mathbf{a} = 1110111100010000 \cdots.$$

The memory size never exceeds two bits, so in effect we have a five stage binary feedback register with period 16. We claim that

$$\pi^{-1} \pmod{q} = 3 \quad \text{and} \quad \mathrm{ord}_q(\pi) = 16.$$

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

Table 8.2: The states of an AFSR with $R = \mathbb{Z}[\pi]$, $\pi = 2^{1/2}$ and $q = 3\pi - 1$.

To see this, observe that $a + b\pi \equiv (3a + b)\pi \pmod{q}$ for any integers a, b. 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 $R/(q)$ is isomorphic to the integers modulo 17, and every element has multiplicative order dividing 16. Since $\pi^8 = 16$ is not congruent to one modulo q, π is primitive. The exponential representation of this sequence is thus

$$a_i = (3^i \pmod{3\pi - 1}) \pmod 2.$$

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

$$\frac{-15 - 45\pi}{\pi^{16} - 1} = \frac{-15 - 45\pi}{255} = \frac{-1}{17} + \frac{-3}{17}\pi = \frac{-1}{3\pi - 1}.$$

The third representation shows that the sequence **a** can be realized by alternating symbols from the 2-adic expansions of the rational numbers $-1/17$ and $-3/17$. But generating the sequence this way would require a pair of length 4 FCSRs each needing one bit of extra memory, for a total 10 state bits.

Alternatively, we could treat the entire sequence as a 2-adic number. From this point of view, it is the 2-adic number $-9/257$, which is generated by an FCSR of length 8 requiring one extra bit of memory.

Alternatively, we could treat the entire sequence as a power series over $\mathbb{F}_2$. From this point of view, it is the rational number

$$\frac{-x^4 + x^3 + 1}{(x+1)^{11}},$$

which is generated by an LFSR of length 11.

Of course there may be a more efficient way to realize the sequence as the output of an AFSR based on another ring.

8.6.c $R = \mathbb{Z}[\pi, \gamma]$ with $\pi = \sqrt{2}$ and, $\gamma^2 = \gamma + 1$

Suppose that $R = \mathbb{Z}[\pi, \gamma]$ with $\gamma^2 = \gamma+1$ and $\pi^2 = p = 2$. Here $F = \mathbb{Q}[\pi, \gamma]$ is a degree 4 extension of $\mathbb{Q}$. Also, γ reduces modulo π to a primitive cube root of 1, so $K = \mathbb{F}_4$. Let $S = \{0, 1, \gamma, 1+\gamma\}$. 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 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 8.3. Each output symbol is of the form $a + b\gamma$, with $a, b \in \{0, 1\}$. We claim that

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

$$\pi^{-1} \pmod{q} = 2\gamma + 3 \quad \text{and} \quad \text{ord}_q(\pi) = 400.$$

To see this, check that

$$a + b\pi \equiv ((2\gamma + 3)a + b)\pi \pmod{q}$$

for any integers a and b. 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

$$a_i = (3^i \pmod{3\pi - 1}) \pmod 2.$$

Finally, by Theorem 8.2.2, the rational representation of the sequence is

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

8.6.d $R = \mathbb{Z}$, $\pi = 3$

Next we consider an example that illustrates the behavior of an AFSR when $q_0 \neq 1$. Let $R = \mathbb{Z}$, $\pi = 3$, and $S = \{0, 1, 2\}$. Let $q = 43 = \pi^3 + 2\pi^2 - 2$. Then q is the connection element of a 3-stage AFSR over $\mathbb{Z}$, π, and S. Its coefficients are 1, 2, and 0. If the register is in state $(a_0, a_1, a_2; z)$, then the new element a_3 and memory z' are determined by the equation

$$2a_3 + 3z' = \sigma = a_0 + 2a_1 + z.$$

This AFSR has period 42, which is equivalent to 3 being a primitive element modulo 43. The memory is always -1, 0, or 1, so in effect it has four 3-ary cells and period 42. A full period of the register is given in Table 8.4.

One period of the output sequence from this register is

$$\mathbf{a} = 000210111022020100101222012111200202122121 \cdots.$$

We have $\pi^{-1} \pmod q = 29$ and $2^{-1} \pmod 3 = 2$, so the exponential representation is

$$a_i = 2(27 \times 29^i \pmod{43}) \pmod 3,$$

where the $27 = 3^{-39}$ introduces an appropriate shift.

register	mem	i	register	mem	i	register	mem	i
000	1	0	010	1	14	112	0	28
002	-1	1	100	1	15	120	1	29
021	-1	2	001	0	16	200	2	30
210	1	3	010	0	17	002	0	31
101	1	4	101	0	18	020	0	32
011	0	5	012	-1	19	202	0	33
111	0	6	122	-1	20	021	0	34
110	1	7	222	0	21	212	0	35
102	0	8	220	2	22	122	0	36
022	-1	9	201	2	23	221	1	37
220	1	10	012	0	24	212	1	38
202	1	11	121	0	25	121	1	39
020	1	12	211	1	26	210	2	40
201	1	13	111	1	27	100	2	41

Table 8.4: One period of an AFSR over $\mathbb{Z}$ with $\pi = 3$, and $q = 43 = \pi^3 + 2\pi^2 - 2$.

8.6.e $R = \mathbb{Z}$, $p = \pi = 2$, revisited

Suppose again that $R = \mathbb{Z}$ so $F = \mathbb{Q}$. Let $\pi = p = 2$, so $K = \mathbb{F}_2$, and $S = \{0, 1\}$. Again suppose that $q = 27$, but now we write $q = 16\pi - 5$, so that $q_1 = 16$ and $q_0 = -5$. This representation of q gives rise to an AFSR with one stage. If we start the register in the initial state (0), and with initial memory 4, then the sequence of states of the register is given in Table 8.5.

The output sequence is

$$\mathbf{a} = 000101101111010010\cdots,$$

identical to the output in Section 8.6.a. The memory size never exceeds four bits, so in effect we have a five stage binary feedback register with period 18, a more memory efficient implementation than the previous one (and the best possible for sequences of period 18). The remaining analysis is the same as before.

Similarly, we can take use the representation $q = 27 = 6\pi^2 + 3 \cdot \pi - 3$, which gives rise to a length 2 AFSR with coefficients $q_0 = -3$, $q_1 = 3$, and $q_2 = 6$. If we start the register in the initial state $(0, 0)$, and with initial memory 2, then the sequence of states of the register is given in Table 8.6. The output sequence is again $\mathbf{a} = 000101101111010010\cdots$. The memory size never exceeds three bits, so again we have a five stage binary feedback register with period 18.

register	mem	i	register	mem	i
0	4	0	1	7	9
0	2	1	1	9	10
0	1	2	1	10	11
1	-2	3	0	13	12
0	7	4	1	4	13
1	1	5	0	10	14
1	6	6	0	5	15
0	11	7	1	0	16
1	3	8	0	8	17

Table 8.5: The states of an AFSR with $R = \mathbb{Z}$, $\pi = 2$, and $q = 27$.

register	mem	i	register	mem	i
00	2	0	11	4	9
00	1	1	11	5	10
01	-1	2	10	7	11
10	1	3	01	5	12
01	2	4	10	4	13
11	1	5	00	5	14
10	5	6	01	1	15
01	4	7	10	2	16
11	2	8	00	4	17

Table 8.6: The states of an AFSR with $R = \mathbb{Z}$, $\pi = 2$, and $q = 27$.

8.6.f $R = \mathbb{F}_2[x]$, $\pi = x^2 + x + 1$

LFSRs arise as AFSRs with $R = \mathbb{F}_2[x]$ and $\pi = x$. Yet another general class of AFSRs arises by again letting $R = \mathbb{F}_2[x]$ but taking π to be a polynomial other than x. As usual we can take S to be any complete set of representatives modulo π, but it is natural to take $S = \{f(x) : \deg(f) < \deg(\pi)\}$. For example, we can let $\pi = x^2+x+1$. When we build AFSRs over this R and π we obtain a sequence over $\mathbb{F}_2[x]/(\pi) = \mathbb{F}_4$. Let us take the connection element $q = x^4 + x^3 + 1 = \pi^2 + x\pi + x$. An AFSR with connection element q has two stages, with coefficients 1 and x. If the register is in state $(a_0, a_1; z)$, then the new bit a_2 and memory z' are determined by the equation

$$xa_2 + \pi z' = \sigma = a_0 + xa_1 + z.$$

If we start the register in the initial state $(1, 0; 0)$, then the output sequence has period 15. In particular, π is primitive modulo q. The memory is always 0 or 1, so this is in effect a 5 bit feedback register with period 15. A full period of the register is given in Table 8.7.

register		mem	i	register		mem	i
1	0	0	0	$x+1$	x	0	8
0	$x+1$	1	1	x	0	1	9
$x+1$	0	1	2	0	x	1	10
0	1	0	3	x	1	1	11
1	1	0	4	1	$x+1$	1	12
1	x	1	5	$x+1$	$x+1$	0	13
x	x	0	6	$x+1$	1	1	14
x	$x+1$	0	7				

Table 8.7: One period of an AFSR over $\mathbb{F}_2[x]$ with $\pi = x^2 + x + 1$, and $q = \pi^2 + x\pi + x$.

One period of the output sequence from this register is

$$\mathbf{a} = 1, 0, x+1, 0, 1, 1, x, x, x+1, x, 0, x, 1, x+1, x+1, \cdots.$$

As we see in Chapter 15, this sequence is a pseudonoise sequence over $\mathbb{F}_4$ but is not an m-sequence. The inverse of π modulo q is $x^3 + x^2 + x$ and the inverse of $q_0 = x$ modulo π is $x + 1$, so the exponential representation of this sequence is

$$a_i = (x+1)((x^3+x^2)(x^3+x^2+x)^i \pmod{x^4+x^3+1}) \pmod{x^2+x+1},$$

where $x^3 + x^2 = \pi^{-13}$ introduces an appropriate shift.

8.6.g $R = \mathbb{F}_2[x, y]/(y^2 + x^3 + 1)$, $\pi = x^2 + y$

More generally, we can let R be a finitely generated ring over a finite field such that $R/(\pi)$ is finite. In this example we let $R = \mathbb{F}_2[x, y]/(y^2 + x^3 + 1)$ and $\pi = x^2 + y$. We can represent every element of R in the form $f_0(x) + yf_1(x)$ where f_0 and f_1 are polynomials. Then every element of $R/(\pi)$ can be represented by a polynomial in x alone. We have $x^4 + x^3 + 1 = \pi^2 + y^2 + x^3 + 1 \in (\pi, y^2 + x^3 + 1)$, so $\mathbb{F}_2[x]/(x^4 + x^3 + 1)$ maps onto $R/(\pi)$. But $x^4 + x^3 + 1$ is irreducible, so $\mathbb{F}_2[x]/(x^4 + x^3 + 1) = \mathbb{F}_{16}$ is a field. Since the only ideal in a field is the zero ideal, this map is an isomorphism. That is, $R/(\pi) = \mathbb{F}_{16}$. We can take $S = \{f(x) : \deg(f) < 4\}$. When we build AFSRs over R and π we obtain sequences over $\mathbb{F}_{16}$.

For example, let us take the connection element

$$q = \pi^2 + y\pi + 1 = x^4 + 1 + yx^2.$$

Consider a two stage AFSR with connection element q, with coefficients 1 and y. If the register is in state $(a_0, a_1; z)$, then the new element a_2 and memory z' are determined by the equation

$$a_2 + \pi z' = \sigma = a_0 + y a_1 + z.$$

If we start the register in the initial state $(1, 1; 0)$, then the output sequence has eventual period 14. In particular, π is not primitive modulo q. The memory is always of the form $f_0(x) + \pi f_1(x)$, where $\deg(f_0) \leq 3$ and $\deg(f_1) \leq 1$. The cells of the register part are polynomials of degree at most 3. Thus this is in effect a 14 bit feedback register with period 14 — not very good. The first 16 states of the register are given in Table 8.8. Note that state 1 and state 15 are identical.

register		mem	i
1	1	0	0
1	x^2+1	1	1
x^2+1	x^3+x^2+1	$\pi + x^2 + 1$	2
x^3+x^2+1	x^2+x	$x\pi + x^3 + x^2$	3
x^2+x	0	$\pi + x^2$	4
0	x	1	5
x	x^3+1	x	6
x^3+1	x^3+x^2+x+1	$(x+1)\pi + x^3 + 1$	7
x^3+x^2+x+1	x^3+x^2+x	$x\pi + x^3 + x^2$	8
x^3+x^2+x	x^3+1	$x\pi + x^3 + x^2$	9
x^3+1	x^3+x^2+1	$(x+1)\pi + x^3 + x + 1$	10
x^3+x^2+1	x^2	$x\pi + x^3 + x^+x2$	11
x^2	x^3+x	$\pi + x^2 + x$	12
x^3+x	$x+1$	$\pi + x^2 + x + 1$	13
$x+1$	1	x	14
1	x^2+1	1	15

Table 8.8: One period of an AFSR over $\mathbb{F}_2[x]$ with $\pi = x^2 + x + 1$, and $q = \pi^2 + x\pi + x$.

It is not obvious how to do computation in the ring R. The main entries in the AFSR are polynomials in x of degree at most 3, and the memory is a sum of powers of π whose coefficients are polynomials in x of degree at most 3. To compute a state change, we need to know how to express monomials yx^i, with $0 \leq i \leq 3$, as sums of powers of π whose coefficients are polynomials in x of degree at most 3. First note that we have $\pi^2 = x^4 + y^2 = x^4 + x^3 + 1$, so that

$$x^4 = \pi^2 + x^3 + 1. \tag{8.18}$$

i	yx^i
0	$\pi + x^2$
1	$x\pi + x^3$
2	$\pi^2 + x^2\pi + x^3 + 1$
3	$(x+1)\pi^2 + x^3\pi + x^3 + x + 1$

Table 8.9: Monomials yx^i as sums of powers of π.

Now we can repeatedly multiply $y = \pi + x^2$ using equation (8.18) to obtain the monomials yx^i, given in Table 8.9.

One period of the output sequence from this register is

$$1, x^2+1, x^3+x^2+1, x^2+1, 0, x, x^3+1, x^3+x^2+x+1, x^3+x^2+x, x^3+1, x^3+x^2+1, x^2, x^3+x, x+1.$$

There is no exponential representation for this sequence since it is not periodic.

Now consider what happens if instead of S we use the set of $T = \{a+bx+cy+dxy : a, b, c, d \in \{0, 1\}\}$ as our set of representatives modulo π. We can generate a sequence with an AFSR based on the same connection element $q = \pi^2 + y\pi + 1$ starting with the same initial state. However, now the sequence is periodic with period 127. The first 16 states are given in Table 8.10.

In this case the memory is always of the form $a + bx + c\pi$ with $a, b, c \in \{0, 1\}$, so we have an 11 bit register with period 127. In computing in this setting, we use the identities $y^2 = 1 + xy + x\pi$ and $xy^2 1 + x + xy + x\pi + \pi^2$.

We also note that in both these examples the state change function is $\mathbb{F}_2$-linear when we think of the state as a bit vector. By Theorem 10.6.2, from the point of view of linear span these AFSRs are equivalent to LFSRs.

8.6.h Dependence of the period on S

It is not the case that for every choice of S the period is the order of the cyclotomic coset of u modulo q. For example, let $R = \mathbb{Z}$ 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, $-2/3$ has the coefficient sequence $0101\underline{01}\cdots$, with period 2. With respect to S_2, $-2/3$ has the coefficient sequence $41144411\underline{14441}\cdots$, with (eventual) period 6. Now suppose u and q are arbitrary relatively prime integers. By Corollary 7.4.3 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. Consider the set of residues $S_3 = \{0, k\}$ for k odd and relatively prime to u. Suppose

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

register		mem	i
1	1	0	0
1	$y+1$	0	1
$y+1$	$xy+y$	x	2
$xy+y$	$y+1$	π	3
$y+1$	1	$x+1$	4
1	x	0	5
x	$xy+1$	0	6
$xy+1$	$xy+y+1$	$\pi+x$	7
$xy+y+1$	$xy+y+1$	$\pi+1$	8
$xy+y+1$	$xy+1$	$\pi+1$	9
$xxy+1$	$xy+y+x+1$	$\pi+x+1$	10
$xy+y+x+1$	$y+x+1$	$\pi+1$	11
$y+x+1$	$xy+x+1$	$x+1$	12
$xy+x+1$	$x+1$	$\pi+x$	13
$x+1$	$y+1$	1	14
$y+1$	$xy+y+x+1$	1	15

Table 8.10: The first 16 states of an AFSR over $\mathbb{F}_2[x]$ with $\pi = x^2+x+1$, and $q = \pi^2 + x\pi + x$.

where $a_i \in S_3$ and $b_i = a_i/k$. 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 order of 2 modulo q and order of 2 modulo 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 q, AFSR with connection element q, and initial state of the AFSR. 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 $\pi = 1 + \sqrt{-1}$ be a root of the quadratic equation $x^2 - 2x + 2 = 0$. Then the ring $R = \mathbb{Z}[\pi]$ equals the Gaussian domain $\mathbb{Z}[\sqrt{-1}]$ and π is prime. We have that $R/(\pi) \cong Z/(2)$, and so may choose complete sets of residues $S_1 = \{0, 1\}$ and $S_2 = \{0, 3\}$. Consider an AFSR over (R, π, S_1) with connection element $q = \pi^3 + \pi - 1$ (so the length $r = 3$), initial state $(a_0, a_1, a_2; z) = (1, 1, 1; 1 - \pi)$. Since $\pi^2 = 2\pi - 2$ and $2 = \pi(2 - \pi)$, we have

$$\sigma = a_0 q_3 + a_1 q_2 + a_2 q_1 + z = 2 + (1 - \pi) = 1 + (2 - \pi) = 1 + \pi(1 - \pi).$$

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 8.11: The states of an AFSR over the Gaussian domain using S_2.

Therefore the feedback element is $a_3 = 1$ and the updated memory is $z = 1-\pi$, 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 the same u, q, π, but replace S_1 by S_2. We then have an AFSR over (R, π, S_2). In terms of S_2, the initial state is $(3, 3, 3; 1-\pi)$. The first 25 iterations are displayed in Table 8.11.

The output sequence has period 24. In terms of K, one period is

$$\mathbf{a} = 111110001110101111001010\cdots .$$

By Theorem 8.2.2, the rational representation for the output is $(\pi^4 - \pi^3 - 3)/q$, which is a reduced rational representation.

8.7 Exercises

1. Let $N \geq 2$ be a natural number and let q be a positive integer that is relatively prime to N. We have observed that by allowing the multipliers to be arbitrary we may construct many different AFSRs with various lengths that have connection element q. Give a tight bound in terms of the multipliers for the maximum number of N-ary symbols needed to store the extra memory.

2. In this exercise, use the bound proved in exercise 8.7.1.

a. Find the length 1 AFSR with connection integer q and nonnegative multipliers that minimizes the amount of storage space needed for the state (including the extra memory, but not including the fixed multipliers).
b. For which q does the optimal length 1 AFSR result in a memory savings over the FCSR with connection integer q (with multipliers in $\{0, 1\}$)?

3. Let $p > 0$ be square free, $\pi^2 = -p$, and $R = \mathbb{Z}[\pi]$. Let $q = \pi^n - 1$ with $n \equiv 2 \pmod 4$

a. Characterize the $u \in R$ such that the π-adic expansion of u/q is periodic. (Hint: Use Exercise 7.8.4.)
b. Show by counting that there exists $u \in R$ such that $u/(\pi^n - 1)$ has a periodic π-adic expansion whose period does not divide n.
c. Show that the hypotheses of Theorem 8.5.1 do not always hold.

Chapter 9 d-FCSRs

In this chapter we consider d-FCSRs, a special case of AFSRs that is nearly as amenable to analysis as LFSRs and FCSRs.

9.1 Binary d-FCSRs

d-FCSRs were introduced in [106] and [58] and further analyzed in [58, 61, 63]. In this section we consider the special case of binary d-FCSRs; the general case is analyzed beginning in Section 9.2. A binary d-FCSR consists of a shift register with cell contents $a_0, a_1, \ldots, a_{m-1}$, feedback connections $q_m, q_{m-1}, \ldots, q_1$, and memory cells $m_0, m_1, \ldots, m_s$, each of which is a 0 or 1. The operation of a d-FCSR is similar to that of the FCSR except that each "carried" bit is delayed $d-1$ steps before being added.

This is best understood using the ring $\mathbb{Z}[\pi]$ which consists of polynomials in π (with integer coefficients), subject to the formal relation $\pi^d = 2$. The ring $\mathbb{Z}[\pi]$ contains the integers $\mathbb{Z}$ and it can be embedded as a subring of the real numbers $\mathbb{R}$ by mapping π to the positive $\sqrt[d]{2}$. However there are also other embeddings into the complex numbers. Any $z \in \mathbb{Z}[\pi]$ may be uniquely expressed as a polynomial $z = z_0 + z_1\pi + \cdots + z_{d-1}\pi^{d-1}$ with $z_i \in \mathbb{Z}$ by making use of the equation $\pi^d = 2 \cdot \pi^0$ whenever higher powers of π are encountered. Let us say that such an element z is *nonnegative* if each $z_i \geq 0$. (This is stronger than saying that the associated real number is nonnegative.) Using the binary expansion of each z_i, any nonnegative element $z \in \mathbb{Z}[\pi]$ can be uniquely expressed as a polynomial

$$z = \sum_{i=0}^{e} z_i' \pi^i$$

with coefficients $z_i' \in \{0, 1\}$ and $e \geq 0$. (For this reason, it is possible to implement the binary d-FCSR using binary logic hardware. See Figure 9.1.) Addition and multiplication preserve nonnegative elements, and are performed as for integers, except that carried bits are advanced d steps because

$$1 + 1 = 2 = 0 + 0\pi + 0\pi^2 + \cdots + 0\pi^{d-1} + 1\pi^d.$$

So it is best not to think of these coefficients as lying in the field $\mathbb{F}_2$. The operations $(\bmod \pi)$ and $(\operatorname{div} \pi)$ make sense in this ring. If $z = z_0 + z_1\pi + \cdots + z_{d-1}\pi^{d-1}$ then $z \pmod \pi = z_0 \pmod 2 \in \mathbb{F}_2$, and we will say that z is *odd* if $z \pmod \pi = 1$. (For example, $-1 = 1 - \pi^d$ so $-1 \pmod \pi = 1$.) Similarly $z \operatorname{div} \pi = z_1 + z_2\pi + \cdots + z_{d-1}\pi^{d-2}$.

We represent the memory by a nonnegative element $z \in \mathbb{Z}[\pi]$. The *connection element*

$$q = -1 + q_1\pi + q_2\pi^2 + \cdots + q_m\pi^m \in \mathbb{Z}[\pi]. \tag{9.1}$$

is associated to the feedback connections. Then $q \in \mathbb{Z}[\pi]$ is odd, and $q+1$ is non-negative. The output sequence is given by the *linear recurrence with delayed carry*

$$\pi z_t + a_t = z_{t-1} + \sum_{i=0}^{m-1} q_i a_{t-i} \tag{9.2}$$

(for $t \geq m$), with initial memory $z = z_{m-1}$. This equation can be solved for z_t and a_t by first finding $a_t \in \{0,1\}$: it is the right hand side of equation (9.1) reduced $\pmod{\pi}$.

To be explicit, the operation of the d-FCSR may be described as follows: Form the integer sum

$$\sigma' = \sum_{i=0}^{m-1} q_i a_{m-i}.$$

Write σ' as a nonnegative element of $\mathbb{Z}[\pi]$ – that is, as a polynomial in π with binary coefficients – using $2 = \pi^d$. (It is this fact which gives rise to the delay by d steps in the carry operation.) Using addition in $\mathbb{Z}[\pi]$ form the (nonnegative) sum $\sigma = z + \sigma' \in \mathbb{Z}[\pi]$. Shift the contents of the register cells to the right by one step. Place the bit $a_m = \sigma \pmod{\pi}$ in the leftmost register cell. Replace the memory by $z' = \sigma \text{ (div } \pi) = (\sigma - a_m)/\pi \in \mathbb{Z}[\pi]$. Thus the new state $(a'_0, a'_1, \ldots, a'_{m-1}; z')$ is related to the old state $(a_0, a_1, \ldots, a_{m-1}; z)$ by $a'_i = a_{i+1}$ for $0 \leq i \leq m-1$ and $\pi z' + a'_m = z + \sum_{i=1}^{m} q_i a_{m-i}$, which shows that the output is given by (9.2).

Implementation. Figure 9.1 illustrates a binary d-FCSR (for $d = 2$). Since addition in $\mathbb{Z}[\pi]$ is needed, it is slightly more convenient to break the addition into two parts. The part labeled Σ adds the $q_i a_{m-i}$ inputs as integers and outputs the result σ' according to its binary expansion. The remaining part is an adder in $\mathbb{Z}[\pi]$, which we now describe.

(This diagram involves a slight change of notation, with $z_i \in \{0,1\}$ now denoting the contents of the i-th memory cell, rather than the full value of the memory after the i-th iteration.) Each symbol Σ represents a full adder with 3 inputs, cascaded so as to form a ripple counter. With each clock cycle the current contents z of the memory is added to the integer σ' which is presented at the input to the adder according to its binary expansion. The result σ is returned to the memory (which involves modifying only the even numbered memory cells). Then the contents of the memory are shifted one step to the right, thus outputting the lowest order bit $\sigma \pmod{\pi}$ and retaining the higher order bits, $\sigma \text{ (div } \pi)$ (with the highest order bit, z_6 in the following example, set to 0).

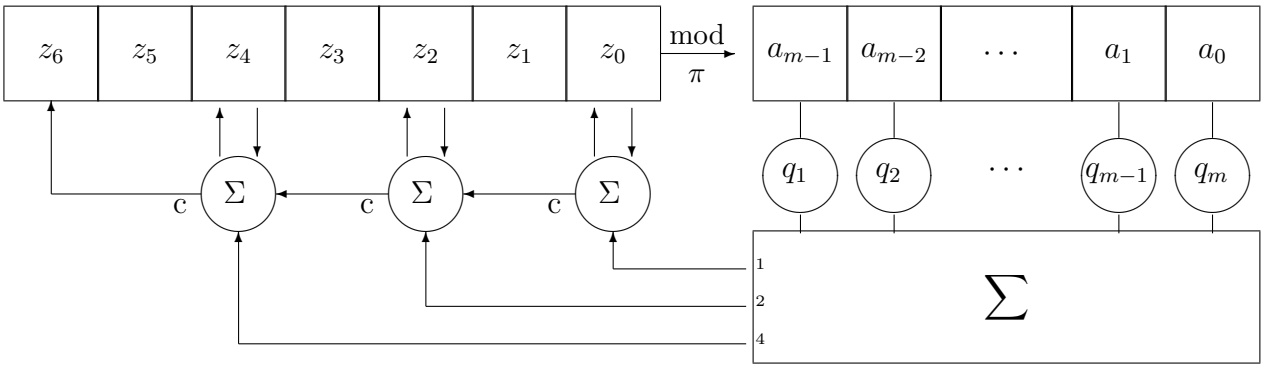


Figure 9.1: d-FCSR

Let $wt(q+1)$ denote the number of nonzero q's involved in the feedback. We see from Figure (9.1) (or from the change of state equations above) that the memory will decrease until $z_i = 0$ for all $i > d\log_2(wt(q+1)) + d$, so no memory overflow will occur provided the shift register is provided with memory cells $z_0, z_1, \ldots, z_s$ where $s \geq d\log_2(wt(q+1)) + d$. The deeper analysis of a d-FCSR is completely parallel to that of an FCSR, however it requires the π-adic formalism, developed in Section 5.5, which we now review in this special case.

Let $\mathbb{Z}_\pi$ be the ring of π-adic integers consisting of all formal power series in π,

$$\alpha = \sum_{i=0}^{\infty} a_i \pi^i \tag{9.3}$$

with $a_i \in \{0, 1\}$. Addition and multiplication are performed using the relation $\pi^d = 2$ whenever necessary. In particular, $\mathbb{Z}_\pi$ contains the 2-adic integers $\mathbb{Z}_2$. Since

$$-1 = 1 + \pi^d + \pi^{2d} + \pi^{3d} + \cdots$$

we see that $\mathbb{Z}_\pi$ also contains $\mathbb{Z}[\pi]$. In fact, $\mathbb{Z}_\pi$ contains all fractions $\alpha = u/q$ with $u, q \in \mathbb{Z}[\pi]$ provided that q is odd, (meaning that $q \equiv 1 \pmod{\pi}$) in which case we refer to equation (9.3) as "the" π-adic expansion of u/q, and we write $\mathbf{a} = a_0, a_1, \cdots = \mathbf{seq}_\pi(u/q)$. Such fractions are precisely the elements of $\mathbb{Z}_\pi$ whose π-adic expansions are eventually periodic. The following result (originally proven in [58]) is a special case of Theorem 8.5.1

Theorem 9.1.1. *Suppose an m-stage (Fibonacci) d-FCSR with connection integer q has initial state $(a_0, a_1, \ldots, a_{m-1}; z)$. Set $q_0 = -1$ and*

$$h = m\pi^m r - \sum_{k=0}^{m-1} \sum_{i=0}^{k} q_i a_{k-i} \pi^k$$

Then the output sequence $\mathbf{a} = a_0, a_1, \cdots$ *of the d-FCSR is the coefficient sequence for the* π*-adic expansion of the fraction* $\alpha = -h/q$*, that is,* $\mathbf{a} = \mathbf{seq}_\pi(-h/q)$*. It can also be expressed as*

$$a_i = \pi^{-i}h \pmod{q} \pmod{\pi}.$$

In Corollary 9.5.2 we prove that the (periodic) output sequences are also given by

$$a_i = b^{-i}A \pmod{N} \pmod{2}$$

for appropriate A, b, where N is the norm of the ideal $(q) \subset \mathbb{Z}[\pi]$. (See [63].) In this way maximal length sequences may be obtained when 2 is not necessarily primitive $\pmod{N}$.

9.2 General d-FCSRs

Let $N \geq 2$ and $d \geq 1$ be integers such that the polynomial $x^d - N$ is irreducible over the rational numbers $\mathbb{Q}$, as described in Section 5.5.c. Let $\pi \in \mathbb{C}$ be a root of this polynomial in an extension field of $\mathbb{Q}$. Here we consider AFSRs based on the ring $R = \mathbb{Z}[\pi]$ consisting of polynomials in π with integer coefficients. It is an integral domain in which every prime ideal is maximal. In fact it is an order (cf. Section 3.4.c), but not necessarily the maximal order, in its fraction field $F = \mathbb{Q}(\pi)$ of elements u/v with $u, v \in \mathbb{Z}[\pi]$. As remarked in Section 5.5.c, $F \cong \mathbb{Q}[\pi]$ because $1/\pi = \pi^{d-1}/N$. So

As described in Section 5.5.c, the set $S = \{0, 1, 2, \cdots, N-2\}$ is a complete set of representatives for the quotient $R/(\pi)$. Consequently the ring R_π (sometimes denoted $\mathbb{Z}_\pi$) of π-adic integers consists of formal power series $u = u_0 + u_1\pi + \cdots$ in π with coefficients $u_i \in S$. If $d = 1$ then $\pi = N$, $\mathbb{Z}[\pi] = \mathbb{Z}$, $\mathbb{Q}[\pi] = \mathbb{Q}$, and the π-adic integers are just the N-adic integers $\mathbb{Z}_N$.

By Proposition 5.5.5, the intersection $F \cap R_\pi$ can be described in two ways: as a subset of R_π it consists of formal power series $a = \sum_{i=0}^{\infty} a_i\pi^i$ (with $0 \leq a_i \leq N-1$) whose coefficient sequence $\mathbf{seq}_\pi(a) = (a_0, a_1, a_2, \ldots)$ is eventually periodic. As a subset of F it consists of fractions u/q (with $u, q \in R$) whose denominator $q = \sum_{i=0}^{d-1} q_i\pi^i$ (with $q_i \in \mathbb{Z}$) is coprime to π or equivalently, if q_0 is relatively prime to N. In summary, we have a diagram

$$\begin{array}{ccccc} F = \mathbb{Q}(\pi) & = & \mathbb{Q}[\pi] & \subset & F_\pi \\ \bigcup & & \bigcup & & \bigcup \\ R = \mathbb{Z}[\pi] & \subset & F \cap R_\pi & \subset & R_\pi \end{array}$$

Definition 9.2.1. *A d-FCSR is an AFSR based on this choice of* $(R = \mathbb{Z}[\pi], \pi, S)$*. A d-FCSR sequence is the (eventually periodic) output from a d-FCSR, that is,* $\mathbf{seq}_\pi(u/q)$ *where* $u \in R$*.*

As in Section 5.5.c the different embeddings $\sigma_1, \cdots, \sigma_d : F \to \mathbb{C}$ are determined by $\sigma_i(\pi) = \zeta^i\pi$ where $\zeta \in \mathbb{C}$ is a primitive dth root of unity. Then $|\sigma_i(\pi)| > 1$ and so Theorem 8.3.2 gives another proof that the output of any AFSR based on (R, π, S) is eventually periodic.

9.3 Relation between the norm and the period

As in Section 9.2 suppose that $x^d - N$ is irreducible over $\mathbb{Q}$ and that $\pi \in \mathbb{C}$ is a root of this polynomial. In our analysis a central role is played by the algebraic norm of the connection element. As in Theorem 3.4.4, the norm $\mathsf{N}_{\mathbb{Q}}^F(a) \in \mathbb{Q}$ of an element $a = \sum_{i-0}^{d-1} a_i \pi^i \in F$ is defined to be the product of the images of a under the embeddings σ_i of F in $\mathbb{C}$. Equivalently, it is the determinant of the linear transformation $f_a(b) = ab$ on the field F (when it is considered as a d-dimensional vector space over $\mathbb{Q}$). With respect to the basis $\{1, \pi, \pi^2, \ldots, \pi^{d-1}\}$ the matrix of f_a is given by

$$M_a = \begin{pmatrix} a_0 & Na_{d-1} & \cdots & Na_2 & Na_1 \\ a_1 & a_0 & \cdots & Na_3 & Na_2 \\ & & \vdots & & \\ a_{d-2} & a_{d-3} & \cdots & a_0 & Na_{d-1} \\ a_{d-1} & a_{d-2} & \cdots & a_1 & a_0 \end{pmatrix} \tag{9.4}$$

If $a \in R$, then $\mathsf{N}_{\mathbb{Q}}^F(a) \in \mathbb{Z}$ is an integer. For example, if $d = 2$, then

$$\mathsf{N}_{\mathbb{Q}}^F(a_0 + a_1\pi) = a_0^2 - Na_1.$$

If $d = 3$, then

$$\mathsf{N}_{\mathbb{Q}}^F(a_0 + a_1\pi + a2\pi^2) = a_0^3 + Na_1^3 + N^2a_2^3 - 3Na_0a_1a_2.$$

By Lemma 3.4.8, if $q \in R$ is a non-unit, then the absolute value $|\mathsf{N}_{\mathbb{Q}}^F(q)|$ is equal to the number of elements in the quotient ring $R/(q)$. The following lemma is used to determine the period.

Lemma 9.3.1. *Let $q \in R$. Then the element $\delta = |\mathsf{N}_{\mathbb{Q}}^F(q)|/q$ is in R.*

Proof. Let $q = \sum_{i=0}^{d-1} a_i \pi^i$ with $a_i \in \mathbb{Z}$. As in Section 9.2, let ζ be a primitive dth root of unity, so that

$$\delta = \pm \prod_{j=1}^{d-1} \sigma_j(q) = \pm \prod_{j=1}^{d-1} \sum_{i=0}^{d-1} a_i \zeta^{ij} \pi^i.$$

For any multi-index $I = (i_1, i_2, \ldots, i_{d-1})$ (where $0 \le i_j \le d-1$), let $\mu(I) = E = (e_0, e_1, \ldots, e_{d-1})$ where for each j, e_j is the number of occurrences of j among the components of I. It follows that

$$\delta = \pm \sum_{E} \left(\sum_{I:\mu(I)=E} \zeta^{\sum_{j=1}^{d-1} j i_j} \right) \prod_{i=0}^{d-1} a_i^{e_i} \pi^{ie_i}$$

where the first sum is over all multi-indices $E = (e_0, e_1, \ldots, e_{d-1})$ such that $0 \le e_j \le d-1$ and $\sum e_j = d-1$. For each such multi-index E and element $a \in \mathbb{C}$, let

$$c_E(a) = \sum_{I:\mu(I)=E} a^{\sum_{j=1}^{d-1} j i_j}.$$

Since the a_i are integers, to prove the lemma it suffices to show that $c_E(\zeta)$ is an integer for each E. Note that q and π have now left the picture.

The element $c_E(\zeta)$ lies in the field $\mathbb{Q}[\zeta]$, which is a Galois extension of the rational numbers. For an integer k, let τ_k be the homomorphism on $\mathbb{Q}[\zeta]$ induced by $\tau_k(\zeta) = \zeta^k$. The Galois group $\mathrm{Gal}(\mathbb{Q}[\zeta]/\mathbb{Q})$ consists of all τ_k such that k is relatively prime to d [83, p. 195].

We claim that $c_E(\zeta)$ is a rational number. By Galois theory, to show this it suffices to show that $c_E(\tau_k(\zeta)) = c_E(\zeta)$ for every τ_k in the Galois group of $\mathbb{Q}[\zeta]$ over $\mathbb{Q}$. Let $k\ell \equiv 1 \pmod{d}$. For any multi-index $I = (i_1, \ldots, i_{d-1})$, let $i'_j = i_{\ell j}$, and $I' = (i'_1, \ldots, i'_{d-1})$ (here the index ℓj in $i_{\ell j}$ is taken modulo d). The mapping from I to I' is a permutation of the multi-indices I with $\mu(I) = E$. Thus

$$\begin{aligned} c_E(\zeta^k) &= \sum_{I:\mu(I)=E} \zeta^{\sum_{j=1}^{d-1} k j i_j} \\ &= \sum_{I:\mu(I)=E} \zeta^{\sum_{j=1}^{d-1} j i_{\ell j}} \\ &= \sum_{I:\mu(I')=E} \zeta^{\sum_{j=1}^{d-1} j i'_j} \\ &= c_E(\zeta). \end{aligned}$$

This proves the claim.

Furthermore, $c_E(\zeta)$ is integral over $\mathbb{Z}$ (ζ is integral over $\mathbb{Z}$ since it is a root of the integral polynomial $x^d - 1$. The sum and product of integral elements are integral.) But the only elements of $\mathbb{Q}$ that are integral over $\mathbb{Z}$ are the integers. Thus $c_E(\zeta) \in \mathbb{Z}$, proving the lemma. □

Lemma 9.3.1 says that for any $q \in R$ the norm $\mathsf{N}_{\mathbb{Q}}^F(q)$ is divisible by q. It follows that the composition $\mathbb{Z} \to R \to R/(q)$ induces a well-defined ring homomorphism

$$\psi : \mathbb{Z}/(\mathsf{N}_{\mathbb{Q}}^F(q)) \to R/(q).$$

Proposition 9.3.2. *Suppose that $(q) = (r)^t$ for some $r \in R$ and $t \in \mathbb{Z}^+$, and that $\mathsf{N}_{\mathbb{Q}}^F(r)$ is (an ordinary, or "rational") prime. Then $(r) \subset R$ is a prime ideal, $\mathsf{N}_{\mathbb{Q}}^F(q) = \mathsf{N}_{\mathbb{Q}}^F(r)^t$, and the mapping $\psi : \mathbb{Z}/(\mathsf{N}_{\mathbb{Q}}^F(q)) \to R/(q)$ is an isomorphism.*

Proof. First consider the case $t = 1$. The ring homomorphism ψ is nontrivial since it takes 1 to 1. So its image is a subring of $\mathbb{Z}/(\mathsf{N}_{\mathbb{Q}}^F(q))$, which is a field, so ψ is surjective. Moreover the cardinality of $R/(q)$ is $|\mathsf{N}_{\mathbb{Q}}^F(q)|$ so ψ is an isomorphism.

Now we proceed by induction on t. For any s let us write

$$\psi = \psi_s : \mathbb{Z}/(\mathsf{N}_{\mathbb{Q}}^F(r^s)) \to R/(r^s).$$

This induces a mapping between short exact sequences

$$\begin{array}{ccccccccc}
0 & \longrightarrow & R/(r^{k-1}) & \xrightarrow{\cdot r} & R/(r^k) & \longrightarrow & R/(r) & \longrightarrow & 0 \\
 & & \uparrow{\scriptstyle \psi_{k-1}} & & \uparrow{\scriptstyle \psi_k} & & \uparrow{\scriptstyle \psi_1} & & \\
0 & \longrightarrow & \mathbb{Z}/(\mathsf{N}_{\mathbb{Q}}^F(r^{k-1})) & \xrightarrow{\cdot r} & \mathbb{Z}/(\mathsf{N}_{\mathbb{Q}}^F(r^k)) & \longrightarrow & \mathbb{Z}/(\mathsf{N}_{\mathbb{Q}}^F(r)) & \longrightarrow & 0
\end{array}$$

The horizontal maps are additive group homomorphisms giving short exact sequences. The vertical maps are ring homomorphisms. Since the vertical homomorphisms on the ends are isomorphisms by induction, the homomorphism in the middle is a group isomorphism as well. But since it is also a ring homomorphism, it is a ring isomorphism as well. Finally, the norm map is multiplicative so

$$\mathsf{N}_{\mathbb{Q}}^F(r^t) = (\mathsf{N}_{\mathbb{Q}}^F(r))^t.$$

□

Proposition 9.3.2 says two things:

- If $u \in R$ then there exists an integer $m \in \mathbb{Z}$ such that $m \equiv u \pmod{q}$. Moreover m may be chosen so that $0 \leq m < |\mathsf{N}_{\mathbb{Q}}^F(q)|$.
- If $a, b \in \mathbb{Z}$ are integers, then $a \equiv b \pmod{\mathsf{N}_{\mathbb{Q}}^F(q)}$ (as integers) if and only if $a \equiv b \pmod{q}$ (as elements of R).

9.4 Periodicity

As in Section 9.2, assume that $x^d - N$ is irreducible over $\mathbb{Q}$ and that $\pi \in \mathbb{C}$ is a root. In this section we identify the fractions $u/q \in F$ such that $\mathbf{seq}_\pi(u/q)$ is strictly periodic. Fix

$$q = q_0 + q_1\pi + \cdots + q_{d-1}\pi^{d-1} \in R$$

which is invertible in $R/(\pi)$. (Thus $q_0 \pmod{N}$ is invertible in $\mathbb{Z}/(N)$.) Let $\delta = |\mathsf{N}_{\mathbb{Q}}^F(q)|/q \in R$.

Lemma 9.4.1. *Let $u \in R$. Set $\delta u = z_0 + z_1\pi + \cdots + z_{d-1}\pi^{d-1}$, with $z_i \in \mathbb{Z}$. Then* $\mathbf{seq}_\pi(u/q)$ *is strictly periodic if and only if*

$$-|\mathsf{N}_{\mathbb{Q}}^F(q)| \le z_i \le 0.$$

Proof. We can uniquely write $\delta u = \sum_{i=0}^{d-1} z_i\pi^i$ with $z_i \in \mathbb{Z}$. Then in $\mathbb{Q}[\pi]$,

$$\frac{u}{q} = \frac{\delta u}{|\mathsf{N}_{\mathbb{Q}}^F(q)|} = \sum_{i=0}^{d-1} \frac{z_i}{|\mathsf{N}_{\mathbb{Q}}^F(q)|}\pi^i.$$

It follows that the π-adic expansion of u/q is the interleaving of the N-adic expansions of the $z_i/|\mathsf{N}_{\mathbb{Q}}^F(q)|$. Thus the former sequence is strictly periodic if and only if each of the latter sequences is strictly periodic. But the N-adic expansion of $z_i/|\mathsf{N}_{\mathbb{Q}}^F(q)|$ is strictly periodic if and only if

$$-|\mathsf{N}_{\mathbb{Q}}^F(q)| \le z_i \le 0. \qquad \square$$

In the next section we make (implicit) use of the vector space isomorphism

$$\tau : F \to \mathbb{Q}^d \tag{9.5}$$

given by $\tau(u_0 + u_1\pi + \cdots + u_{d-1}\pi^{d-1}) = (u_0, u_1, \ldots, u_{d-1})$. Then $\tau(R) = \mathbb{Z}^d$ which is a lattice in $\mathbb{Q}^d$. Suppose that $q \in R$ is invertible mod π. Define the half open parallelepiped for q to be

$$P = \left\{ \sum_{i=0}^{d-1} a_i q\pi^i : \quad a_i \in \mathbb{Q} \text{ and } -1 < a_i \le 0 \right\} \subset \mathbb{Q}[\pi]. \tag{9.6}$$

Also define the closed parallelepiped for q to be

$$P^c = \left\{ \sum_{i=0}^{d-1} a_i q\pi^i : \ a_i \in \mathbb{Q} \text{ and } -1 \le a_i \le 0 \right\} \subset \mathbb{Q}[\pi]. \tag{9.7}$$

The terminology refers to the fact that in equation (9.5) the image $\tau(P) \subset \mathbb{Q}^d$ is the parallelepiped spanned by the vectors $-q$, $-q\pi$, $-q\pi^2$, …, $-q\pi^{d-1}$. Define

$$\Delta = P \cap R \quad \text{ and } \Delta^c = P^c \cap R$$

to be the sets of lattice points in the half open and closed parallelepipeds. Applying the projection to $R/(q)$ gives a mapping

$$\phi : \Delta \subset R \to R/(q). \tag{9.8}$$

Theorem 9.4.2. *If $q \in R$ is invertible mod π, then the mapping ϕ is a one to one correspondence. For any element $u \in R$, the coefficient sequence of*

$$\frac{u}{q} = \sum_{i=0}^{\infty} a_i \pi^i \tag{9.9}$$

is strictly periodic if and only if $u \in \Delta^c$. If $u \in \Delta$, then

$$a_i = \overline{q}^{-1} \left(\pi^{-i} u \pmod{q} \right) \pmod{\pi}. \tag{9.10}$$

This last equation needs a bit of explanation. Using the fact that ϕ is a one to one correspondence, for any element $a \in R/(q)$ define

$$a \pmod{\pi} = \phi^{-1}(a) \pmod{\pi} \in R/(\pi) = \mathbb{Z}/(N), \tag{9.11}$$

meaning that we first lift a to Δ and then reduce modulo π. The image of π in $R/(q)$ is invertible, so the product $\pi^{-i}u$ is defined in $R/(q)$. Hence

$$\pi^{-i} u \pmod{\pi} = \phi^{-1}(\pi^{-i} u \pmod{q}) \pmod{\pi}$$

is defined in $\mathbb{Z}/(n)$. Moreover, the element

$$\overline{q} = q \pmod{\pi} \in R/(\pi) = \mathbb{Z}/(N)$$

is also invertible; by equation (9.4) it coincides with $q_0 \pmod{N}$. So the product in equation (9.10) makes sense in $\mathbb{Z}/(N)$. As usual, the mapping defined by equation (9.11), $R/(q) \to \mathbb{Z}/(N)$ is not a ring homomorphism and in fact it depends on the choice $\Delta \subset R$ of a complete set of representatives for the elements of $R/(q)$.

Proof. First we show that the above mapping $\phi : \Delta \to R/(q)$ is one to one and onto. The idea is that existence and uniqueness of representatives modulo q in P corresponds to existence and uniqueness modulo $|\mathsf{N}_{\mathbb{Q}}^F(q)|$ in the (half-open) hypercube

$$\delta P = \left\{ \sum_{i=0}^{d-1} a_i \pi^i : \ a_i \in \mathbb{Q} \text{ and } \ -|\mathsf{N}_{\mathbb{Q}}^F(q)| < a_i \leq 0 \right\}$$

(the image of P under the linear map $F \to F$ which is given by multiplication by δ). First we show that ϕ is onto. For any $w \in \mathbb{Q}$ define $\hat{w} \in \mathbb{Q}$ to be the unique rational number such that

$$-|\mathsf{N}_{\mathbb{Q}}^F(q)| < \hat{w} \leq 0$$

and such that

$$w - \hat{w} = e|\mathsf{N}_{\mathbb{Q}}^{F}(q)|$$

for some integer e. If w is an integer then so is $\hat{w}$. Now let $b \in R$ be arbitrary, and write

$$\delta b = \sum_{i=0}^{d-1} b_i \pi^i$$

for some integers b_i. Set

$$c = \sum_{i=0}^{d-1} \hat{b}_i \pi^i.$$

We claim that $\delta^{-1}c \in \Delta$ and that it is congruent to b modulo q. In fact,

$$\begin{aligned} b - \delta^{-1}c &= \delta^{-1}(\delta b - c) \\ &= \delta^{-1} \sum_{i=0}^{d-1} (b_i - \hat{b}_i)\pi^i \\ &= \delta^{-1} \sum_{i=0}^{d-1} e_i |\mathsf{N}_{\mathbb{Q}}^{F}(q)| \pi^i \\ &= q \sum_{i=0}^{d-1} e_i \pi^i \end{aligned}$$

(for some integers e_i). This last expression is in R, hence $\delta^{-1}c \in R$. But $\delta^{-1}c$ is also in P, so $\delta^{-1}c \in \Delta$ and (again by the last expression) it is congruent to b modulo (q). To prove ϕ is one to one, suppose that $b, b' \in \Delta$ are congruent modulo q. Then $b - b' = cq$ for some $c \in R$, so

$$\delta b - \delta b' = c\delta q = c|\mathsf{N}_{\mathbb{Q}}^{F}(q)|.$$

But both δb and $\delta b'$ lie in the half-open hypercube δP whose sides all have length $|\mathsf{N}_{\mathbb{Q}}^{F}(q)|$. So no coordinate can differ by as much as $|\mathsf{N}_{\mathbb{Q}}^{F}(q)|$, hence $c = 0$ and $b = b'$.

Next suppose that $b \in R$. By Lemma 9.4.1, b/q has a periodic π-adic expansion if and only if

$$\delta b \in \left\{ \sum_{i=0}^{d-1} a_i \pi^i \middle| \; a_i \in \mathbb{Q} \text{ and } \; -|\mathsf{N}_{\mathbb{Q}}^{F}(q)| \leq a_i \leq 0 \right\} = \delta P^c,$$

which holds if and only if $b \in P^c$.

To verify equation (9.10), first observe that reducing equation (9.9) modulo π gives

$$a_0 \equiv \overline{q}^{-1} u \pmod{\pi}.$$

The sequence $\{a_1, a_2, \ldots\}$ is also strictly periodic and corresponds to a π-adic integer

$$\frac{u'}{q} = \sum_{i=0}^{\infty} a_{i+1} \pi^i$$

with the same denominator q (and with some numerator $u' \in \Delta$). Then

$$\frac{u'}{q}\pi = \frac{u}{q} - a_0$$

or $u'\pi = u - a_0 q$. Reducing modulo q gives

$$u' \equiv \pi^{-1} u \pmod{q}.$$

Equation (9.10) now follows by induction. □

Remark. A *face* $U \subset P$ of the parallelepiped P in equation (9.6) is the set of points obtained by setting some of the coefficients a_j equal 0. The set of lattice points $\Delta \cap U$ in any face corresponds under equation (9.8) to an additive subgroup of $R/(q)$. If $\mathsf{N}_{\mathbb{Q}}^F(q)$ is prime then there are no additive subgroups other than $\{0\}$. Hence, if $\mathsf{N}_{\mathbb{Q}}^F(q)$ is prime, then all the nonzero elements of Δ lie in the interior of the parallelepiped. Let

$$P_0 = \left\{ \sum_{i=0}^{d-1} v_i q \pi^i \mid \; v_i \in \mathbb{Q} \text{ and } \; -1 < v_i < 0 \right\} \subset \mathbb{Q}[\pi] \tag{9.12}$$

be the open parallelepiped associated with q, and define

$$\Delta_0 = P_0 \cap \mathbb{Z}[\pi].$$

Corollary 9.4.3. *If $\mathsf{N}_{\mathbb{Q}}^F(q)$ is prime and u/q is strictly periodic, then either q divides u or $u \in \Delta_0$.*

The following lemma is used in the proof of Proposition 17.3.3 to analyze arithmetic correlations.

Lemma 9.4.4. *Suppose $b \in \Delta$. Then the representative of $-b$ modulo q in Δ is*

$$c = -q\frac{\pi^d - 1}{\pi - 1} - b.$$

Proof. The element b is in Δ if and only if each π-adic coordinate z_i of δb is between $-|\mathsf{N}_{\mathbb{Q}}^F(q)|$ and 0. If this holds, then

$$-|\mathsf{N}_{\mathbb{Q}}^F(q)| < -|\mathsf{N}_{\mathbb{Q}}^F(q)| - z_i < 0.$$

But $-|\mathsf{N}_{\mathbb{Q}}^F(q)| - z_i$ is the ith π-adic coordinate of $-|\mathsf{N}_{\mathbb{Q}}^F(q)|(1+\pi+\pi^2+\cdots+\pi^{d-1}) - \delta b = \delta c$. Thus c is in Δ and is congruent to $-b$ modulo q. □

9.5 Elementary description of d-FCSR sequences

In some cases it is possible to describe the d-FCSR sequences in elementary terms (without reference to algebraic number fields). As in Section 9.2, suppose that $\pi^d = N$ and that $x^d - N$ is irreducible over $\mathbf{Q}$. Let $q \in R$ and suppose that

- q is invertible in $\mathbb{Z}_\pi$,
- d is relatively prime to $\mathsf{N}_{\mathbb{Q}}^F(q)$,
- $(q) = (r^t)$ for some $r \in R$, and
- $\mathsf{N}_{\mathbb{Q}}^F(r)$ is (a "rational", or ordinary) prime.

Recall that all these conditions are satisfied if N is prime and $\mathsf{N}_{\mathbb{Q}}^F(q)$ is prime.

Let $Q = |\mathsf{N}_{\mathbb{Q}}^F(q)| \in \mathbb{Z}$. By Proposition 9.3.2 we have a natural isomorphism of rings, $\psi : \mathbb{Z}/(N) \to R/(q)$. Each of these rings has a canonical set of representatives: let

$$T = \{0, -1, -2, \ldots, -(Q-1)\} \subset \mathbb{Z}$$

and let $\Delta \subset R$ be the parallelepiped as defined as in Section 9.4. Then T is a complete set of representatives for $\mathbb{Z}/(Q)$ and Δ is a complete set of representatives for $R/(q)$. By abuse of notation we also write $\psi : T \to \Delta$ for the resulting one-to-one correspondence. Set

$$\delta = \sum_{i=0}^{d-1} s_i \pi^i.$$

Theorem 9.5.1. *Let $u \in \Delta$. Then in $\mathbb{Z}/(N)$,*

$$(\overline{q}^{-1}u) \pmod{\pi} = \overline{\mathsf{N}_{\mathbb{Q}}^F(q)}^{-1}(s_0\psi^{-1}(u) \pmod{Q}) \pmod{N}.$$

The left side of this equation makes sense since the image $\overline{q} \in R/(\pi)$ of $q \in R$ is invertible. On the right side of the equation, the element $\psi^{-1}(u) \in \mathbb{Z}/(Q)$ is multiplied by $s_0 \pmod{Q}$ then lifted to the set T of representatives. Then the result is reduced modulo N. The image $\overline{\mathsf{N}_{\mathbb{Q}}^F(q)} \in \mathbb{Z}/(N)$ of $\mathsf{N}_{\mathbb{Q}}^F(q) \in \mathbb{Z}$ is invertible, so the product on the right hand side makes sense. This result may be best explained in terms of the following two diagrams. Both the upper and lower squares in the

first diagram commute. Although the upper square in the second diagram commutes, the lower square does not. Theorem 9.5.1 supplies "correction factors" which are needed in order to make it commute.

$$\begin{array}{ccc} \mathbb{Z}/(Q) & \xrightarrow{\psi} & R/(q) \\ {\scriptstyle \text{mod Q}}\uparrow & & \uparrow{\scriptstyle \text{mod q}} \\ \mathbb{Z} & \longrightarrow & R \\ {\scriptstyle \text{mod N}}\downarrow & & \downarrow{\scriptstyle \text{mod } \pi} \\ \mathbb{Z}/(N) & \xrightarrow{\cong} & R/(\pi) \end{array} \qquad \begin{array}{ccc} \mathbb{Z}/(Q) & \xrightarrow{\psi} & R/(q) \\ \uparrow & & \uparrow \\ T & \xrightarrow{\psi} & \Delta \\ \downarrow & & \downarrow \\ \mathbb{Z}/(N) & \xrightarrow{\cong} & R/(\pi) \end{array}$$

Replacing u with $\overline{\pi}^{-i}u$ in Theorem 9.5.1 gives the following corollary which describes the d-FCSR sequence completely in terms of operations on ordinary integers.

Corollary 9.5.2. *Let $u \in \Delta$ and let $e = \psi^{-1}(\pi) \in \mathbb{Z}/(Q)$. Write*

$$\frac{u}{q} = \sum_{j=0}^{\infty} a_j \pi^j$$

for the π-adic expansion of the fraction u/q. Then the coefficient sequence $\mathbf{seq}_\pi(u/q) = a_0, a_1, \ldots$ is strictly periodic and moreover

$$a_j = \overline{\mathsf{N}}_{\mathbb{Q}}^{F}(q)^{-1}(Ab^i \pmod{Q}) \pmod{N}$$

where $b = e^{-1}$ and

$$A = s_0\psi^{-1}(u) \in \mathbb{Z}/(Q).$$

The proof of theorem 9.5.1 occupies the next subsection.

9.5.a Proof of theorem 9.5.1

Continue with the same notation as in Section 9.4, that is,

$$\pi^d = N, \quad q \in R, \quad Q = |\mathsf{N}_{\mathbb{Q}}^{F}(q)|, \quad \delta = \mathsf{N}_{\mathbb{Q}}^{F}(q)/q = \sum_{i=0}^{d-1} s_i \pi^i, \quad \text{and } e = \psi^{-1}(\pi) \in \mathbb{Z}/(Q).$$

Let $\zeta \in \mathbb{C}$ be a primitive d-th root of unity and let U denote any one of the following rings:

$$\mathbb{Z}, \quad R, \quad \mathbb{Z}[\zeta], \quad \mathbb{Z}[\pi, \zeta].$$

The integer $Q \in \mathbb{Z}$ generates an ideal $(Q)_U$ in the ring U. If $a, b \in \mathbb{Z}$ we write

$$a \equiv b \pmod{Q} \text{ in } R$$

to mean that $a - b \in (Q)_U$.

Lemma 9.5.3. *For any $a, b \in \mathbb{Z}$ we have*

$$a \equiv b \pmod{Q} \text{ in } U \Longleftrightarrow a \equiv b \pmod{Q} \text{ in } \mathbb{Z}.$$

Proof. Both π and ζ are integral over $\mathbb{Z}$ so each $c \in U$ is integral over $\mathbb{Z}$. If $a \equiv b \pmod{Q}$ in U, then $a - b = cQ$ for some $c \in UR$. Then $c = (a-b)/Q$ is a rational number which is integral over $\mathbb{Z}$, hence $c \in \mathbb{Z}$. □

Lemma 9.5.4. *The inclusion $\mathbb{Z}[\zeta] \to \mathbb{Z}[\pi, \zeta]$ induces a ring isomorphism*

$$\Psi : \mathbb{Z}[\zeta]/(Q) \cong \mathbb{Z}[\pi, \zeta]/(q).$$

Proof. The proof is similar to that of Proposition 9.3.2. Since $q|Q$ in $\mathbb{Z}[\pi, \zeta]$ the mapping Ψ is well defined. To see that it is surjective it suffices to show that π is in the image of ϕ. But $\pi \equiv e \pmod{q}$ in R, hence also in $\mathbb{Z}[\pi, \zeta]$. That is, $\Psi(e) = \pi$. Finally, the cardinality of both rings is $|Q|^{\phi(d)}$ (where ϕ is the Euler totient function), hence Ψ is also injective. □

Lemma 9.5.5. *Let $u \in R$ and set*

$$\delta u = \sum_{i=0}^{d} z_i \pi^i.$$

Then

$$z_k \pi^k \equiv z_0 \pmod{q}$$

(in R) and

$$z_k e^k \equiv z_0 \pmod{Q}$$

(in $\mathbb{Z}$) for all $k = 0, 1, \ldots, d$.

Proof. Let $\hat{z}_j$ denote the j-th Fourier coefficient of the finite sequence of complex numbers

$$(z_0, z_1\pi, z_2\pi^2, \ldots z_{d-1}\pi^{d-1}),$$

which may be considered to be a function $f : \mathbb{Z}/(d) \to \mathbf{C}$. In other words, if $\zeta \in \mathbf{C}$ denotes a primitive d-th root of unity, then

$$\hat{z}_j = \sum_{i=0}^{d-1} z_i \pi^i \zeta^{ij}.$$

Evidently, the Fourier coefficient $\hat{z}_j$ lies in the ring $\mathbf{Z}[\pi, \zeta]$. The homomorphism $\sigma_j : R \to \mathbf{C}$ may be extended to this ring by setting $\sigma_j(\zeta) = \zeta$. Then

$$\sigma_j(\delta) = \prod_{i \neq j} \sigma_i(q)$$

which implies that, in the ring $\mathbb{Z}[\pi,\zeta]$, the element $\sigma_j(\delta)$ is divisible by q for $j = 1, 2, \ldots, d-1$. Hence

$$\begin{aligned}\hat{z}_j &= \sum_{i=0}^{d-1} z_i \pi^i \zeta^{ij} \\ &= \sigma_j(\delta u) \\ &= \sigma_j(\delta)\sigma_j(u) \\ &\equiv 0 \pmod{q}.\end{aligned}$$

The Fourier inversion formula gives

$$\begin{aligned}dz_k\pi^k &= \textstyle\sum_{j=0}^{d-1} \hat{z}_j \zeta^{-kj} \\ &\equiv \hat{z}_0 \pmod{q},\end{aligned}$$

which is independent of k. Since d is invertible mod $\mathsf{N}_{\mathbb{Q}}^{F}(q)$ it is also invertible in $R/(q)$, meaning that $da = 1 + yq$ for some $a, y \in R$. Hence d is also invertible in $\mathbb{Z}[\pi,\zeta]/(q)$, which shows that

$$z_k\pi^k \equiv z_0\pi^0 \pmod{q} \text{ in } \mathbb{Z}[\pi,\zeta]$$

for all k. We need to show the same holds in R. However,

$$\begin{aligned}& z_k\pi^k \equiv z_0 \pmod{q} \text{ in } \mathbb{Z}[\pi,\zeta] \\ \Rightarrow\ & z_k e^k \equiv z_0 \pmod{q} \text{ in } \mathbb{Z}[\pi,\zeta] \\ \Rightarrow\ & z_k e^k \equiv z_0 \pmod{N} \text{ in } \mathbb{Z}[\zeta] \\ \Rightarrow\ & z_k e^k \equiv z_0 \pmod{N} \text{ in } \mathbb{Z} \\ \Rightarrow\ & z_k e^k \equiv z_0 \pmod{q} \text{ in } R \\ \Rightarrow\ & z_k\pi^k \equiv z_0 \pmod{q} \text{ in } R\end{aligned}$$

by Lemma (9.5.4), Lemma (9.5.3), and Proposition (9.3.2). □

Proof of theorem 9.5.1. As in the previous section set

$$\delta u = \sum_{k=0}^{d-1} z_k\pi^k \quad \text{and} \quad \delta = \sum_{k=0}^{d-1} s_k\pi^k.$$

Lemma 9.5.5 gives $\delta u \equiv dz_0 \pmod{q}$ and $\delta \equiv ds_0 \pmod{q}$, hence $dz_0 \equiv ds_0 u \pmod{q}$. Since d is invertible modulo q, this implies $z_0 \equiv s_0 u \pmod{q}$. The isomorphism $\psi^{-1} : R/(q) \to \mathbb{Z}/(Q)$ is the identity on integers, so we obtain

$$z_0 = s_0\psi^{-1}(u) \text{ in } \mathbb{Z}/(Q).$$

Now lift these elements to the set $T \subset \mathbb{Z}$ of representatives of $\mathbb{Z}/(Q)$ and reduce modulo N to obtain

$$z_0 \pmod N = (s_0\psi^{-1}(u) \pmod Q) \pmod N.$$

But $z_0 \pmod N = \delta u \pmod \pi \in \mathbb{Z}/(N)$. That is,

$$(\delta u) \pmod \pi = (s_0\psi^{-1}(u) \pmod Q) \pmod N.$$

Multiplying by $\overline{\mathsf{N}_{\mathbb{Q}}^F(q)}^{-1} \in \mathbb{Z}/(N)$ gives

$$(\overline{q}^{-1}u) \pmod \pi = \overline{\mathsf{N}_{\mathbb{Q}}^F(q)}^{-1}(s_0\psi^{-1}(u) \pmod Q) \pmod N$$

where $\overline{q} \in R/(\pi)$ is the image of q, and where the right hand side is interpreted as follows: first compute $s_0\psi^{-1}(u) \in \mathbb{Z}/(Q)$, then lift this to the set T of representatives, then reduce modulo N, then multiply by $\overline{\mathsf{N}_{\mathbb{Q}}^F(q)}^{-1}$. This completes the proof of Theorem 9.5.1. □

9.6 An Example

Consider a binary d-FCSR (meaning that $N = 2$) with $\pi^2 = 2$ (so $d = 2$) and with connection element $q \in R$ which is invertible in $\mathbb{Z}_\pi$. Let $Q = |\mathsf{N}_{\mathbb{Q}}^F(q)|$. Then

$$\overline{\mathsf{N}_{\mathbb{Q}}^F(q)}^{-1} \pmod 2 = 1$$

and Corollary 9.5.2 says that the strictly periodic portions of the d-FCSR sequence will be described by

$$a_j = Ab^j \pmod Q \pmod 2$$

for certain $A, b \in \mathbb{Z}/(Q)$. Let us take $q = 5 + 2\pi$. Then $\mathsf{N}_{\mathbb{Q}}^F(q) = 17$ which is prime, so the parallelogram contains 16 elements in its interior. It is illustrated in Figure 1.

The isomorphism

$$\psi : \mathbb{Z}[\pi]/(q) \to \mathbb{Z}/(17)$$

maps π to $m = 6$. Hence $m^{-1} = 3 \in \mathbb{Z}/(17)$ which is primitive, so we obtain a maximal length output sequence. The element δ equals $5-2\pi$ hence $s_0 = 5$. Each element in $\mathbb{Z}[\pi]/(q)$ has a unique representative u in the parallelogram; these representatives are listed in the second column of Table 2. Each element in $\mathbb{Z}/(17)$ has a unique representative in the set $T = \{-1, -2, \ldots, -16\}$; these representatives h are listed in the third column. The correspondence between the second and third column is given by Theorem 9.5.1. That is, $h = s_0\psi^{-1}(u)$ (since $\overline{q} \equiv \overline{\mathsf{N}}_{\mathbb{Q}}^F(q) \equiv 1 \pmod 2$). The fourth column (which is the d-FCSR sequence under consideration) is the third column modulo 2, and it coincides with the second column modulo π.

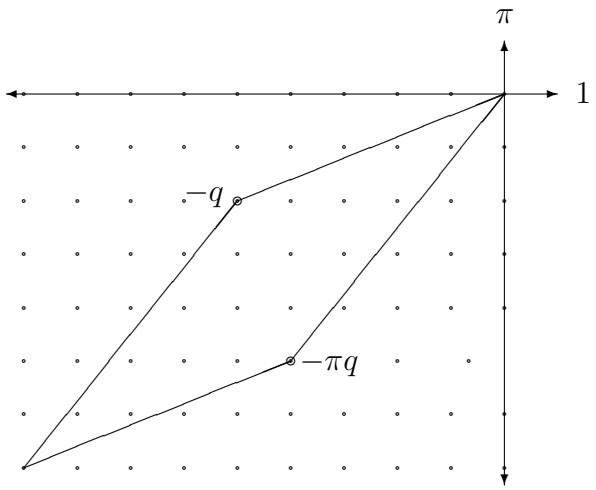


Figure 9.2: Parallelogram for $q = 5 + 2\pi$.

i	$\pi^{-i} \pmod q$	$5 \cdot 6^{-i} \pmod{17}$	output
0	$-4 - 2\pi$	-12	0
1	$-2 - 2\pi$	-2	0
2	$-2 - \pi$	-6	0
3	$-1 - \pi$	-1	1
4	$-3 - 3\pi$	-3	1
5	$-5 - 4\pi$	-9	1
6	$-6 - 5\pi$	-10	0
7	$-5 - 3\pi$	-13	1
8	$-5 - 5\pi$	-5	1
9	$-7 - 5\pi$	-15	1
10	$-7 - 6\pi$	-11	1
11	$-8 - 6\pi$	-16	0
12	$-6 - 4\pi$	-14	0
13	$-4 - 3\pi$	-8	0
14	$-3 - 2\pi$	-7	1
15	$-4 - 4\pi$	-4	0

Table 9.1: Model states for $q = 5 + 2\pi$.

9.7 Exercises

1. Let $\pi^3 = 2$ and $R = \mathbb{Z}[\pi]$. Consider the connection element $q = 1 - \pi + 3\pi^2$.
   (a) Show that the multiplicative group of $\mathbb{Z}[\pi]/(q)$ is cyclic.
   (b) Find a maximum period periodic output sequence from a 3-FCSR with connection element q.
   (c) Find the elementary representation of this sequence as in Section 9.5.
2. Let $\pi^2 = 7$ and $q = 8 + \pi$. Find all u for which u/q has a strictly periodic π-adic expansion.
3. Prove that under the hypotheses of Section 9.5, if π is primitive modulo q, then the second half of a period of a binary d-FCSR sequence with connection element q is the bitwise complement of the first half.

Chapter 10 Galois Mode, Linear Registers, and Related Circuits

In this chapter we examine the issues involved in efficient implementation of feedback shift registers. In general feedback shift registers can be implemented either in hardware or in software. The choice depends on the application, the resources available, and the set of output symbols. Binary sequence generators (which are typically used in high speed cryptographic applications) can often be implemented very efficiently in hardware, often with only a small number of gates, but less efficiently in software – they either waste space or require slow packing and unpacking of words. Generators with outputs in a set $\{0, 1, \cdots, N-1\}$ with N requiring 8 or more bits (which are typically used in pseudo-Monte Carlo and quasi-Monte Carlo simulation) are often slow in hardware but relatively fast in software where they can make use of built-in word operations.

10.1 Galois mode LFSRs

Throughout this section we fix a commutative ring R with unit. A *Galois mode linear feedback shift register* or *modular shift register* of length m over R, with "multipliers" $q_1, q_2, \cdots, q_m \in R$ is a sequence generator (cf. Definition 5.1.2) whose state is an element

$$\mathbf{s} = (h_0, h_1, \cdots, h_{m-2}, h_{m-1}) \in R^m = \Sigma,$$

whose output is $\mathbf{out}(\mathbf{s}) = a_0 = h_0$, and whose state change operation τ is given by

$$(h_0, h_1, \cdots, h_{m-2}, h_{m-1}) \longrightarrow (h_1 + q_1 h_0, h_2 + q_2 h_0, \cdots, h_{m-1} + q_{m-1} h_0, q_m h_0). \tag{10.1}$$

It is convenient to think of a Galois mode LFSR as the circuit depicted in Figure 6.1.

As with LFSRs, because we like to think of bits as flowing out to the right, the order of the components a_i in the diagram is the reverse of the order when we write the state as a vector. When thinking of LFSRs as physical devices as in the figure, we sometimes list the components of the state in descending order of their indices. We write $\boxed{h_{m-1}}\boxed{h_{m-2}}\boxed{\cdots}\boxed{h_1}\boxed{h_0}$ for the *machine state* to distinguish the two notations.

The connection polynomial for this shift register is defined to be

$$q(x) = -1 + q_1 x + q_2 x^2 + \cdots + q_m x^m.$$

We assume that $q_m \neq 0$ (otherwise periodic states of the shift register may be described using only those cells to the right of the first nonzero coefficient q_s.) It is not immediately apparent how to describe the output of this machine: the cell contents $(h_{m-1}, \cdots, h_1, h_0)$ are not simply shifted to the right; they are modified with each clock tick. So the following fact may come as a surprise.

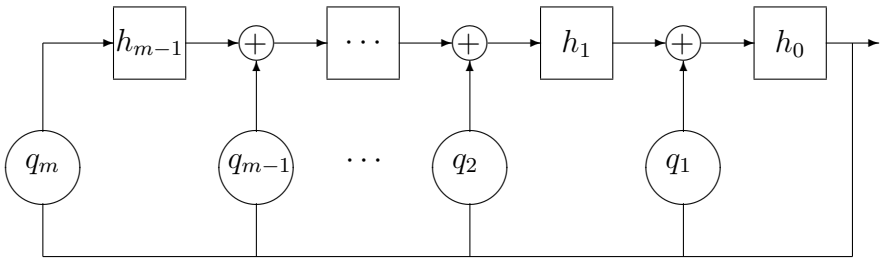


Figure 10.1: Galois LFSR.

Theorem 10.1.1. *The output sequence* $\mathbf{a} = a_0, a_1, a_2, \cdots$ *of the Galois LFSR satisfies the linear recurrence*

$$a_n = q_1 a_{n-1} + q_2 a_{n-2} + \cdots + q_m a_{n-m} \tag{10.2}$$

for all $n \geq m$. *If* q_m *is invertible in the ring* R *then* $\mathbf{a} = \mathbf{seq}(-h(x)/q(x))$ *is the (coefficient sequence of the) power series expansion of the rational function* $-h(x)/q(x) \in R[[x]]$, *that is,*

$$\frac{-h(x)}{q(x)} = a_0 + a_1 x + a_2 x^2 + \cdots \in R[[x]] \tag{10.3}$$

where $h(x) = h_0 + h_1 x + \cdots + h_{m-1} x^{m-1}$ *corresponds to the initial loading. By Theorem 6.4.1 the output sequence is strictly periodic. The output sequence may also be described by the equation (cf. Section 6.6.a),*

$$a_n = x^{-n} h \pmod{q} \pmod{x} \tag{10.4}$$

for all $n \geq 0$.

Proof. In the ring $R[x]/(q)$, using equation (6.15) calculate

$$\begin{aligned} x^{-1}h(x) &= (q_1 + q_2 x + \cdots + q_m x^{m-1})h_0 + h_1 + h_2 x + \cdots + h_{m-1} x^{m-2} \\ &= (q_1 h_0 + h_1) + (q_2 h_0 + h_2)x + \cdots + (q_{m-1} h_0 + h_{m-1}) x^{m-2} + q_m h_0 x^{m-1} \end{aligned}$$

Since this polynomial has degree $< m$ it is the unique representative of $x^{-1}h$ in $R[x]/(q)$ and we see from equation (10.1) that it exactly corresponds to the change of state operation. The constant term is

$$a_1 = q_1 h_0 + h_1 = x^{-1} h(x) \pmod{q} \pmod{x}.$$

It follows by induction that $a_n = x^{-n} h \pmod{q} \pmod{x}$ which proves equation (10.4). By running the displayed equations in the proof of Theorem 6.6.2 backward, it follows that the output sequence $a_0, a_1, \cdots$ satisfies the linear recurrence (10.2). Finally, Proposition 6.6.1 says that

$$\sum_{i=0}^{\infty} a_i x^i = -h(x)/q(x).$$

□

The Galois LFSR is sometimes preferred in applications because the additions are performed in parallel by separate adders. Another difference between the Galois LFSR and the Fibonacci LFSR concerns the relation between the initial loading and the numerator $-h(x)$ of the rational function (10.3): in the Galois case the initial loading $h_0, h_1, \cdots, h_{m-1}$ exactly corresponds to $h(x) = h_0 + h_1 x + \cdots + h_{m-1}x^{m-1}$. In the Fibonacci case the initial loading $a_0, a_1, \cdots, a_{m-1}$ corresponds to the numerator $f(x)$ in equation (6.5).

10.2 Division by $q(x)$ in $R[[x]]$

Suppose $q(x) = -1 + q_0 x + \cdots + q_m x^m$ is a polynomial and $h(x)$ is either a polynomial or a power series, and we wish to calculate $-h(x)/q(x)$ as a power series. According to Theorem 10.1.1 this can be accomplished by setting up a Galois LFSR with connection polynomial q and by initializing the cells with the coefficients $h_0, h_1, \cdots$. If $\deg(h) = N > \deg(q)$ then we need a Galois LFSR with cells extending to the left of the final coefficient q_m.

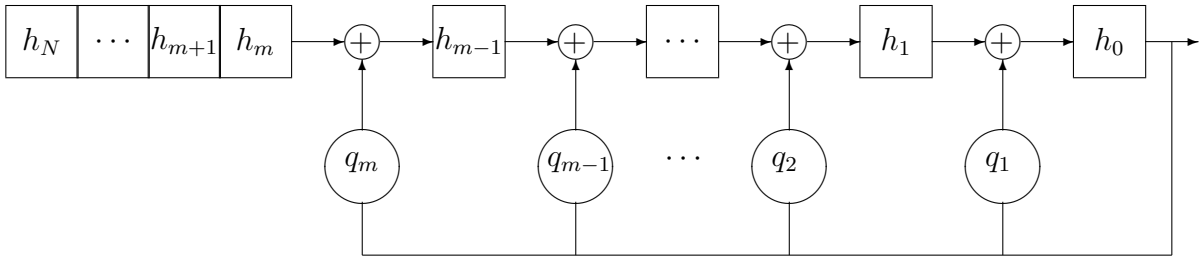


Figure 10.2: Division by $q(x)$.

It is not necessary to store the coefficients h_j (for $j \geq m$) in separate memory cells: they may be presented to the leftmost cell of the LFSR as a signal that is synchronized with the LFSR clock. In this way polynomials of arbitrarily high degree (or even power series) $h(x)$ may be divided by $q(x)$. An alternative to initializing the LFSR cells is to calculate the quotient $-h(x)x^m/q(x)$ (where $\deg(q) = m$), for this may be achieved by initializing the LFSR cells to zero and presenting the whole polynomial (or power series) $h(x)$ at the leftmost cell. Of course, after m clock cycles, the state illustrated in Figure 10.2 will be reached.

10.3 Galois mode FCSR

As in Chapter 7 we fix a positive integer N and consider a feedback with carry shift register for which the cells contain entries from $R = \mathbb{Z}/(N)$. The register has fixed "multipliers"

$q_1, q_2, \cdots, q_m \in R$. A state consists of cells $a_0, a_1, \cdots, a_{m-1} \in \{0, 1, \cdots, N-1\} \subset R$ (separated by adders) and memory (or "carry") integers, $c_1, c_2, \cdots, c_m$. The state is written as $s = (a_0, a_1, \cdots, a_{m-1}; c_1, c_2, \cdots, c_m)$.

The change of state can be described as follows. Calculate $\sigma_m = c_m + q_m a_0$ and $\sigma_j = c_j + a_j + q_j a_0$ in the j-th adder for $1 \leq j < m$. The new values are then

$$a'_{j-1} = \sigma_j \pmod N \quad c'_j = \sigma_j \text{ (div } N) = \lfloor \sigma_j / N \rfloor$$

so that

$$N c'_j + a'_{j-1} = c_j + a_j + q_j a_0 \ (1 \leq j < m) \quad \text{and} \quad N c'_m + a'_{m-1} = c_m + q_m a_0. \tag{10.5}$$

This procedure is illustrated in the schematic of Figure 10.3 in which the symbol Σ denotes an integer adder $\pmod N$ with carry.

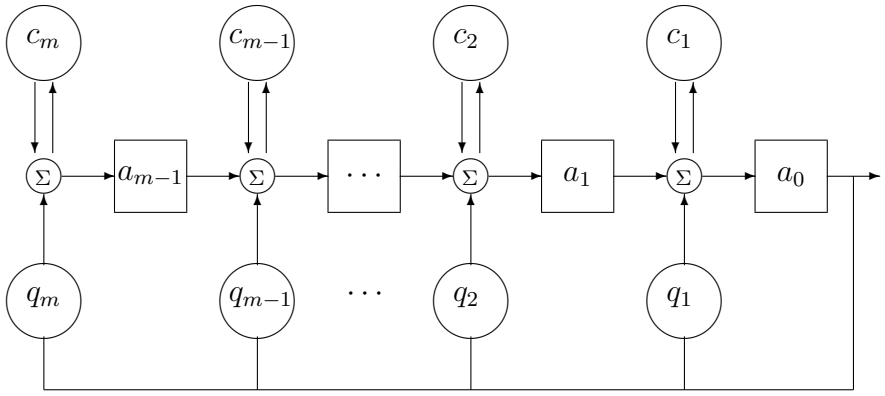


Figure 10.3: Galois FCSR.

Define the *connection integer*

$$q = -1 + q_1 N + q_2 N^2 + \cdots + q_m N^m. \tag{10.6}$$

It is relatively prime to N so it is invertible in the N-adic integers $\mathbb{Z}_N$.

Theorem 10.3.1. *Suppose an m-stage Galois FCSR with connection integer q has initial state $(a_0, a_1, \cdots, a_{m-1}; c_1, c_2, \cdots, c_m)$. Set*

$$h = a_0 + (a_1 + c_1)N + \cdots + (a_{m-1} + c_{m-1})N^{m-1} + c_m N^m. \tag{10.7}$$

Then the output sequence $b_0, b_1, \cdots$ of the FCSR is $\mathbf{seq}_N(-h/q)$*, that is,*

$$\frac{-h}{q} = b_0 + b_1 N + b_2 N^2 + \cdots \in \mathbb{Z}_N$$

and $b_i = N^{-i}h \pmod q \pmod N$*. Conversely, if* $\mathbf{b} = b_0, b_1, \cdots$ *is a periodic sequence (with* $b_i \in \mathbb{Z}/(N)$*) with corresponding* N*-adic number* $\beta = -h/q$ *then* q *is the connection integer of a Galois FCSR that generates the sequence* $\mathbf{b}$*.*

Proof. As the FCSR iterates its state change, it goes through an infinite sequence of states,

$$\{\mathbf{s}(n) = (a_0(n), a_1(n), \cdots, a_{m-1}(n); c_1(n), c_2(n), \cdots, c_m(n) : n \geq 0\}.$$

Each $\mathbf{s}(n)$ can be considered as an initial state. Thus for each $n \geq 0$ there is an associated output sequence $(b_0(n), b_1(n), b_2(n) \cdots) = (a_0(n), a_0(n+1), a_0(n+2) \cdots)$. Each such sequence has an associated N-adic number

$$B(n) = \sum_{i=0}^{\infty} b_i(n) N^n = \sum_{i=0}^{\infty} a_0(n+i) N^n. \tag{10.8}$$

We also have a sequence of integers

$$\begin{aligned} h(n) &= a_0(n) + (a_1(n) + c_1(n))N + \cdots + (a_{m-1}(n) + c_{m-1}(n))N^{m-1} + c_m(n)N^m \\ &= \sum_{i=0}^{m} (a_i(n) + c_i(n))N^i, \end{aligned}$$

where we take $c_0(n) = a_m(n) = 0$ for convenience. We show that for all n, $h(n) + qB(n) = 0$. By multiplying out the formulas (10.6) and (10.8) for q and $B(n)$ we see that $h(n) + qB(n) \equiv a_0(n) - b_0(n) \pmod N = 0$. That is, each $h(n) + qB(n)$ is divisible by N. Now, compute

$$\begin{aligned} h(k) + qB(k) &= \sum_{i=0}^{m} (a_i(k) + c_i(k))N^i + q(a_0(k) + NB(k+1)) \\ &= a_0(k) + \sum_{i=1}^{m} (a_{i-1}(k+1) + Nc_i(k+1) - q_i a_0(k))N^i + qa_0(k) + qNB(k+1) \\ &= \sum_{i=1}^{m} (a_{i-1}(k+1) + Nc_i(k+1))N^i + qNB(k+1) \\ &= N(h(k+1) + qB(k+1)). \end{aligned}$$

This allows us to prove by induction that for all i, N^i divides every $h(k) + qB(k)$. Thus $h(k) + qB(k) = 0$ as we claimed. □

As with the Galois LFSR, the additions are performed in parallel in a Galois FCSR. The relationship between the numerator h and the initial loading of the FCSR is simpler and more direct than for the Fibonacci mode FCSR. If $q_i = 0$, then the memory cell c_i may be omitted because no carry can occur at that location. If $q_m = 1$, then c_m may be omitted for the same reason, and it follows from Corollary 7.2.2 that (under these conditions $q_m = 1, c_m = 0$) the output sequence will be strictly periodic.

Proposition 10.3.2. *If the memory c_j at the j-th cell is in the range $0 \le c_j \le N-1$, then it will remain within that range forever. If $c_j \le N$, then it will remain $\le N$ forever. If $c_j > N$, then it will drop exponentially until it falls within the range $0 \le c_j \le N$, where it will remain thereafter.*

Proof. The state change equation gives

$$c'_j = \left\lfloor \frac{c_j + a_j + q_j a_0}{N} \right\rfloor \le \left\lfloor \frac{c_j + (N-1) + (N-1)^2}{N} \right\rfloor = \left\lfloor \frac{c_j}{N} \right\rfloor + N - 1.$$

The result follows. □

There are two stable states, the all-zero state (or *bottom*) state, and the *top* state where $c_i = q_i$ $(1 \le i \le m-1)$, $c_m = q_m - 1$, and $a_i = N-1$ $(0 \le i \le m-1)$. For the top state, $h = q$ so the output from the top state is the all $N-1$'s sequence. That is, it is the N-adic representation of the integer -1. For any periodic state other than the top state, if $q_i = 0$ then the memory cell c_i will eventually drop to 0 and it will remain there forever. So the periodic states must satisfy $c_i = 0$ whenever $q_i = 0$. Let us say that a state satisfying this condition is an *admissible* state. Then the collection of admissible states is in one to one correspondence with the elements of $\mathbb{Z}/(q)$, the correspondence being given by equation (10.7).

10.4 Division by q in the N-adic numbers

Let $q = -1 + q_1 N + \cdots + q_m N^m$ and let $h = h_0 + h_1 + \cdots + h_N$ be the base-N expansion of two integers. We wish to calculate the quotient $-h/q$ in the N-adic numbers. Just as in Section 10.2 this may be accomplished using the Galois mode FCSR of Figure 10.3 by initializing the cell and memory values to 0, and then presenting the coefficient sequence for h as a signal (synchronized with the FCSR clock) to the leftmost adder. After m steps the FCSR will acquire the state shown in Figure 10.4, so by Theorem 10.3.1 the output sequence will be the N-adic expansion of $-h/q$.

10.5 Galois mode d-FCSR

In this section we describe Galois mode circuitry for a binary d-FCSR, so this is the situation of an AFSR based on the ring $R = \mathbb{Z}[\pi]$ with $\pi^d = 2$ and $S = \{0, 1\}$. A general (non-binary)

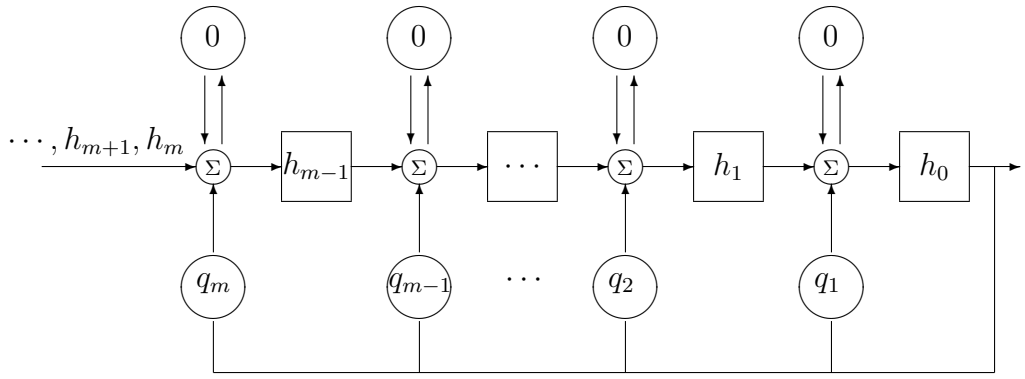


Figure 10.4: Division by q in $\mathbb{Z}_N$.

Galois d-FCSR (see Section 9.2) may be constructed similarly. In the Galois architecture for a binary d-FCSR, the carried bits are delayed $d-1$ steps before being fed back, so the output of the memory or "carry" cell c_i is fed into the register cell a_{i+d-2}. (Recall that the register cells are numbered starting from a_0.) If there are r feedback multipliers $q_1, \ldots, q_r$ and r carry cells $c_1, \ldots, c_r$ then $r+d-1$ register cells $a_0, \ldots, a_{r+d-2}$ are evidently needed since c_r will feed into a_{r+d-2}. This is illustrated in Figure (10.5) for $d=2$. If $d \geq 3$ the situation is more complicated and a some number t of additional memory cells $c_{r+1}, \ldots, c_{r+t}$ are needed, which feed into t additional register cells $a_{r+d-1}, \ldots, a_{r+t+d-2}$. It is not at all obvious at first glance whether the amount t of extra memory can be chosen to be finite without incurring a memory overflow during the operation of the shift register. However we show later (in Theorem 10.5.1) that this is indeed the case and henceforth we suppose that t has been chosen as described there, sufficiently large so as to avoid any memory overflow.

Suppose a general (Galois) d-FCSR is initially loaded with given values (a_0, a_1, ..., $a_{r+t+d-2}$; c_1, c_2, ..., c_{r+t}). The register operates as follows. (To simplify notation, set $q_j = 0$ for $j \geq r+1$, set $c_j = 0$ for $j \leq 0$ and also for $j \geq r+t+1$, and set $a_j = 0$ for $j \geq r+t+d-1$.) For each j (with $1 \leq j \leq r+t+d-1$) form the *integer* sum $\sigma_j = a_0 q_j + a_j + c_{j-d+1}$; it is between 0 and 3. The new values are given by $a'_{j-1} = \sigma_j \pmod 2$ and $c'_j = \sigma_j \text{ (div 2)}$. That is,

$$2c'_j + a'_{j-1} = a_0 q_j + a_j + c_{j-d+1} \quad \text{for} \quad 1 \leq j \leq r+t+d-1 \tag{10.9}$$

(Note, for example, that these equations say $a'_{r+t+d-2} = c_{r+t}$.)

Assume $q_r \neq 0$ and define the connection element

$$q = -1 + \sum_{i=1}^{r} q_i \pi^i \in \mathbb{Z}[\pi]. \tag{10.10}$$

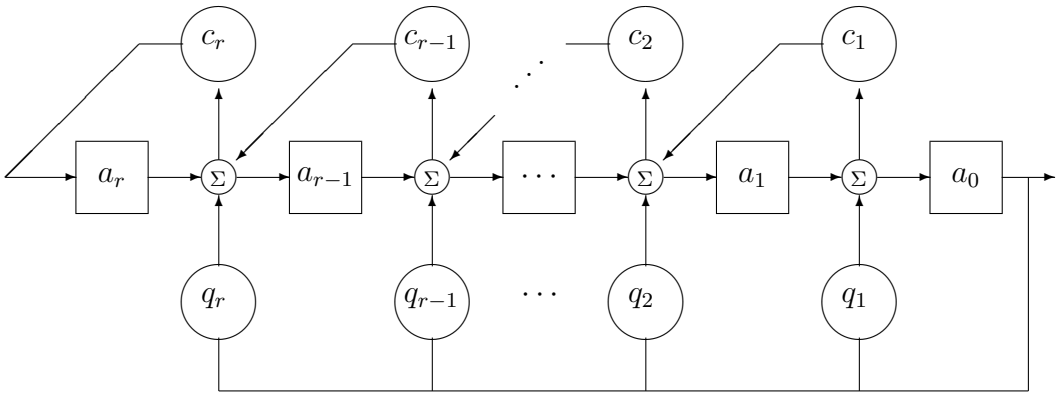


Figure 10.5: Galois 2-FCSR

Theorem 10.5.1. *Suppose a Galois d-FCSR with connection element $q \in \mathbb{Z}[\pi]$ has initial state $(a_0, \ldots, a_{r+t+d-2}; c_1, c_2, \ldots, c_{r+t})$. Set*

$$h = \sum_{i=0}^{r+t+d-2} (a_i + c_{i-d+1})\pi^i \in \mathbb{Z}[\pi]. \tag{10.11}$$

If t satisfies

$$t > \frac{2}{d} - \log_2(2^{1/d} - 1),$$

then no memory overflow will occur, and the output sequence $\mathbf{b} = (b_0, b_1, \ldots)$ coincides with $\mathbf{seq}_\pi(-h/q)$, the π-adic expansion of the fraction $-h/q \in \mathbb{Z}_\pi$. Conversely if $\mathbf{b} = (b_0, b_1, \ldots)$ is a periodic sequence with corresponding π-adic number $\beta = -h/q$ and if $q + 1$ is nonnegative, then q is the connection element of a Galois d-FCSR which generates the sequence $\mathbf{b}$. The change of state is given by $h \mapsto \pi^{-1}h$. Hence the output sequence is:

$$b_i = \pi^{-i}h \pmod{q} \pmod{\pi}.$$

Proof. First we make our model mathematically simpler by thinking of the state as two doubly infinite sequences $(\cdots, a_{-1}, a_0, a_1, \cdots; \cdots, c_{-1}, c_0, c_1, c_2, \cdots)$, We assume that initially

$$c_i = a_j = 0 \text{ for } i \leq 0 \text{ and } j \leq -1 \tag{10.12}$$

and we update the state using equation (10.9) for all j. Then it is immediate that equation (10.12) holds throughout the infinite execution of the d-FCSR. It can also be seen that if initially

$$c_i = a_j = 0 \text{ for } i \geq r + t \text{ and } j \geq r + t + d - 1, \tag{10.13}$$

then equation (10.13) holds throughout the infinite execution of the d-FCSR [61]. We omit the details and assume this is the case.

As the d-FCSR iterates its state change, it goes through an infinite sequence of states,

$$\mathbf{s}(n) = (\cdots, a_{-1}(n), a_0(n), a_1(n), \cdots; \cdots, c_{-1}(n), c_0(n), c_1(n), c_2(n), \cdots), n \geq 0.$$

Each $\mathbf{s}(n)$ can be thought of as an initial state. Thus for each $n \geq 0$ there is an associated output sequence $(b_0(n), b_1(n), b_2(n)\cdots) = (a_0(n), a_0(n+1), a_0(n+2)\cdots)$. Each such sequence has an associated π-adic number

$$B(n) = \sum_{i=0}^{\infty} b_i(n)\pi^n = \sum_{i=0}^{\infty} a_0(n+i)\pi^n. \tag{10.14}$$

We also have a sequence of elements of R

$$h(n) = \sum_i (a_i(n) + c_{i-d+1}(n))\pi^i = \sum_{i=0}^{r+t+d-1} (a_i(n) + c_{i-d+1}(n))\pi^i.$$

We show that for all n, $h(n) + qB(n) = 0$. By multiplying out the formulas (10.10) and (10.14) for q and $B(n)$ we see that $h(n) + qB(n) \equiv a_0(n) - b_0(n) \pmod{\pi} = 0$. That is, each $h(n) + qB(n)$ is divisible by π. Now, compute

$$\begin{aligned}
h(k) + qB(k) &= \sum_i (a_i(k) + c_{i-d+1}(k))\pi^i + q(a_0(k) + \pi B(k+1)) \\
&= a_0(k) + \sum_{i\geq 1} (a_{i-1}(k+1) + 2c_i(k+1) - q_i a_0(k))\pi^i + qa_0(k) + q\pi B(k+1) \\
&= \sum_{i\geq 1} (a_{i-1}(k+1) + 2c_i(k+1))\pi^i + q\pi B(k+1) \\
&= \pi \sum_{i\geq 0} a_i(k+1)\pi^i + \sum_{i\geq 1} c_i(k+1)\pi^{i+d} + q\pi B(k+1) \\
&= \pi \sum_{i\geq 0} a_i(k+1)\pi^i + \pi \sum_{i\geq d} c_{i-d+1}(k+1)\pi^i + q\pi B(k+1) \\
&= \pi(h(k+1) + qB(k+1)).
\end{aligned}$$

This allows us to prove by induction that for all i, π^i divides every $h(k) + qB(k)$. Thus $h(k) + qB(k) = 0$ as we claimed. □

Corollary 10.5.2. *There is a mapping from the set of periodic states of the Fibonacci d-FCSR with connection element q to the set of periodic states of the Galois d-FCSR with connection element q so that corresponding states produce the same output.*

Our understanding of the Galois d-FCSR architecture still leaves much to be desired. We do not know how to intrinsically characterize the strictly periodic states. We don't even know how to find a class of "admissible" states for which the output is strictly periodic (as we did in the case of the FCSR). We do not know an optimal estimate on the amount of memory needed for the d-FCSR (except in the case $d = 2$). We do not know how to describe the contents of each cell as a function of time.

10.6 Linear registers

The relationship between Fibonacci and Galois modes of LFSRs, FCSRs, and d-FCSRs can be made clearer by considering various generalizations. We start by considering the state of an LFSR or the basic register of an FCSR or d-FCSR to be a vector over some ring (the ring containing the output symbols for an LFSR and $\mathbb{Z}$ for an FCSR or d-FCSR).

Consider first an LFSR over a ring R in Fibonacci mode with connection polynomial $q(x) = \sum_{i=1}^{m} q_i x^i - 1 \in R[x]$. As in Section 6.2, the state change is given by multiplication by a matrix

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & & \vdots & & \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ q_m & q_{m-1} & \cdots & q_2 & q_1 \end{pmatrix}, \tag{10.15}$$

where the state is treated as a column vector. Similarly, the state change of the associated Galois mode representation is given by multiplication by the matrix

$$\begin{pmatrix} q_1 & 1 & 0 & \cdots & 0 \\ q_2 & 0 & 1 & \cdots & 0 \\ & & \vdots & & \\ q_{m-1} & 0 & \cdots & 0 & 1 \\ q_m & 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{10.16}$$

Both modes can be generalized by choosing an arbitrary $m \times m$ matrix M over R and defining the state change operation $\mathbf{s} \to M\mathbf{s} = \mathbf{s}'$. Furthermore, the output function can be replaced by an arbitrary linear function of the state, $\mathbf{s} \to \mathbf{v} \cdot \mathbf{s}$ where $\mathbf{v}$ is a length m vector over R and $\cdot$ is the inner product. We call such a sequence generator a *linear feedback register* or *LFR* of length m. A special case of this construction, called a *ring LFSR* was treated by Mrugaski, Rajski, and Tyszer [149].

Theorem 10.6.1. *Suppose R is a field. If $\mathbf{a} = a_0, a_1, \cdots$ is the output sequence from a length m linear feedback register over R based on matrix M and vector $\mathbf{v}$, then $\mathbf{a}$ is also generated by a LFSR with connection polynomial* $\det(I - xM)$.

Proof. For $n = 0, 1, \cdots =$, let $\mathbf{s}(n) = (s_0(n), \cdots, s_{m-1}(n))$ denote the state at time n. Let $S(x, k) = (S_0(x, k), \cdots, S_{m-1}(x, k)) = \sum_{n=0}^{\infty} \mathbf{s}(n + k)x^n$ be the generating function of the state starting at time k. It suffices to show that $\det(I - xM)S(x, 0)$ is a vector of polynomials of degree less than m.

For every $k \geq 0$ we have the relation

$$S(x, k) = \mathbf{s}(k) + xS(x, k + 1) = \mathbf{s}(k) + xMS(x, k).$$

Thus

$$(I - xM)S(x, k) = \mathbf{s}(k)$$

is a vector of scalars.

Recall from linear algebra that for any square matrix L over any ring, there is a matrix $adj(L)$, the *adjugate* of L, satisfying $adj(L)L = \det(L)I$, and that $adj(L)$ can be expressed as a polynomial in the entries of L of degree less than the dimension of L (Cramer's rule). It follows that

$$\det(I - xM)S(x, k) = adj(I - xM)\mathbf{s}(k)$$

is a vector of polynomials of degree less than m as we wanted. □

Thus in some sense LFRs are no more powerful than LFSRs. However, it may be possible to find an LFR that has a more efficient implementation — e.g., fewer total taps so that fewer operations are needed in a software implementation, or a smaller maximum row Hamming weight so that update time is smaller in a hardware implementation.

A similar generalization can be made of FCSRs. As with Galois mode FCSRs, we need a vector of carries. First we choose a nonnegative integer N and let $T = \{0, 1, \cdots, N-1\}$. The state consists of a vector $\mathbf{s}$ of elements of T, representing the main register, and a vector $\mathbf{c} = (c_0, c_1, \cdots, c_{m-1})$ of integers, representing the carry. The state change is determined by an $m \times m$ integer matrix M and the output is determined by an integer vector $\mathbf{v}$. The state change is given by

$$(\mathbf{s}, \mathbf{c}) \rightarrow (M\mathbf{s} + \mathbf{c} \pmod N, M\mathbf{s} + \mathbf{c} \text{ (div } N)) = (\mathbf{s}', \mathbf{c}'),$$

and the output is $(\mathbf{v} \cdot \mathbf{s}) \pmod N$. We call the resulting sequence generator an *N-ary feedback with carry register (FCR)* of length m. This model was studied by Arnault, Berger, Lauradoux, Minier, and Pousse in a special case [2].

Theorem 10.6.2. *If $\mathbf{a} = a_0, a_1, \cdots$ is generated by a length m N-ary FCR based on matrix M and vector $\mathbf{v}$, then $\mathbf{a}$ is also generated by an N-ary FCSR with connection integer* $\det(I - NM)$.

Proof sketch. The proof is similar to the proof of Theorem 10.6.1. With notation analogous to the notation above, for every $k \geq 0$ we have the relation

$$S(N,k) = \mathbf{s}(k) + NS(N,k+1).$$

We also have

$$\mathbf{s}(n+k+1) + N\mathbf{c}(n+k+1) = M\mathbf{s}(n+k) + \mathbf{c}(n+k).$$

Thus

$$S(N,k+1) = \sum_{n=0}^{\infty} \mathbf{s}(n+k+1)N^t = \sum_{n=0}^{\infty} (M\mathbf{s}(n+k) + \mathbf{c}(n+k) - N\mathbf{c}(n+k+1))N^t = MS(N,k) + \mathbf{c}(k)$$

It follows that

$$\det(I - NM)S(N,k) = adj(I - NM)\mathbf{s}(k)$$

is an integer vector. □

Again, it may be possible to find an FCR that has a more efficient implementation

10.7 Exercises

1. Describe a Galois LFSR whose output is the Fibonacci sequence 1,1,2,3,5,8,

2. Consider the linear register in Figure 10.6 that combines the Galois feedback and the Fibonacci feedback.

   a. Show that this register has matrix representation

$$\begin{pmatrix} p_1 & 1 & 0 & 0 \\ p_2 & 0 & 1 & 0 \\ p_3 & 0 & 0 & 1 \\ p_4 + r_4 & r_3 & r_2 & r_1 \end{pmatrix}$$

   and that its characteristic polynomial is

$$x^4 - x^3(p_1 + r_1) - x^2(r_2 + p_2 - r_1p_1) \\ -x(r_3 + p_3 - r_2p_1 - r_1p_2) - (r_4 + p_4 - r_3p_1 - r_2p_2 - r_1p_3).$$

   b. Show that the output sequence is an m-sequence provided this polynomial is primitive.

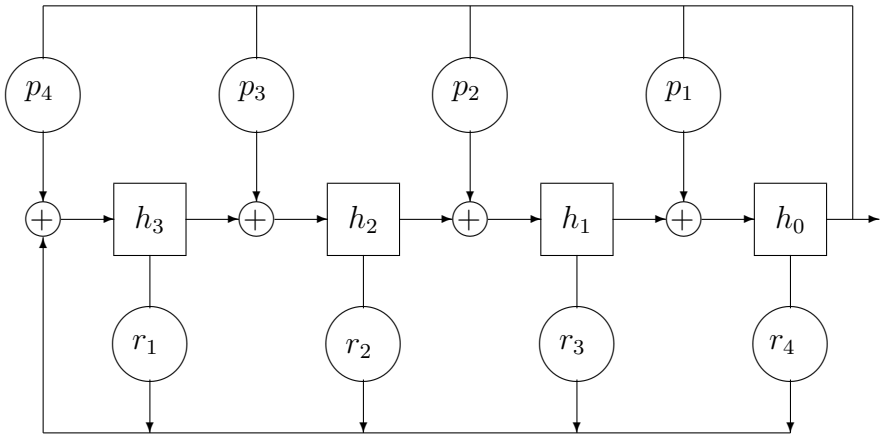


Figure 10.6: Galois/Fibonacci LFSR.

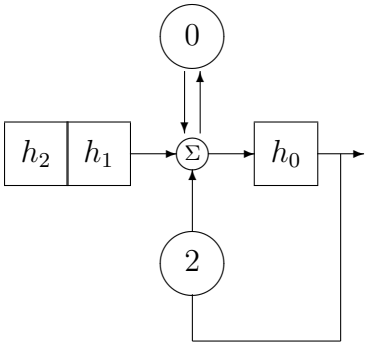


Figure 10.7: Division by 5 in $\mathbb{Z}_3$.

3. Consider a 1-stage Galois FCSR over $\mathbb{Z}/(3)$ with $q = -1+2\cdot 3 = 5$. Let $h = 1+2\cdot 3+2\cdot 3^2 = 25$. As in Section 10.4 imagine h being used as an input signal to the FCSR, by adding two "phantom" cells, so that after running the register for one step it attains the state shown in Figure 10.7. Show that the output of the register is the sequence $1, 1, 2, 2, 2, \cdots$ and verify that this is the 3-adic expansion of -5.

4. a. Use a computer algebra program such as Mathematica or Maple to find all the maximal length binary FCSRs of length 128 having 4 taps. In other words, find all the prime numbers p with $2^{128} \leq p+1 \leq 2+2^2+\cdots+2^{128}$ such that 2 is a primitive root modulo p and so that

the dyadic weight of $p+1$ is four, i.e. $p = -1 + 2^a + 2^b + 2^c + 2^d$. (There are fewer than 200 such primes.)

b. Take the first of the FCSRs that you found in the previous exercise and, starting with the "impulse" initial state = $\boxed{0}\boxed{0}\boxed{0}\boxed{\cdots}\boxed{0}\boxed{1}$ consider the state vector to be the binary expansion of a 128-bit integer. List the (decimal equivalent) of the first 100 such integers that are generated by your FCSR.

5. Let R be a ring. Generalize the notion of a linear feedback register to an affine feedback register over R by allowing state change operations of the form

$$\mathbf{s} \to M\mathbf{s} + \mathbf{w},$$

where M is an $m \times m$ matrix over R and $\mathbf{w}$ is a constant vector over R. If R is a field and $\mathbf{a}$ is an output sequence from this affine feedback register, then prove that there is a LFSR of length at most $m+1$ that outputs $\mathbf{a}$.

Part III

Pseudo-Random and Pseudo-Noise Sequences

Chapter 11 Measures of Pseudorandomness

Golomb's randomness postulates. In [53], [54], S. Golomb proposed three desirable criteria that one might ask of a binary pseudorandom sequence **a**.

1. It should be balanced,
2. it should have the run property (Section 11.2.b), and
3. it should have an ideal autocorrelation function.

These concepts are described below. Since 1967, new applications have created an enormous demand for pseudorandom sequences that exhibit additional, more sophisticated randomness properties. Some of these requirements are known to be incompatible with others. Currently, for any given application, one generally draws up a list of required randomness properties and then goes about trying to find or to design a pseudorandom sequence that meets these requirements. Although Golomb's list looks rather minimal by today's standards, it is an amazing fact that there are still only a handful of known techniques for constructing sequences with all three of these properties. In this section we will describe some of the most common measures of randomness.

11.1 Why pseudo-random?

Random numbers (in one sense or another) have applications in computer simulation, Monte Carlo integration, cryptography, randomized computation, radar ranging, and other areas. In each case you need a sequence of numbers (or of bits) that "appears" to be "random", yet is repeatable. Of course these are contradictory requirements. If we know the sequence beforehand, then it is not random. So instead, we settle for sequences that are "pseudo-random", meaning that they behave in the same way that a truly random sequence would behave, when subject to various tests (usually statistical tests), but they are generated by a specific repeatable rule. In other areas, such as spread spectrum communication, we want sequences that pass specific statistical tests that allow us to draw specific conclusions about the systems that use them.

Repeatability is important in the case of spread-spectrum communications and cryptography where a receiver must know the same sequence as a sender in order to decode a message. But it is also essential for simulations, Monte Carlo integration, and sometimes even randomized computation. Suppose we are simulating traffic flow along a highway, by having cars enter the highway at random times and places. We use a truly random generator (which counts cosmic rays or perhaps measures activity on the internet) to create this "arriving traffic" data. We are hoping to alleviate a bottleneck that develops at a certain spot, say, by widening the road or by eliminating a bend

in the road. After making these changes to the highway, we run the simulation again and discover the bottleneck has disappeared. Is this apparent success the result of our changes to the road, or is it an artifact of the pattern of arrival of new traffic? We can only be certain if we can run both simulations with the same pattern of arriving traffic.

We can consider a variety of measures of randomness. These include statistical measures such as large period, balance, uniform distribution of blocks, chi-square, autocorrelation and others, and security measures such as linear complexity or linear span, N-adic complexity, nonlinearity, and resilience. In some cases statistical tests for families of sequences must be considered, such as pairwise shifted cross-correlations.Typically these generators are chosen so as to give sequences with extremely long periods, say, 10^{100}, which makes it impossible to perform these tests on the actual sequence of numbers. Instead, one hopes to be able predict or theoretically compute the outcome of the test based on the knowledge of how the sequence was constructed. There is often a tradeoff — in order to pass many tests it may be necessary to make the sequence generators very complex, making it harder to analyze the the sequences with respect to the randomness measures.

Depending on the application, some of these measures may be more important than others. For spread spectrum applications, correlation coefficients are vital. For cryptographic applications, complexity measures with respect to classes of sequence generators give some estimate as to the difficulty of breaking a given cipher. For Monte Carlo integration, certain statistical measures (such as the discrepancy) provide precise bounds on errors in integration.

It is essentially hopeless to generate pseudo-random sequences that pass all statistical tests. The pseudo-random generators themselves provide statistical tests that the sequences cannot pass. Short of this, in complexity theoretic cryptography we typically want sequence generators that pass all polynomial time computable statistical tests. There are various constructions of such sequence generators, all depending on some believable but unconfirmed complexity theoretic assumptions. In general such generators sacrifice speed to a degree that is unacceptable for the applications we have in mind, and we do not consider them in this book.

It is common to speak of a "random" sequence or other object, but we must be clear what we mean by this. There is no set of sequences that can be identified as the random ones. Rather, when we say that the random sequence has some statistical property, we are making a statement about the expectation of a statistical measure.

A further problem is that the very properties that are taken to me a sequence is random are typically uncommon. For example, we typically ask that a binary sequence of even length n have equal numbers of 0s and 1s. But the probability that a sequence has this property is

$$\frac{\binom{n}{n/2}}{2^n} \sim \sqrt{\frac{2}{\pi n}},$$

so such sequences is are quite special.

11.2 Sequences based on an arbitrary alphabet

Suppose $\mathbf{a} = a_0, a_1, \cdots$ is a sequence with period T (so $a_T = a_0$) whose elements a_i are taken from some finite alphabet A. A *block* $b = (b_0, b_1, \cdots, b_{k-1})$ of length k is an ordered sequence of k elements, $b_i \in A$. An *occurrence* of the block b in (a single period of) the sequence $\mathbf{a}$ is an index $i \leq T-1$ such that $(a_i, a_{i+1}, \cdots, a_{i+k-1}) = b$. (In the parlance of cryptography, an occurrence of a block b is called a *k-gram*.) A *run* of length k is a block of k consecutive identical symbols that is not contained in a longer block of consecutive symbols.

11.2.a Distribution of blocks.

What is the expected number $E[P(k, \mathbf{b})]$ of occurrences of a given block $\mathbf{b}$ of length k in sequences $\mathbf{a}$ of period T? There are $|A|^T$ possible such sequences. The number $N(\mathbf{b}, i)$ of occurrences of b starting at position i among all sequences is $|A|^{T-k}$ since choosing the k values for the block $\mathbf{b}$ leaves $T-k$ values free. Thus the total number of occurrences of $\mathbf{b}$ among all sequences is $T|A|^{T-k}$. So the average number of occurrences of $\mathbf{b}$ in any one sequence is

$$E[P(k, \mathbf{b})] = \frac{T}{|A|^k}.$$

Let us say the sequence $\mathbf{a}$ is *equidistributed to order r* if, for every k $(1 \leq k \leq r)$ and for every block $\mathbf{b}$ of length k, the number $N(\mathbf{b})$ of occurrences of $\mathbf{b}$ within a single period of $\mathbf{a}$ satisfies

$$\left\lfloor T/|A|^k \right\rfloor \leq N(\mathbf{b}) \leq \left\lceil T/|A|^k \right\rceil.$$

A sequence is equidistributed only with low probability: with high probability some blocks will occur more often and some blocks will occur less often. However for many purposes, such as sampling, equidistribution of blocks is a desirable property. In cryptography, if an external observer knows the extent to which a sequence is not equidistributed, she can often exploit this knowledge to predict the sequence on the basis of partial information.

The periodic sequence $\mathbf{a}$ is *balanced* if it is equidistributed to first order ($r = 1$), meaning that each symbol in the alphabet occurs either $\lfloor T/|A| \rfloor$ or $\lceil T/|A| \rceil$ times in a single period of $\mathbf{a}$. Now suppose that $\mathbf{a}$ is a periodic sequence whose symbols a_i are in a finite group G. A related notion of balance arises by considering a nontrivial character χ of the group G.

Definition 11.2.1. *If $\mathbf{a} = (a_0, a_1, \cdots)$ is a periodic sequence of period T with entries $a_i \in G$ then the* imbalance *of $\mathbf{a}$ with respect to a nontrivial character χ is*

$$Z_\chi(\mathbf{a}) = \sum_{g \in G} \mu(g)\chi(g) = \sum_{i=0}^{T-1} \chi(a_i),$$

where $\mu(g)$ is the number of occurrences of $g \in G$ in a single period of $\mathbf{a}$. The imbalance of an eventually periodic sequence is the imbalance of one complete period in the periodic part. An eventually periodic sequence $\mathbf{a}$ is balanced with respect to χ *if $|Z_\chi(\mathbf{a})| \leq 1$.*

If G is the cyclic group $\mathbb{Z}/(N)$, then its group of characters is also cyclic and is generated by the character $\chi(i) = \zeta^i$ where $\zeta \in \mathbb{C}$ is a primitive N-th root of unity. In this case we write $Z(\mathbf{a})$ for $Z_\chi(\mathbf{a})$, simply called the *imbalance of* $\mathbf{a}$. If $a \in \mathbb{Q}$ is a rational number with N-adic expansion $a = a_0 + a_1 N + \cdots$ then define the imbalance of a to be the imbalance of its coefficient sequence (which is eventually periodic).

Proposition 11.2.2. *Let G be a finite Abelian group with $|G| = N$ elements. Let $\mathbf{a}$ be a periodic sequence of elements $a_i \in G$ and period T. Suppose that $\mathbf{a}$ is balanced with respect to every nontrivial character $\chi : G \to \mathbb{C}$. Then $\mathbf{a}$ is balanced. Conversely, suppose that $\mathbf{a}$ is balanced and that*

$$T \pmod N \in \{-1, 0, 1\}.$$

That is, the period of $\mathbf{a}$ differs from a multiple of N at most by one. Then $\mathbf{a}$ is balanced with respect to every nontrivial character χ.

Proof. Let $\mu(g)$ be the number of times that $g \in G$ occurs in a single period of the sequence $\mathbf{a}$. Consider its Fourier transform, $\widehat{\mu}$, which is a function on the set of characters. For any character χ we have:

$$Z_\chi(\mathbf{a}) = \sum_{i=0}^{T-1} \chi(a_i) = \sum_{g \in G} \mu(g)\chi(g) = \widehat{\mu}(\chi)$$

The Fourier inversion formula says that for any $g \in G$,

$$\mu(g) = \frac{1}{N}\left(T + \sum_{\chi \neq \mathbf{1}} \widehat{\mu}(\chi)\overline{\chi}(g)\right).$$

Let us write $T = Nh + r$ with $0 \leq r \leq N-1$ and write $\epsilon = \epsilon(g)$ for the above sum, so that

$$\mu(g) = h + \frac{1}{N}(r + \epsilon).$$

The hypothesis of the Proposition implies that $|\epsilon| \leq N - 1$ since there are $N - 1$ non-trivial characters of G. Therefore $1 - N \leq r + \epsilon \leq 2N - 2$. But $\mu(g)$ is an integer so $r + \epsilon$ is a multiple of N, hence the only possibilities for $r + \epsilon$ are 0 or N. These give $\mu(g) = h$ or $\mu(g) = h + 1$. Thus, the sequence $\mathbf{a}$ is balanced.

The converse statement (which requires that $T \pmod N \in \{-1, 0, 1\}$) follows from Proposition 2.3.2. □

Proposition 11.2.2 implies, for example, that m-sequences and de Bruijn sequences (see Section 11.2.c below) are balanced with respect to every nontrivial character.

11.2.b Run property.

What is the expected number $E[Q(k)]$ of runs of length k in sequences $\mathbf{a}$ of period T? Let $Q(k,i)$ be the number of occurrences, among all possible sequences, of a run of length k starting at position i. Let us suppose that $k \leq T-2$. There are $|A|$ possible choices for the symbol that repeats in the run. There are $|A|-1$ possible choices for the symbol that precedes the run, and $|A|-1$ possible choices for the symbol that immediately follows the run, and $|A|$ choices for each of the remaining $T-k-2$ symbols. So

$$Q(k,i) = |A|(|A|-1)^2|A|^{T-k-2}.$$

Adding these up over all possible start positions i and dividing by the total number of sequences, $|A|^T$ gives

$$E[Q(k)] = \frac{T(|A|-1)^2}{|A|^{k+1}}.$$

Let us say that a periodic sequence $\mathbf{a}$ has the *run property* if the number $N(k)$ of runs of length k in the sequence satisfies

$$\left\lfloor \frac{T(|A|-1)^2}{|A|^{k+1}} \right\rfloor \leq N(k) \leq \left\lceil \frac{T(|A|-1)^2}{|A|^{k+1}} \right\rceil.$$

11.2.c de Bruijn sequences

The study of these sequences goes back to [20]. See also [73] Section 9.

Definition 11.2.3. *The sequence* $\mathbf{a}$ *is a* de Bruijn sequence *of span* k *if every block of length* k *occurs exactly once in (each period of)* $\mathbf{a}$.

Proposition 11.2.4. *Suppose* $\mathbf{a}$ *is a de Bruijn sequence of span* k. *Then*

1. *The period of* $\mathbf{a}$ *is* $T = |A|^k$.
2. *For any* $t \leq k$ *and for any block* b *of length* t, *the number of occurrences of* b *within a single period of* $\mathbf{a}$ *is* $|A|^{k-t}$, *hence* $\mathbf{a}$ *is equidistributed to order* k.
3. *For any* $m \leq k-2$ *the number of runs of length* m *within a single period of* $\mathbf{a}$ *is exactly*

$$|A|^{k-m-1}(|A|-1)^2. \tag{11.1}$$

The number of runs of length $k-1$ *is* $|A|(|A|-2)$. *The number of runs of length* k *is* $|A|$. *There are no runs of length greater than* k.

Proof. Part (1) is left to the reader, and part (2) just counts the number of ways of completing b to a block of length k. For part (3), first consider the case $1 \le m \le k-2$. A run of length m is a string of the form $xyy\cdots yz$ where x and z are distinct from y. By part (2), for each such choice of x, y, z there are $|A|^{k-m-2}$ occurrences of this block, and there are $|A|(|A|-1)^2$ such blocks, which gives equation (11.1). If $m = k-1$ then each block $xyy\cdots y$ with $x \neq y$ occurs once. Having chosen $y \in A$ there is a single value $x = x_0$ for x such that another y occurs after this block. Thus, we have two forbidden values $x \neq y, x_0$ for x, giving a count of $|A|(|A|-2)$ runs. Next, suppose $m = k$. For each y there is a single block of k consecutive $y's$ within a single period, giving $|A|$ such runs. Finally, if there were a run of length greater than k consisting of a single element y then there would be two or more occurrences of the block $y, y, \cdots, y$ of length k, which is a contradiction. □

11.2.d Punctured de Bruijn sequences

If $\mathbf{a}$ is a deBruijn sequence of span k, then for each $b_0 \in A$ the block $b_0, b_0, \cdots, b_0$ of length k occurrs exactly once. The periodic sequence $\mathbf{a}'$ is said to be a *punctured* or *modified* de Bruijn sequence, or a *pseudonoise sequence*, if it is obtained from a de Bruijn sequence $\mathbf{a}$ by deleting a single b_0 from the single occurrence of this block $b_0 b_0 \cdots b_0$ (in each period). It follows that such a punctured de Bruijn sequence $\mathbf{a}'$ has period $|A|^k - 1$. From Proposition 11.2.4 we conclude the following. For any $t \le k$ the number of occurrences of a fixed string b of length t is $|A|^{k-t}$ except for the single string $b_0 b_0 \cdots b_0$ of length k, which occurs $|A|^{k-1} - 1$ times. For $t = 1$, $A = \{0, 1, \cdots, N-1\}$, and $b_0 = 0$, this is the N-ary generalization of Golomb's first randomness postulate. For any $m \le k-1$ the number of runs of length m in $\mathbf{a}'$ is $|A|^{k-m-1}(|A|-1)^2$, and the number of runs of length k is $|A| - 1$. This is the the generalization of Golomb's second randomness postulate.

11.2.e Shift register generation of de Bruijn sequences

Suppose $\mathbf{a}$ is a de Bruijn sequence or a punctured de Bruijn sequence of span k with elements in some finite alphabet A. Consider a generalized (nonlinear) feedback shift register as in Figure 11.1.

The cell contents are elements of A, and the feedback function $f : A^k \to A$ is an arbitrary function. The output from the rightmost cell will be an eventually periodic sequence. It is easy to see that there exists a unique such $f : A^k \to A$ so that the output coincides with the de Bruijn sequence $\mathbf{a}$ as follows. Let $b = (b_0, b_1, \cdots, b_{k-1})$ be a block of size k. It occurs once in each period of the sequence $\mathbf{a}$ and it will be followed by some element, let us call it b_k. Imagine loading the shift register with the initial loading given by the block b. If the shift register is to output the sequence $\mathbf{a}$, then after shifting one step, the new contents of the leftmost cell must be b_k. So we are forced to define $f(b_0, b_1, \cdots, b_{k-1}) = b_k$. Such a definition will not lead to any contradictions because the block b occurs only once in (each period of) the sequence $\mathbf{a}$. Moreover, in the case

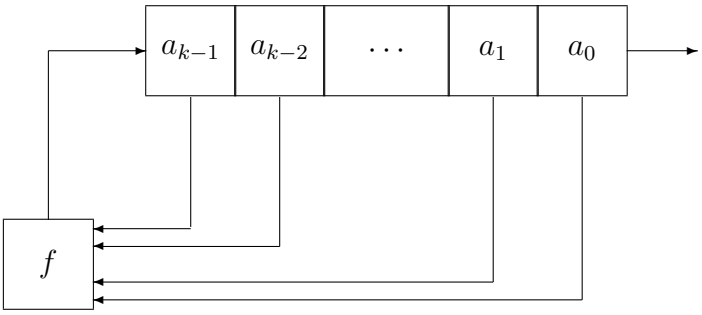


Figure 11.1: Generalized Feedback Shift Register of Length k.

of a de Bruijn sequence, this prescription defines f on all elements of A^k (that is, on all blocks of size k). In the case of a punctured de Bruijn sequence, there is a single block $(b_0, b_0, \cdots, b_0)$ of size k that does not occur in $\mathbf{a}$. In this case we define $f(b_0, b_0, \cdots, b_0) = b_0$, so that the state $(b_0, b_0, \cdots, b_0)$ is "stable". Starting the shift register at any other state will produce (a shift of) the sequence $\mathbf{a}$.

More generally, if $\mathbf{a}$ is any periodic sequence, then there is a minimal block length k so that every block of length k occurs at most once in $\mathbf{a}$. We call k the *nonlinear span* of $\mathbf{a}$. By a similar construction to the one in the preceding paragraph, we can construct a nonlinear feedback register of length k that outputs $\mathbf{a}$ (just fill in any undetermined values of f arbitrarily). This is a smallest nonlinear feedback shift register that outputs $\mathbf{a}$.

In some sense, all binary de Bruijn sequences are "known": there are iterative procedures for starting with a de Bruijn sequence $\mathbf{a}$ of span k and producing de Bruijn sequences of span $k+1$ that contain $\mathbf{a}$ as a subsequence. (See [1, 7, 120, 164].) However one does not necessarily derive from this process a simple prescription for finding the feedback function f. Only in the case of a binary *m-sequence* does one have a good understanding of the construction of the feedback function f (see Section 13). One of the long-standing unsolved problems in the theory of de Bruijn sequences is that of finding a simple prescription for those feedback functions f which produce de Bruijn and punctured de Bruijn sequences. In Section 15.1.c we describe a procedure for generating punctured de Bruijn sequences over a non-prime field, using an algebraic feedback shift register.

11.3 Correlations

11.3.a Classical correlations

Let G be a finite Abelian group. Let $\mathbf{a}$ and $\mathbf{b}$ be periodic sequences, with the same period T, of elements of G. For any integer τ let $\mathbf{a}^\tau$ be the τ-shift of $\mathbf{a}$, that is, $\mathbf{a}^\tau = (a_\tau, a_{\tau+1}, \cdots)$. Let $\chi : G \to \mathbb{C}^\times$ be a character of G. (See Section 2.3.a.) It takes values in the unit circle and $\overline{\chi}(g) = \chi(g)^{-1} = \chi(g^{-1})$.

Definition 11.3.1. *The (periodic)* cross-correlation function *of the sequences* $\mathbf{a}, \mathbf{b}$ *with respect to the character* χ *is the function*

$$\mathcal{C}_{\mathbf{a},\mathbf{b}}(\tau) = \sum_{i=0}^{T-1} \chi(a_i)\overline{\chi}(b_{i+\tau}). \tag{11.2}$$

The (periodic) autocorrelation *function* $\mathcal{A}_\mathbf{a}(\tau) = \mathcal{C}_{\mathbf{a},\mathbf{a}}(\tau)$ *of the sequence* $\mathbf{a}$ *is its cross-correlation with itself.*

A sequence $\mathbf{a}$ has *ideal autocorrelation* function if for every nontrivial character χ and for every nonzero shift τ we have: $|\mathcal{A}_\mathbf{a}(\tau)| \leq 1$. If $G = \mathbb{Z}/(2) = \{0, 1\}$, and $T = 2^k - 1$ for some k, then the stronger condition that $\mathcal{A}_\mathbf{a}(\tau) = -1$ for all τ that are not multiples of T is Golomb's third randomness postulate.

Remarks. In the case of binary sequences ($G = \{0, 1\}$) there is a single nontrivial character χ, and the cross-correlation $\mathcal{C}_{\mathbf{a},\mathbf{b}}(\tau)$ is the number of agreements minus the number of disagreements in a single period of $\mathbf{a}$ and $\mathbf{b}^\tau$. In most applications the group G is cyclic, so the group $\widehat{G}$ of characters of G consists of all the powers of a single ("primitive") generator χ, cf. Section 2.3.a. In these cases, it is traditional to consider only the cross-correlation with respect to such a primitive character χ. However, as we shall see, correlations tend to be remarkably insensitive to change of character.

11.3.b Expected correlation values

Let G be a finite Abelian group and let χ be a nontrivial character of G. Fix a positive integer T and consider the collection $\mathcal{S}$ of all periodic sequences of elements of G with period T. Let τ be an integer (to be interpreted as a "shift") with $0 \leq \tau \leq T - 1$. In this section we consider the expected value of the autocorrelation function, $\mathcal{A}_\mathbf{a}(\tau)$, averaged over all elements $\mathbf{a} \in \mathcal{S}$ and the expected value of the cross-correlation, $\mathcal{C}_{\mathbf{a},\mathbf{b}}(\tau)$, averaged over all pairs of elements $\mathbf{a}, \mathbf{b} \in \mathcal{S}$.

Proposition 11.3.2. *Let χ be a nontrivial character of G. The expected value of the autocorrelation $\mathcal{A}_{\mathbf{a}}(\tau)$ with respect to χ is*

$$E[\mathcal{A}_{\mathbf{a}}(\tau)] = \begin{cases} T & \text{if } \tau = 0 \\ 0 & \text{otherwise} \end{cases}$$

The expected value for the square of $\mathcal{A}_{\mathbf{a}}(\tau)$ is

$$E[\mathcal{A}_{\mathbf{a}}(\tau)^2] = \begin{cases} T^2 & \text{if } \tau = 0 \\ 2T & \text{if } T \text{ is even and } \tau = T/2 \text{ and } \chi^2 = 1 \\ T & \text{otherwise.} \end{cases}$$

The variance of $\mathcal{A}_{\mathbf{a}}(\tau)$ is

$$V[A_{\mathbf{a}}(\tau)] = \begin{cases} 0 & \text{if } \tau = 0 \\ 2T & \text{if } T \text{ is even and } \tau = T/2 \text{ and } \chi^2 = 1 \\ T & \text{otherwise.} \end{cases}$$

The expected value, the expected value of the square, and the variance of the cross-correlation with respect to χ, $\mathcal{C}_{\mathbf{a},\mathbf{b}}(\tau)$, of a pair of sequences $\mathbf{a}$ and $\mathbf{b}$ are

$$E[C_{\mathbf{a},\mathbf{b}}(\tau)] = 0, \quad E[C_{\mathbf{a},\mathbf{b}}(\tau)^2] = T, \quad \text{and} \quad V[C_{\mathbf{a},\mathbf{b}}(\tau)] = T.$$

Proof. There are $|G|^T$ sequences in $\mathcal{S}$. For any such sequence $\mathbf{a}$ the autocorrelation with shift $\tau = 0$ is

$$\mathcal{A}_{\mathbf{a}}(0) = \sum_{i=0}^{T-1} 1 = T,$$

so the expected value is T as well. If $T = 1$, then this is the only case, so we may now assume $T \geq 2$.

If $\tau \neq 0$ then the expected value of $\mathcal{A}_{\mathbf{a}}(\tau)$ is

$$E[\mathcal{A}_{\mathbf{a}}(\tau)] = \frac{1}{|G|^T} \sum_{\mathbf{a} \in \mathcal{S}} \sum_{i=0}^{T-1} \chi(a_i)\overline{\chi}(a_{i+\tau}) = \frac{1}{|G|^T} \sum_{i=0}^{T-1} \sum_{\mathbf{a} \in \mathcal{S}} \chi(a_i)\chi(-a_{i+\tau}).$$

For any $g, h \in G$ and $0 \leq i < T$, there are $|G|^{T-2}$ sequences $\mathbf{a} \in \mathcal{S}$ such that $a_i = g$ and $a_{i+\tau} = h$. This gives

$$E[\mathcal{A}_{\mathbf{a}}(t)] = \frac{1}{|G|^T} \sum_{i=0}^{T-1} |G|^{T-2} \sum_{g \in G} \sum_{h \in G} \chi(g)\chi(-h) = \frac{1}{|G|^2} \sum_{i=0}^{T-1} \left(\sum_{g \in G} \chi(g) \right)^2 = 0$$

by Proposition 2.3.2.

Similarly, if $\tau = 0$ and if $\mathbf{a} \in \mathcal{S}$ is any sequence then $\mathcal{A}_{\mathbf{a}}(0)^2 = T^2$. If $0 < \tau \leq T-1$ then

$$\begin{aligned} E[\mathcal{A}_{\mathbf{a}}(\tau)^2] &= \frac{1}{|G|^T} \sum_{\mathbf{a}\in\mathcal{S}} \sum_{i=0}^{T-1} \chi(a_i)\overline{\chi}(a_{i+\tau}) \sum_{j=0}^{T-1} \overline{\chi}(a_j)\chi(a_{j+\tau}) \\ &= \frac{1}{|G|^T} \sum_{i=0}^{T-1} \sum_{j=0}^{T-1} H(i,j,\tau) \end{aligned}$$

where

$$H(i,j,\tau) = \sum_{\mathbf{a}\in\mathcal{S}} \chi(a_i)\overline{\chi}(a_{i+\tau})\overline{\chi}(a_j)\chi(a_{j+\tau}).$$

There are five cases to consider. For $T \geq 3$, they are listed in Table 11.1 together with the number N of pairs i, j with $0 < i, j \leq T-1$ for which each case occurs. The expected values of the autocorrelation follow from the data in Table 11.1.

Type	Description	N $(\tau \neq T/2)$	N $(\tau = T/2)$	H $\chi^2 \neq 1$	H $\chi^2 = 1$
(a)	$i, j, i+\tau, j+\tau$ all distinct	$T(T-3)$	$T(T-2)$	0	0
(b)	$i = j$ and $i+\tau = j+\tau$	T	T	$\|G\|^T$	$\|G\|^T$
(c)	$i \equiv j+\tau$ and $i+\tau \not\equiv j \pmod{T}$	T	0	0	0
(d)	$i \equiv j+\tau$ and $i+\tau \equiv j \pmod{T}$	0	T	0	$\|G\|^T$
(e)	$j \equiv i+\tau$ and $j+\tau \not\equiv i \pmod{T}$	T	0	0	0

Table 11.1: Numbers of Occurrences of Values of H

In case (a), the index i can be chosen arbitrarily (T choices), then j is chosen with three forbidden values, $i, i+\tau, i-\tau \pmod{T}$. (If $\tau = T/2$ then this amounts to only two forbidden values.) If $T \geq 4$ then the four values $g = a_i$, $h = a_{i+\tau}$, $g' = a_j$, $h' = a_{j+\tau}$ may be chosen independently so

$$H(i,j,\tau) = |G|^{T-4} \sum_{g\in G} \sum_{h\in G} \sum_{g'\in G} \sum_{h'\in G} \chi(g)\overline{\chi}(h)\overline{\chi}(g')\chi(h') = 0$$

again by Proposition 2.3.2. In case (b) the index i is chosen arbitrarily so there are T choices. The values $g = a_i$ and $h = a_{i+\tau}$ may be chosen independently so in this case

$$H(i,j,\tau) = |G|^{T-2} \sum_{g\in G} \sum_{h\in G} \chi(g)\overline{\chi}(h)\overline{\chi}(g)\chi(h) = |G|^T.$$

Cases (c) and (e) are similar, as is the cross-correlation calculation. In case (d) we find

$$H(i,j,\tau) = \sum_{\mathbf{a}\in\mathcal{S}} \chi^2(a_i)\chi^2(a_j)$$

which is 0 if $\chi^2 \neq 1$ and is $|G|^T$ if $\chi^2 = 1$.

If $T = 2$, then $\tau = 1 = T/2$ and the table is still valid. If $T = 3$, then $\tau \neq T/2$ and case (a) reduces to 0, so the table is also still valid.

The variance of a random variable X is $V[x] = E[X^2] - E[X]^2$, so the results on the variances follow immediately. □

11.3.c Arithmetic correlations

In this section we define a with carry analog of the usual notion of cross-correlations.

For any N-ary sequence $\mathbf{b}$, let $\mathbf{b}^\tau$ be the sequence formed by shifting $\mathbf{b}$ by τ positions, $b_i^\tau = b_{i+\tau}$. In terms of Section 11.2.a, the ordinary cross-correlation with shift τ of two N-ary sequences $\mathbf{a}$ and $\mathbf{b}$ of period T is the imbalance of the coefficient sequence of the difference between the power series associated with $\mathbf{a}$ and the power series associated with $\mathbf{b}^\tau$. In the binary case this is the number of zeros minus the number of ones in one period of the bitwise exclusive-or of $\mathbf{a}$ and the τ shift of $\mathbf{b}$ [54]. The arithmetic cross-correlation is the with-carry analogue.

Definition 11.3.3. *Let* $\mathbf{a}$ *and* $\mathbf{b}$ *be two eventually periodic* N*-ary sequences with period* T *and let* $0 \leq \tau < T$*. Let* a *and* b^τ *be the* N*-adic numbers whose coefficients are given by* $\mathbf{a}$ *and* $\mathbf{b}^\tau$*, so* $\mathbf{a} = \mathbf{seq}_N(a)$ *and* $\mathbf{b}^\tau = \mathbf{seq}_N(b^\tau)$*. Then* $\mathbf{seq}_N(a - b^\tau)$ *is eventually periodic and its period divides* T*. Fix a non-trivial character* χ *of* $\mathbb{Z}/(N)$*,* $\chi(x) = \zeta^x$ *with* ζ *a primitive complex* N*th root of unity. The* arithmetic cross-correlation *(with shift* τ*) of* $\mathbf{a}$ *and* $\mathbf{b}$ *is*

$$\mathcal{C}^A_{\mathbf{a},\mathbf{b}}(\tau) = Z(a - b^{(\tau)}), \tag{11.3}$$

where Z *is the imbalance with respect to* χ *(see Definition 11.2.1). When* $\mathbf{a} = \mathbf{b}$*, the arithmetic cross-correlation is called the* arithmetic autocorrelation *of* $\mathbf{a}$ *and is denoted* $\mathcal{A}^A_{\mathbf{a}}(\tau)$*.*

The sequences $\mathbf{a}, \mathbf{b}$ (of period T) are said to have *ideal arithmetic correlations* if,

$$C^A_{\mathbf{a},\mathbf{b}}(\tau) = \begin{cases} T & \text{if } \mathbf{a} = \mathbf{b}^\tau \\ 0 & \text{otherwise} \end{cases}$$

for each τ with $0 \leq \tau < T$. A family of sequences is said to have ideal arithmetic correlations if every pair of sequences in the family has ideal arithmetic correlations.

11.3.d Expected arithmetic correlations

In this section we investigate the expected values of the arithmetic autocorrelations and cross-correlations for a fixed shift. We also compute the second moment and variance of the arithmetic cross-correlations. Computing the second moment and variance of the arithmetic autocorrelation is an open problem.

First we do some initial analysis. Fix a period T. By Theorem 5.4.4, the N-ary sequences of period T are the coefficient sequences $\mathbf{a}$ of rational numbers of the form $a = -f/(N^T - 1)$ with $0 \leq f \leq N^T - 1$.

Lemma 11.3.4. *If a and b are distinct N-adic numbers whose coefficient sequences $\mathbf{seq}_N(a), \mathbf{seq}_N(b)$ are periodic with period T, and $a - b \in \mathbb{Z}$, then $\{a, b\} = \{0, -1\}$.*

Proof. First note that the N-adic expansion of an integer is strictly periodic if and only the integer is 0 or -1. Let $a = -f/(N^T-1)$ and $b = -g/(N^T-1)$ with $0 \leq f, g \leq N^T-1$. Assume that $f > g$. Then $a-b = -(f-g)/(N^T-1)$ is strictly periodic and nonzero, so $-1 = a-b = -(f-g)/(N^t-1)$. Thus $f = g + (N^T - 1)$. The only possibility is that $f = N^T - 1$ and $g = 0$. That is, $a = -1$ and $b = 0$. The case when $g > f$ is similar. □

Next fix a shift τ. Then the τ shift of $\mathbf{a}$ corresponds to a rational number

$$a_\tau = c_{f,\tau} + \frac{-N^{T-\tau} f}{N^T - 1},$$

where $0 \leq c_{f,\tau} < N^{T-\tau}$ is an integer.

Now let $\mathbf{b}$ be another periodic N-ary sequence with the same period T and corresponding to the rational number

$$b = \frac{-g}{N^T - 1}.$$

Then the arithmetic cross-correlation between $\mathbf{a}$ and $\mathbf{b}$ with shift τ is

$$\begin{aligned} \mathcal{C}^A_{\mathbf{a},\mathbf{b}}(\tau) &= Z\left(\frac{-f}{N^T - 1} - \left(c_{g,\tau} + \frac{-N^{T-\tau} g}{N^T - 1}\right)\right) \\ &= Z\left(\frac{N^{T-\tau} g - f}{N^T - 1} - c_{g,\tau}\right). \end{aligned} \tag{11.4}$$

Theorem 11.3.5. *For any τ, the expected arithmetic autocorrelation, averaged over all sequences $\mathbf{a}$ of period T, is*

$$E[\mathcal{A}^A_{\mathbf{a}}(\tau)] = \frac{T}{N^{T-\gcd(\tau,T)}}.$$

The expected cross-correlation, averaged over all pairs of sequences $\mathbf{a}$ *and* $\mathbf{b}$ *is*

$$E[\mathcal{C}^A_{\mathbf{a},\mathbf{b}}(\tau)] = \frac{T}{N^T}.$$

Proof. If the τ shift of $\mathbf{b}$ equals $\mathbf{a}$, then $\mathcal{C}^A_{\mathbf{a},\mathbf{b}}(\tau) = T$. Otherwise $\mathbf{a}$ and $\mathbf{b}^\tau$ are distinct periodic sequences. In particular, by Lemma 11.3.4 $a - b_\tau$ is an integer only if $\{a, b_\tau\} = \{0, -1\}$.

First we consider the autocorrelation. Let

$$S = \sum_{f=0}^{N^T-1} Z\left(\frac{(N^{T-\tau}-1)f}{N^T-1} - c_{f,\tau}\right).$$

It follows from equation (11.4) that the expected arithmetic autocorrelation is $E[\mathcal{A}^A_{\mathbf{a}}(\tau)] = S/N^T$.

By the first paragraph of this proof $a - a_\tau$ is an integer only if $a_\tau = a$. When it is not an integer, the periodic part of

$$\frac{(N^{T-\tau}-1)f}{N^T-1} - c_{f,\tau}$$

is the same as the periodic part of

$$\frac{(N^{T-\tau}-1)f \pmod{N^T-1}}{N^T-1},$$

where we take the reduction modulo N^T-1 in the set of residues $\{-(N^T-2), -(N^T-3), \cdots, -1, 0\}$. In particular, this latter rational number has a strictly periodic N-adic expansion with period T, so we can compute its contribution to S by considering the first T coefficients.

Let $d = \gcd(T, T-\tau) = \gcd(T, \tau)$. Then $\gcd(N^T-1, N^{T-\tau}-1) = N^d - 1$. Then the set of elements of the form $(N^{T-\tau}-1)f \pmod{N^T-1}$ is the same as the set of elements of the form $(N^d-1)f \pmod{N^T-1}$. Thus

$$S = \sum_{f=0}^{N^T-1} Z\left(\frac{(N^d-1)f \pmod{N^T-1}}{N^T-1}\right).$$

Now consider the contribution to S from the ith term in the expansion in each element in the sum, say corresponding to an integer f. If we multiply f by N^{T-i} modulo N^T-1, this corresponds to cyclically permuting the corresponding sequence to the right by $T-i$ places. This is equivalent to permuting to the left by i positions, so the elements in the ith place become the elements in the 0th place. Moreover, multiplying by N^{T-i} is a permutation modulo N^T-1, so the distribution of values contributing to S from the ith terms is identical to the distribution of values from the 0th term.

To count the contribution from the 0th position, let

$$D = \frac{N^T - 1}{N^d - 1}$$

and $f = u + vD$ with $0 < u < D$ and $0 \leq v < N^d - 1$. Then

$$(N^d - 1)f \pmod{N^T - 1} = (N^d - 1)u \pmod{N^T - 1} = (N^d - 1)u - (N^T - 1).$$

Thus

$$\frac{(N^d - 1)f \pmod{N^T - 1}}{N^T - 1} = \frac{(N^d - 1)u}{N^T - 1} - 1. \tag{11.5}$$

In particular, the contribution to S from the 0th position depends only on u. Thus we can count the contributions over all g with $0 < u < D$, and then multiply by $N^d - 1$. The contribution from the 0th position for a particular u is given by reducing the right hand side of equation (11.5) modulo N. We have

$$\begin{aligned} \frac{(N^d - 1)u}{N^T - 1} - 1 &= (1 + N^T + N^{2T} + \cdots)(N^d - 1)u - 1 \\ &\equiv -u - 1 \pmod{N}. \end{aligned}$$

Since

$$-(1 + N^d + N^{2d} + \cdots + N^{T-d}) \leq -u - 1 \leq -2,$$

as u varies its reduction modulo N takes each value in $\{0, 1, \cdots, N-1\}$ exactly

$$N^{d-1} + N^{2d-1} + \cdots + N^{T-d-1}$$

times. It follows that the contribution to S from the sequences that are not equal to their τ shifts is a multiple of

$$1 + \zeta + \cdots + \zeta^{N-1} = 0.$$

Thus we need to count the number of sequences that are equal to their τ shifts. These are the sequences whose minimal periods are divisors of τ. Of course the minimal periods of such sequences are also divisors of T, so it is equivalent to count the sequences whose minimal period divides d. The number of such sequences is exactly N^d. Thus the expected autocorrelation is

$$E[\mathcal{A}^A_{\mathbf{a}}(\tau)] = \frac{N^d T}{N^T}.$$

Now consider the expected cross-correlation. The set of τ shifts of all T-periodic sequences is just the set of all T-periodic sequences, so we can take $\tau = 0$. Thus $c_{g,\tau} = 0$. Let

$$R = \sum_{f,g=0}^{N^T-1} Z\left(\frac{g - f}{N^T - 1}\right) = \sum_{f,g=0}^{N^T-1} Z\left(\frac{(g - f) \pmod{N^T - 1}}{N^T - 1}\right).$$

Here when we reduce modulo $N^T - 1$ we must take $(g - f) \pmod{N^T - 1} = g - f$ if $g \leq f$ and $(g-f) \pmod{N^T - 1} = g - f - (N^T - 1)$ if $f < g$. By similar arguments to those in the derivation of the expected autocorrelations, we can reduce to counting the distribution of values contributed by the 0th positions. Thus

$$R = TN^T + \sum_{0 \leq g < f \leq N^T - 1} Z\left(\frac{g-f}{N^T-1}\right) + \sum_{0 \leq f < g \leq N^T - 1} Z\left(\frac{g-f-N^T+1}{N^T-1}\right),$$

where the first term on the left hand side accounts for the cases when $f = g$. The rational numbers in the remaining two terms have periodic expansions, so the reductions of these numbers modulo N are coefficients of N^0 in the expansions of the numerators. That is, if $f = f_0 + Nf'$ and $g = g_0 + Ng'$ with $0 \leq f_0, g_0 < N$, then

$$R = TN^T + \sum_{0 \leq g < f \leq N^T - 1} \zeta^{g_0 - f_0} + \sum_{0 \leq f < g \leq N^T - 1} \zeta^{g_0 - f_0 - 1}.$$

We have $g < f$ if and only if either $g' < f'$ or $g' = f'$ and $g_0 < f_0$. If we fix an f and consider all $g < f$ with $g' < f'$, then the g_0 is free to vary over $\{0, 1, \cdots, N-1\}$. Thus the contribution to R from such terms is zero. Similarly, if $f' < g'$ then the contribution is zero. Thus we only need to consider the cases when $f' = g'$. Since there are N^{T-1} choices of f', we have

$$\begin{aligned}
R &= TN^T + N^{T-1} \sum_{f_0=0}^{N-1} \sum_{g_0=0}^{f_0-1} \zeta^{g_0-f_0} + N^{T-1} \sum_{f_0=0}^{N-1} \sum_{g_0=f_0+1}^{N-1} \zeta^{g_0-f_0-1} \\
&= TN^T + N^{T-1} \sum_{f_0=0}^{N-1} \sum_{g_0=0}^{f_0-1} \zeta^{g_0-f_0} + N^{T-1} \sum_{f_0=0}^{N-1} \sum_{g_0=f_0}^{N-2} \zeta^{g_0-f_0} \\
&= TN^T + N^{T-1} \sum_{f_0=0}^{N-1} \sum_{g_0=0}^{N-2} \zeta^{g_0-f_0} \\
&= TN^T + N^{T-1} \frac{\zeta^{-N} - 1}{\zeta - 1} \cdot \frac{\zeta^{N-1} - 1}{\zeta - 1} \\
&= TN^T.
\end{aligned}$$

Thus

$$E[\mathcal{C}^A_{\mathbf{a},\mathbf{b}}(\tau)] = \frac{R}{N^{2T}} = \frac{T}{N^T}.$$ □

Theorem 11.3.6. *For any shift τ, the second moment of the arithmetic cross-correlation, averaged over all pairs of sequences* **a** *and* **b** *is*

$$E[\mathcal{C}^A_{\mathbf{a},\mathbf{b}}(\tau)^2] = T\frac{N^T + 1 - T}{N^T}.$$

The variance is

$$V[\mathcal{C}^A_{\mathbf{a},\mathbf{b}}(\tau)] = T\frac{(N^T+1)(N^T-T)}{N^{2T}}.$$

Proof. As in the computation of the expectation, we can reduce to the case when $\tau = 0$. We let

$$P = \sum_{f,g=0}^{N^T-1} \left| Z\left(\frac{(g-f) \pmod{N^T-1}}{N^T-1}\right)\right|^2,$$

so that the second moment is P/N^{2T}. We proceed by determining the number of pairs f, g with

$$g - f \equiv -h \pmod{N^T-1} \quad \text{and} \quad 0 \le f, g \le N^T - 1 \tag{11.6}$$

for each h with $0 \le h \le N^T - 1$. If $h = N^T - 1$, then equation (11.6) only holds for $g = 0$ and $f = N^T - 1$. Let $h < N^T - 1$. For every $f < N^T - 1$ there is exactly one $g < N^T - 1$ satisfying equation (11.6), namely $f - h \pmod{N^T - 1}$. For $g = N^T - 1$ there is one $f < N^T - 1$ satisfying equation (11.6), namely $f = h$. For $f = N^T - 1$ there is one g with $1 \le g \le N^T - 1$ satisfying equation (11.6), namely $g = N^T - 1 - h$. This accounts for all choices of f and g, and we see that for $0 \le h < N^T - 1$ there are $N^T + 1$ pairs f, g satisfying equation (11.6), and for $h = N^T - 1$ there is one such pair.

Let us first compute as if all h occurred equally often. We switch to thinking about N-ary T-tuples $\mathbf{h} = (h_0, h_1, \cdots, h_{T-1})$ representing single periods. If each h occurred $N^T + 1$ times, then each $\mathbf{h}$ would occur $N^T + 1$ times as single periods of $-h/(N^t - 1)$s. This would give a total contribution to P of

$$\begin{aligned}
(N^T+1)\sum_{\mathbf{h}}\left|\sum_{i=0}^{T-1}\zeta^{h_i}\right|^2 &= (N^T+1)\sum_{\mathbf{h}}\sum_{i=0}^{T-1}\sum_{j=0}^{T-1}\zeta^{h_i-h_j} \\
&= (N^T+1)\sum_{i=0}^{T-1}\sum_{\mathbf{h}}\zeta^{h_i-h_i} + (N^T+1)\sum_{i\neq j}\sum_{\mathbf{h}}\zeta^{h_i-h_j} \\
&= TN^T(N^T+1) + (N^T+1)\sum_{i\neq j}N^{T-2}\sum_{h_i,h_j}\zeta^{h_i-h_j} \\
&= TN^T(N^T+1).
\end{aligned}$$

But we have over counted since the T-tuple $\mathbf{h} = (N-1, N-1, \cdots, N-1)$, corresponding to the N-adic number -1 and the numerator $h = N^T - 1$, only occurs once. For this value of h we have $|Z(-1)|^2 = T^2$, so in fact

$$P = TN^T(N^T+1) - T^2N^T = TN^T(N^T+1-T).$$

It follows that

$$E[\mathcal{C}^A_{\mathbf{a},\mathbf{b}}(\tau)^2] = \frac{P}{N^{2T}} = T\frac{N^T + 1 - T}{N^T}$$

as claimed. The claimed value of the variance follows. □

11.3.e Hamming Correlations and Frequency Hopping

The fascinating history of research on frequency hopping includes the now-famous 1942 patent [129] of actress Hedy Lamarr (Hedy Marky) and composer George Antheil for remote control guidance of torpedos, which used a piano roll to implement a pseudo-random sequence of frequency hops, in order to avoid jamming by the enemy; see [179, 186]. Marky was familiar with issues of torpedo guidance as her ex-husband was a munitions dealer in Austria. After escaping from Austria and emigrating to America, she apparently came up with the frequency hopping idea while playing a piano duet in which one voice would lead and the second voice would follow with exactly the same sequence of notes. The technical expertise for implementing this remarkable idea was provided in part by Antheil's experience with player piano mechanisms, which he had used in his *Ballet Méchanique* (1924).

In a frequency hopping system, the transmitter and receiver jump, in synchronization, from one frequency to another, as directed by a pseudo-random sequence $\mathbf{a} = a_0, a_1, \cdots$ that they share. Thus the symbols of the sequence are drawn from an alphabet $F = \{f_1, f_2, \cdots, f_m\}$, sometimes called a *frequency library*, whose elements correspond in a one to one way with the frequencies that are allocated to the system. The system might implement *fast hops* in which a single information symbol might be transmitted over several hops, or *slow hops*, in which a single hop might involve the transmission of several information symbols.

Now suppose a second transmitter-receiver couple wishes to use the same collection of frequencies for their own communications according to a second pseudo-random sequence $\mathbf{b}$ with the same set of symbols $b_i \in F$, and (in most applications) with the same period, say, T.

A *collision* occurs when both transmitters send messages on the same frequency. Provided this happens relatively rarely, an error correcting code can be used to recover from collisions. In their foundational article [121], Lempel and Greenberger defined the *Hamming (cross)correlation* with shift τ of the sequences $\mathbf{a}, \mathbf{b}$ to be the number of collisions that occur in a single period, that is,

$$\mathcal{C}^{\text{Ham}}_{\mathbf{a},\mathbf{b}}(\tau) = \sum_{i=0}^{T-1} \epsilon(a_i, b_{i+\tau}) \quad \text{where} \quad \epsilon(a_i, b_j) = \begin{cases} 1 & \text{if } a_i = b_j \\ 0 & \text{if } a_i \neq b_j. \end{cases}$$

The *Hamming autocorrelation* function is $\mathcal{A}^{\text{Ham}}_{\mathbf{a}}(\tau) = \mathcal{C}^{\text{Ham}}_{\mathbf{a},\mathbf{a}}(\tau)$.

We need to assign a pseudo-random sequence to each of the transmitters in such a way that the number of collisions never exceeds the error correcting capability of the ECC. In order to synchronize the receiver with the transmitter, we also need to minimize the number of self-collisions, that is, collisions between the sequence $\mathbf{a}$ and its shifts.

Therefore we need to find a collection $\mathcal{S}$ of sequences such that for each pair $\mathbf{a}, \mathbf{b} \in \mathcal{S}$ of sequences, the numbers $\mathcal{A}^{\mathrm{Ham}}(\mathbf{a}), \mathcal{A}^{\mathrm{Ham}}(\mathbf{b}), \mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b})$ are all small, where

$$\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) = \max_{0 \le \tau \le T-1} C^{\mathrm{Ham}}_{\mathbf{a},\mathbf{b}}(\tau) \quad \text{and} \quad \mathcal{A}^{\mathrm{Ham}}(\mathbf{a}) = \max_{1 \le \tau \le T-1} \mathcal{A}^{\mathrm{Ham}}_{\mathbf{a}}(\tau).$$

Lempel and Greenberger showed that there are *a priori* limits on how small these correlations can be. We give their result (in a slightly modified and generalized form) in Theorem 11.3.7 below .

Throughout this section, when considering a sequence of period T of elements in a set F we denote by α, β the unique non-negative integers such that $T = \alpha|F| + \beta$ and $0 \le \beta < |F|$.

Theorem 11.3.7. [121] *Let* $\mathbf{a}, \mathbf{b}$ *be sequences of period* $T \ge 2$ *with symbols drawn from an alphabet* F *of size* $|F| \le T$. *Then*

$$\mathcal{A}^{\mathrm{Ham}}(\mathbf{a}) \ge \begin{cases} \alpha - 1 & \text{if } |F| = T \\ \alpha & \text{otherwise.} \end{cases} \tag{11.7}$$

If $\mathbf{a}$ *and* $\mathbf{b}$ *are balanced then*

$$\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) \ge \begin{cases} \alpha & \text{if } \beta = 0 \\ \alpha + 1 & \text{otherwise.} \end{cases} \tag{11.8}$$

Proof. The maximum value $\mathcal{A}^{\mathrm{Ham}}(\mathbf{a})$ of the autocorrelation is at least as big as its average value $\mathcal{A}^{\mathrm{Ham}}_{av}(\mathbf{a})$ (averaged over all nonzero shifts τ) so

$$\mathcal{A}^{\mathrm{Ham}}(\mathbf{a}) \ge \mathcal{A}^{\mathrm{Ham}}_{av} = \frac{1}{T-1} \sum_{\tau=1}^{T-1} \sum_{i=0}^{T-1} \epsilon(a_i, a_{i+\tau}).$$

For any $f \in F$ let $K(f)$ be the number of times the symbol f occurs in the sequence $\mathbf{a}$, so that $\sum_{f \in F} K(f) = T$. Exchanging the order of summation, and summing over τ gives:

$$\mathcal{A}^{\mathrm{Ham}}_{av}(\mathbf{a}) = \frac{1}{T-1} \sum_{i=0}^{T-1} (K(a_i) - 1) = \frac{1}{T-1} \left(\sum_{f \in F} K(f)^2 - T \right)$$

Then the minimum is attained when the sequence $\mathbf{a}$ is balanced, which is to say that $K(f) \in \{\alpha, \alpha + 1\}$ for every $f \in F$. In this case, $K(f) = \alpha + 1$ occurs for β values of f, and $K(f) = \alpha$ occurs for $|F| - \beta$ values of f. This gives $\mathcal{A}^{\mathrm{Ham}}(\mathbf{a}) \ge \mathcal{A}_0$ where

$$\mathcal{A}_0 = \frac{\alpha}{T-1}(T - |F| + \beta).$$

To further evaluate $\lceil \mathcal{A}_0 \rceil$, we claim that

$$\alpha - 1 \leq \mathcal{A}_0 \leq \alpha$$

and that the first inequality is strict unless $T = |F|$. Multiplying by $T-1$, we need to show that

$$(\alpha - 1)(T-1) \leq \alpha(T - |F| + \beta) \leq \alpha(T-1).$$

The second inequality holds because $\beta - |F| \leq -1$. For the first inequality, multiply out and cancel αT from both sides, leaving

$$-\alpha - T + 1 \leq -\alpha|F| + \alpha\beta$$

or,

$$1 \leq \alpha + \beta + \alpha\beta.$$

Assuming $n \neq 0$, this always holds. Equality holds if and only if $(\alpha, \beta) = (1, 0)$ or $(0, 1)$ which is equivalent to the statements that $T = |F|$ or $T = 1$ respectively.

Next, let us consider lower bounds for $\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b})$ under the hypothesis that $\mathbf{a}, \mathbf{b}$ are balanced. Let $K(f)$ (respectively, $L(f)$) denote the number of times that $f \in F$ occurs in the sequence $\mathbf{a}$ (resp., in the sequence $\mathbf{b}$), so that so that

$$\sum_{f \in F} K(f) = \sum_{f \in F} L(f) = T.$$

The same argument as in the preceding paragraph gives

$$\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) \geq \mathcal{C}^{\mathrm{Ham}}_{av}(\mathbf{a}, \mathbf{b}) = \frac{1}{T} \sum_{i=0}^{T-1} L(a_i) = \frac{1}{T} \left(\sum_{f \in F} K(f) L(f) \right) \tag{11.9}$$

The balance condition means that $K(f), L(f) \in \{\alpha, \alpha + 1\}$. The minimum value for the sum in equation (11.9) occurs when $K(f) = \alpha$ is paired with $L(f) = \alpha + 1$ (or vice versa) as often as possible because

$$\alpha(\alpha + 1) + (\alpha + 1)\alpha < \alpha^2 + (\alpha + 1)^2.$$

This gives

$$\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}_0, \mathbf{b}_0) \geq \mathcal{C}_0 = \begin{cases} \frac{\alpha}{T}(T + \beta) & \text{if} \quad \beta \leq |F|/2 \\ \\ \frac{\alpha+1}{T}(T - |F| + \beta) & \text{if} \quad \beta \geq |F|/2. \end{cases}$$

Since $\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b})$ is an integer and since $0 \leq \beta < |F|$, we conclude, for balanced sequences $\mathbf{a}, \mathbf{b}$, that

$$\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) \geq \lceil \mathcal{C}_0 \rceil = \begin{cases} \alpha & \text{if } \beta = 0 \\ \alpha + 1 & \text{otherwise.} \end{cases}$$

□

Remarks. If $\mathbf{a}, \mathbf{b}$ are not balanced then $\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b})$ may drop below $\alpha + 1$. Indeed, it will be zero if the sequences $\mathbf{a}$ and $\mathbf{b}$ use disjoint sets of symbols from the frequency library F.

In [122] there is an argument giving a lower bound for the quantity

$$M^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) = \max\left\{\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}), \mathcal{A}^{\mathrm{Ham}}(\mathbf{a}), \mathcal{A}^{\mathrm{Ham}}(\mathbf{b})\right\}$$

which is repeated in [179] and leads to the bound

$$M^{\mathrm{Ham}} \geq \frac{\alpha+1}{T}(T - |F| + \beta) = \begin{cases} \alpha & \text{if } \beta = 0 \\ \alpha + 1 & \text{otherwise.} \end{cases}$$

However there is an error in this argument and a counterexample is the following. Let $F = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and consider the following two sequences of period 9,

$$\mathbf{a} = (1\ 1\ 2\ 3\ 1\ 4\ 5\ 6\ 7) \qquad \mathbf{b} = (8\ 8\ 7\ 6\ 8\ 5\ 4\ 3\ 2)$$

Then $\mathcal{A}^{\mathrm{Ham}}(\mathbf{a}) = \mathcal{A}^{\mathrm{Ham}}(\mathbf{b}) = \mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) = M^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) = \alpha = 1$. (These sequences are quite unbalanced since $\mathbf{a}$ does not even use the symbol $8 \in F$, and $\mathbf{b}$ does not use $1 \in F$.)

It always holds that $M^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) \geq \mathcal{A}^{\mathrm{Ham}}(\mathbf{a}) \geq \alpha$, so at best one might hope to improve this bound by 1, for certain values of $\beta > 0$. We do not know precise conditions on β which guarantee that $M^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) \geq \alpha + 1$ but we conjecture that $\beta \geq |F|/2$ suffices.

11.3.f Hamming-Optimal families

Let $\mathcal{S}$ denote a family of sequences of period T over an alphabet F. In [121] Lempel and Greenberger defined an *optimal sequence* to be one that meets the bound $\mathcal{A}^{\mathrm{Ham}}(\mathbf{a}) \geq \alpha$ of Theorem 11.3.7. They defined $\mathcal{S}$ to be an *optimal family* if every sequence $\mathbf{a} \in \mathcal{S}$ is optimal and if every distinct pair of sequences $\mathbf{a}, \mathbf{b} \in \mathcal{S}$ meets the bound $\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) \geq \alpha + 1$ of Theorem 11.3.7.

One might expect that a large family $\mathcal{S}$ will necessarily contain sequences $\mathbf{a}, \mathbf{b}$ with large (Hamming) cross-correlation, as predicted by the Welch bound (see Chapter 14) for the (usual) cross-correlation. A coding-theoretic approach to this question was described in [162]. Let us assume the alphabet F is an Abelian group. If $\mathbf{a} = (a_0, a_1, \cdots) \in \mathcal{S}$, $v \in F$, and $\tau \in \mathbb{Z}$, then let

$$\mathbf{a} + v = (a_0 + v, a_1 + v, \cdots)$$

denote the *v-translated* sequence and let $\mathbf{a}^\tau = (a_\tau, a_{\tau+1}, a_{\tau+2} \cdots)$ denote the *τ-shifted* sequence. Let us assume that the family $\mathcal{S}$ is *translationally closed*, that is, if $\mathbf{a} \in \mathcal{S}$ then $\mathbf{a} + v \in \mathcal{S}$ for any $v \in F$. Let us also assume the sequences in the family $\mathcal{S}$ are shift distinct.

We may consider (a single period of) the sequence $\mathbf{a} \in \mathcal{S}$ to be a vector in F^T. Thus the collection of all sequences in $\mathcal{S}$, together with their left shifts, gives $|\mathcal{S}|T$ vectors in F^T which,

together with the zero vector, may be thought of as a (possibly nonlinear) code $\overline{\mathcal{S}} \subset F^T$. (The bar indicates that we have augmented the family $\mathcal{S}$ by including the cyclic shifts of sequences.) The key observation relating Hamming correlation to coding theory is the following statement, whose proof is immediate.

Lemma 11.3.8. *Let $d_{\min} = d_{\min}^{\mathrm{Ham}}(\overline{\mathcal{S}})$ denote the minimum Hamming distance between any two codewords (i.e. vectors) in $\overline{\mathcal{S}}$. Then for any distinct pair $\mathbf{a}, \mathbf{b} \in \overline{\mathcal{S}}$ the following inequality holds:*

$$\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) \geq T - d_{\min}. \tag{11.10}$$

At this point one could apply various known bounds on $d_{\min}$, the minimum Hamming distance between codewords of a code. In [162], Quang et al apply the *singleton bound*

$$d_{\min} \leq T - \log_{|F|}(|\overline{\mathcal{S}}|) + 1$$

which gives the (rather weak) bound,

$$\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) \geq \log_{|F|}(T|\mathcal{S}|) - 1.$$

However other coding theoretic bounds are known, for example, the *sphere packing bound* which says that $|F|^T \geq |\overline{\mathcal{S}}| V(\lfloor d_{\min}/2 \rfloor)$ where

$$V(r) = 1 + \binom{T}{1}|F| + \binom{T}{2}|F|^2 + \cdots + \binom{T}{r}|F|^r$$

is the volume of a Hamming sphere of radius r in the space F^T.

Other coding theoretic bounds are addressed in [40]. We remark that an earlier well-known paper [156] in this direction derives the following bound for any pair of sequences $\mathbf{a}, \mathbf{b} \in \mathcal{S}$ in a family $\mathcal{S}$ of sequences of period T,

$$\mathcal{C}^{\mathrm{Ham}}(\mathbf{a}, \mathbf{b}) \geq \frac{2ITm - (I+1)I|F|}{(Tm-1)m} \tag{11.11}$$

where $m = |\mathcal{S}|$ and $I = \lfloor Tm/|F| \rfloor$. Even though this bound appears to depend on the size of the family, it can be shown that (for $T > |F|$) this bound lies between $\alpha = \lfloor T/|F| \rfloor$ and $\alpha + 1$, so it is actually independent of the size of the family and in fact, it is essentially the same as the Lempel-Greenberger bound of Theorem 11.3.7.

Examples of optimal families of frequency hopping sequences are discussed in Section 14.9.

11.4 Exercises

1. What is the expected number of runs of length $T-1$ and of length T in sequences of period T over a finite alphabet A?

2. A periodic N-ary sequence is *strictly balanced* if the number of occurrences (in a single period) of each symbol is exactly the same. Prove: if $\mathbf{a}$ is strictly balanced then $Z(\mathbf{a}) = 0$. (Strictly balanced sequences are rare; for example the period must be divisible by N.)

3. Prove that if N is prime and if χ is a nontrivial character of $\mathbb{Z}/(N)$ then an N-ary sequence $\mathbf{a}$ of period $T = kN - 1$ (where $k \in \mathbb{Z}$) is balanced if and only if it is balanced with respect to χ, and it is strictly balanced if and only if it is strictly balanced with respect to χ, that is, $Z_\chi(\mathbf{a}) = 0$.

4. This exercise shows how many terms of a difference must be computed in order to find arithmetic correlations.

Let $\mathbf{b}$ and $\mathbf{c}$ be periodic N-ary sequences with period T. Let b and c be the N-adic numbers associated with $\mathbf{b}$ and $\mathbf{c}$. Let $\mathbf{d} = d_0, d_1, \cdots$ be the sequence associated with $b - c$. Prove that $\mathbf{d}$ is strictly periodic from at least d_T on.

5. Fix a finite set F and an integer $T \geq 2$. Consider the class $\mathcal{P}$ of all periodic sequences of period T with entries in F. For any shift $\tau \neq 0$ show that the expected Hamming autocorrelation among such sequences is

$$E\left[\mathcal{A}_{\mathbf{a}}^{\mathrm{Ham}}(\tau)\right] = T/|F|.$$

6. With $F, T, \mathcal{P}$ as in the previous exercise, show that the expected Hamming cross-correlation between any two sequences $\mathbf{a}, \mathbf{b} \in \mathcal{P}$ is

$$E\left[\mathcal{C}_{\mathbf{a},\mathbf{b}}^{\mathrm{Ham}}(\tau)\right] = T/|F|.$$

Chapter 12 Shift and Add Sequences

Shift and add sequences are important because (a) they arise naturally as m-sequences (see Section 6.8.a and Section 13) or related sequences (see Chapter 14, 15), (b) they often have ideal autocorrelation properties (see Proposition 12.1.3), and (c) they are often (punctured) de Bruijn sequences (see Theorem 12.4.1). At one time it was thought that all shift and add sequences were m-sequences, but this has turned out to be false ([196, 56, 12]; see the discussion in Section 12.2 below), and in this chapter we develop a complete description of the set of all shift and add sequences. In Chapter 15 we describe a class of (algebraic) shift registers that may be used to generate "good" shift and add sequences over non-prime fields. In this chapter we also consider a with-carry version of the shift and add property and develop a complete description of all sequences with this property.

12.1 Basic properties

Let G be a finite Abelian group and let $\mathbf{a} = (a_0, a_1, \cdots)$ be a periodic sequence of elements from G. Let T be the minimal period of $\mathbf{a}$. For any integer τ, $0 \le \tau < T$, let $\mathbf{a}^\tau$ be the τ-shift of $\mathbf{a}$, that is, $\mathbf{a}^\tau = (a_\tau, a_{\tau+1}, \cdots)$. If $\mathbf{b}$ is another periodic sequence with period T, then let $\mathbf{a}+\mathbf{b}$ be the sequence $(a_0+b_0, a_1+b_1, \cdots)$. If R is a ring, $\mathbf{a}$ and $\mathbf{b}$ are sequences of elements in R, and $\alpha, \beta \in R$ (or even $\alpha, \beta \in M$, a module over R), then we denote by $\alpha\mathbf{a} + \beta\mathbf{b}$ the sequence $(\alpha a_0 + \beta b_0, \alpha a_1 + \beta b_1, \cdots)$.

Definition 12.1.1. *The sequence* $\mathbf{a}$ *has the* shift and add property *if for any shift* τ, $0 \le \tau < T$, *either*

1. $\mathbf{a} + \mathbf{a}^\tau = \mathbf{0}$ *(the all-zero sequence) or*
2. *there exists another shift* τ', $0 \le \tau' < T$, *such that* $\mathbf{a} + \mathbf{a}^\tau = \mathbf{a}^{\tau'}$.

If $\mathbf{a}$ is a shift and add sequence and T is the minimal period of $\mathbf{a}$, then the number τ' is uniquely determined by τ and we define the *shift and add rule* $H : \{0, 1, \cdots, T-1\} \to \{0, 1, \cdots, T-1\}$ to be the corresponding function $\tau' = H(\tau)$.

We may similarly define the shift and subtract property. More generally, if $\mathbf{a}$ is a sequence of elements from some module M over a (commutative) ring R we say that $\mathbf{a}$ has the *shift and add property with coefficients in* R if for any $\alpha, \beta \in R$, and for any shift τ, either

$$\alpha\mathbf{a} + \beta\mathbf{a}^\tau = 0$$

or there exists a shift τ' such that

$$\alpha\mathbf{a} + \beta\mathbf{a}^\tau = \mathbf{a}^{\tau'}.$$

That is, the set of left shifts of $\mathbf{a}$ together with the all zero sequence is a module over R.

Lemma 12.1.2. *Suppose G is a finite Abelian group and $\mathbf{a}$ is a periodic sequence of elements of G. Then the following are equivalent:*

1. *$\mathbf{a}$ has the shift and add property.*
2. *$\mathbf{a}$ has the shift and subtract property.*
3. *$\mathbf{a}$ has the shift and add property with coefficients in $\mathbb{Z}$.*

Proof. Every Abelian group is a module over $\mathbb{Z}$, so condition (3) makes sense. Condition (3) implies conditions (1) and (2).

Suppose that $\mathbf{a}$ has the shift and add property. For any element $g \in G$ the sum $g+g+\cdots+g$ ($|G|-1$ times) equals $-g$. Let τ be any shift. Then $\mathbf{a}+\mathbf{a}^\tau+\mathbf{a}^\tau+\cdots+\mathbf{a}^\tau = \mathbf{a}-\mathbf{a}^\tau$ (where $\mathbf{a}^\tau$ occurs $|G|-1$ times). By repeatedly applying the shift and add property, we conclude there exists a shift τ' such that $\mathbf{a}-\mathbf{a}^\tau = \mathbf{a}^{\tau'}$. Thus $\mathbf{a}$ has the shift and subtract property. The converse is similar, so condition (1) is equivalent to condition (2). Now any sum $\alpha\mathbf{a}+\beta\mathbf{a}^\tau$ with $\alpha,\beta \in \mathbb{Z}$ can be written as a series of sums and differences of shifts of $\mathbf{a}$, so condition (3) holds as well. □

Proposition 12.1.3. *Let G be a finite Abelian group and let $\mathbf{a}$ be a periodic sequence of elements of G, with some period T. Suppose $\mathbf{a}$ has the shift and add property. Then the shifted autocorrelations of $\mathbf{a}$ with respect to any nontrivial character χ of G are*

$$\mathcal{A}_{\mathbf{a}}(\tau) = Z_\chi(\mathbf{a}),$$

where $Z_\chi(\mathbf{a})$ is the imbalance of $\mathbf{a}$ (see Definition 11.2.1).

Proof. Fix a shift τ and a nontrivial character χ and compute the autocorrelation,

$$\mathcal{A}_{\mathbf{a}}(\tau) = \sum_{i=0}^{T-1} \chi(a_i)\overline{\chi}(a_{i+\tau}) = \sum_{i=0}^{T-1} \chi(a_i - a_{i+\tau}) = \sum_{i=\tau'}^{\tau'+T-1} \chi(a_i) = Z_\chi(\mathbf{a})$$

for some shift τ' since $\mathbf{a}$ satisfies the shift and subtract property by Lemma 12.1.2. □

Corollary 12.1.4. *Let G be a finite Abelian group and let $\mathbf{a}$ be a periodic sequence of elements of G, with some period T. Suppose*

1. *$T \equiv 0, +1, \text{ or } -1 \pmod{|G|}$,*
2. *the sequence $\mathbf{a}$ is balanced (i.e., equidistributed to order 1), and*
3. *the sequence $\mathbf{a}$ has the shift and add property.*

Then $\mathbf{a}$ has ideal autocorrelations (i.e., $|\mathcal{A}_{\mathbf{a}}(\tau)| \leq 1$ for $\tau \neq 0$, cf. Section 11.3.a).

Proof. Consider the case $T \equiv 1 \pmod{|G|}$; the other two cases are similar. Since $\mathbf{a}$ is balanced, each element in G occurs exactly $\lfloor T/|G| \rfloor$ times in a period except for one element g_0, which occurs $\lfloor T/|G| \rfloor + 1$ times. By Propositions 2.3.2 and 12.1.3, the norm of the autocorrelation is then

$$|\mathcal{A}_{\mathbf{a}}(\tau)| = |Z_\chi(\mathbf{a})| = |\, \lfloor T/|G| \rfloor \cdot 0 + \chi(g_0)| = 1.$$

□

The following lemma is needed in Section 15.1.c where we consider the shift register generation of punctured de Bruijn sequences satisfying the shift and add property.

Lemma 12.1.5. *Let V be a vector space over $\mathbb{F}_p$ and let $\mathbf{a} = a_0, a_1, \cdots$ be a periodic sequence of elements in V. Suppose $\mathbf{a}$ is a punctured de Bruijn sequence with the shift-and-add property. Let $\hat{V}$ be another vector space over $\mathbb{F}_p$ and let $\phi : V \to \hat{V}$ be a set theoretic mapping. Then the following conditions are equivalent:*

1. *The sequence $\phi(\mathbf{a})$ is a punctured de Bruijn sequence with the shift and add property.*
2. *The mapping $\phi : V \to \hat{V}$ is a (linear) isomorphism of vector spaces.*

Proof. It is straightforward to see that condition (2) implies condition (1), so let us assume condition (1) and prove condition (2). Since $\hat{\mathbf{a}} = \phi(\mathbf{a})$ is a punctured de Bruijn sequence, the mapping ϕ must be a one to one correspondence so $\dim(V) = \dim(\hat{V})$. In particular $\phi(0) = 0$ because 0 occurs in $\mathbf{a}$ (and in $\hat{\mathbf{a}}$) fewer times than the other symbols. For every shift τ there exists a unique shift $k = k(\tau)$ such that

$$\phi(\mathbf{a} + \mathbf{a}^\tau) - \phi(\mathbf{a}) = \phi(\mathbf{a}^{k(\tau)}).$$

This follows from the facts that $\mathbf{a}$ is a shift and add sequence, $\hat{\mathbf{a}}$ is a shift and subtract sequence,

$$\mathbf{a} + \mathbf{a}^\tau \neq \mathbf{a},$$

and for all i, $\phi(\mathbf{a}^i) = \phi(\mathbf{a})^i$. So for each i,

$$\phi(a_i + a_{i+\tau}) = \phi(a_i) + \phi(a_{i+k(\tau)}).$$

To show that ϕ is linear it now suffices to prove that $k(\tau) = \tau$ for all τ. Suppose there exists an index τ such that $k(\tau) \neq \tau$. Then whenever i satisfies $a_{i+\tau} = 0$ we obtain

$$\phi(a_{i+k(\tau)}) = 0.$$

In other words, if $a_\ell = 0$ then

$$a_{\ell+k(\tau)-\tau} = 0.$$

The sequence $\mathbf{a}$ contains a unique largest block of zeroes (with $k-1$ zeroes, where k is the span of the de Bruijn sequence). Applying the above implication to each of these zeroes gives another (possibly overlapping) block of $k-1$ zeroes. This is a contradiction unless these two blocks coincide, meaning that $k(\tau) = \tau$. □

12.2 Characterization of shift and add sequences

N. Zierler stated that the sequences over a finite field with the shift and add property are exactly the m-sequences [196]. His proof is valid for sequences over a prime field $\mathbb{F}_p$, but it is incorrect for sequences over non-prime fields. Gong, Di Porto, and Wolfowicz gave the first counterexamples [56]. Subsequently, Blackburn [12] gave a complete characterization of shift and add sequences. In Theorems 12.2.1 and 12.2.2 below, we describe the results of Zierler and Blackburn. Throughout this section we fix a prime number p.

Theorem 12.2.1. [196] *Let* $\mathbf{a} = (a_0, a_1, \cdots)$ *be a nonzero periodic shift and add sequence with entries* $a_i \in \mathbb{F}_p$. *Then* $\mathbf{a}$ *is an m-sequence (cf.* Section *6.8.a,* Section *13).*

Proof. Consider the set A consisting of the sequence $\mathbf{a}$, all its shifts, and the sequence $\mathbf{0}$ consisting of all zeroes. Then A forms a vector space (under termwise addition of sequences) that is closed under the shift operation. By Theorem 6.7.5 part (7) and the fact that $\mathcal{F}_p[x]$ is a principal ideal domain, there exists a polynomial $q(x)$ such that $A = G(q)$, the set of all sequences that satisfy the linear recurrence defined by q. By Lemma 6.7.4, $G(q)$ has $p^{\deg(q)}$ elements, so the period of $\mathbf{a}$ is

$$T = p^{\deg(q)} - 1.$$

By Proposition 6.6.5 the order of q is a multiple of T. But the order of q is no more than $p^{\deg(q)} - 1$ so the order of q is exactly $p^{\deg(q)} - 1$, which implies by Lemma 3.2.10) that q is a primitive polynomial. Therefore $\mathbf{a}$ is an m-sequence. □

Theorem 12.2.2. [12] *Let* V *be a vector space of dimension* e *over* $\mathbb{F}_p$ *and let* $\mathbf{a} = (a_0, a_1, \cdots)$ *be a periodic shift and add sequence with entries in* V. *Then the (minimal) period of* $\mathbf{a}$ *is* $p^n - 1$ *for some* n. *Moreover, if* $F = \mathbb{F}_{p^n}$, *then there exists a primitive element* $\alpha \in F$ *and there exists an* $\mathbb{F}_p$*-linear mapping* $L : F \to V$ *such that* $a_i = L(\alpha^i)$ *for* $i = 0, 1, 2, \cdots$.

Proof. A choice of basis for V gives an isomorphism

$$f = (f_1, f_2, \cdots, f_e) : V \to (\mathbb{F}_p)^e$$

where each $f_j : V \to \mathbb{F}_p$ is linear. Then each sequence $\mathbf{b}_j = f_j(\mathbf{a})$ is either $\mathbf{0}$ or it is an m-sequence by Zierler's theorem (Theorem 12.2.1), whose minimal polynomial, q_j, is primitive. We claim the q_j that arise in this way are all equal. For, consider any two of these (nonzero) sequences, which (by renumbering if necessary) we may denote by $\mathbf{b}_1 = f_1(\mathbf{a})$ and $\mathbf{b}_2 = f_2(\mathbf{a})$. The sum $\mathbf{b}_1 + \mathbf{b}_2 = (f_1 + f_2)(\mathbf{a})$ is also a shift and add sequence over $\mathbb{F}_p$ so its minimum polynomial is primitive and irreducible. By Theorem (6.7.5) parts (2) and (4) the minimum polynomial of $\mathbf{b}_1 + \mathbf{b}_2$ divides the polynomial $gcd(q_1, q_2)$. In other words, it is either 1, q_1, q_2, or $q_1 q_2$. But it is irreducible so it cannot be $q_1 q_2$. On the other hand, if $q_1 \neq q_2$ then the sequence $\mathbf{b}_1 + \mathbf{b}_2$ does not

satisfy the linear recurrence defined by q_1, nor does it satisfy the linear recurrence defined by q_2. This leaves only the possibility that $q_1 = q_2$. This proves that each of the (nonzero) coordinate sequences $f_j(\mathbf{a})$ comes from the same primitive polynomial q. Moreover each coordinate sequence has the same period, $p^n - 1$ where $n = \deg(q)$. Therefore $\mathbf{a}$ has period $p^n - 1$ also.

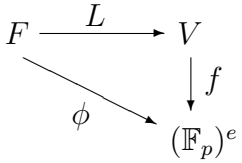
Let $F = \mathbb{F}_{p^n}$, let $L_0 : F \to \mathbb{F}_p$ be a surjective linear mapping (such as the trace), and let $\alpha^{-1} \in F$ be a root of the above minimal polynomial q. If the sequence $\mathbf{b}_j$ is nonzero, then by Proposition 6.6.5 there exists $u_j \in F$ such that the m-sequence $\mathbf{b}_j$ is given by

$$\mathbf{b}_j = (L_0(u_j\alpha^0), L_0(u_j\alpha^1), L_0(u_j\alpha^2), \cdots).$$

If $\mathbf{b}_j = \mathbf{0}$ set $u_j = 0$. Putting these together we obtain a mapping $\phi : F \to (\mathbb{F}_p)^e$ given by

$$\phi(x) = (L_0(u_1x), L_0(u_2x), \cdots, L_0(u_ex)).$$

In the diagram



put

$$L = f^{-1} \circ \phi : F \to V.$$

Then $A = (L(\alpha^0), L(\alpha^1), L(\alpha^2), \cdots)$ □

We remark that Theorem 12.2.2 applies equally well to sequences with values in a vector space V over an arbitrary finite field of characteristic p because V is also a vector space over the prime field $\mathbb{F}_p$.

12.3 Examples of shift and add sequences

12.3.a Window construction

Let G be a finite Abelian group and let $\mathbf{a} = (a_0, a_1, \cdots)$ be a shift-and-add sequence with entries $a_i \in G$ and with period T. For any positive integer j, define the *window sequence* (with window size j)to be the sequence $\mathbf{A} = (A_0, A_1, \cdots)$ of vectors $A_i \in G^j$ by setting

$$A_i = (a_i, a_{i+1}, \cdots, a_{i+j-1}) \in G^j,$$

that is, A_i is the "window" of length j beginning at a_i in the sequence $\mathbf{a}$. Then $\mathbf{A}$ is a shift and add sequence (of period T) because each of its coordinate sequences is a shift and add sequence *with the same shift and add rule.*

Suppose the original sequence $\mathbf{a}$ is equidistributed to order r (see Section 11.2.a), meaning that for any any block b of size $j \le r$, the number of occurrences of b within a single period of the sequence $\mathbf{a}$ is either $\lfloor T/|G|^j \rfloor$ or $\lceil T/|G|^j \rceil$. Suppose $j \le r$. Then the window sequence $\mathbf{A}$ is balanced (see Section 11.2.a) although it likely fails to be equidistributed to any degree greater than one because each symbol is a essentially shift of the preceding one.

12.3.b From m-sequences

Let $\mathbf{a} = (a_0, a_1, \cdots)$ be an m-sequence with entries in a finite field $F = \mathbb{F}_q$ and with period $q^n - 1$. Then $\mathbf{a}$ is a shift and add sequence that is equidistributed to order n (see Proposition 13.1.2). By applying the window construction with windows of size $j \le n$ one obtains a balanced shift and add sequence $\mathbf{A} = (A_0, A_1, \cdots)$ with $A_i \in F^j$, with the same period $q^n - 1$. Each symbol $A \in F^j$ occurs exactly q^{n-j} times in a single period of $\mathbf{A}$ except for the all-zero symbol, which occurs $q^{n-j} - 1$ times.

12.3.c From GMW sequences

Let $K \subset E \subset F$ be finite fields of characteristic p and let h be a power of p. Let $\mathbf{a} = (a_0, a_1, \cdots)$ be the GMW sequence

$$a_i = \mathsf{Tr}_K^E \left((\mathsf{Tr}_E^F(\alpha^i))^h \right)$$

where $\alpha \in F$ is a primitive element. According to Theorem 14.6.1 the sequence $\mathbf{a}$ is a balanced shift-and-add sequence of period $|F| - 1$ that is equidistributed to order $[F : E]$. Applying the window construction with windows of size $j \le [F : E]$ gives a balanced shift-and-add sequence of period $|F| - 1$ of elements in K^j.

12.3.d From function field sequences

Let $\mathbf{a}$ be a (π, q)-*adic* ℓ-*sequence* (see Section 15.1.c) where $\pi(x), q(x) \in F[x]$ are irreducible polynomials over a field F, $\pi(x)$ is primitive mod $q(x)$, and $\deg(q) = k \deg(\pi)$ for some integer $k \ge 1$. In other words, $\mathbf{a}$ is a maximal length function field sequence whose symbols lie in the field $F[x]/(\pi)$ with $|F|^{\deg(\pi)}$ elements. According to Proposition 15.1.8 and Theorem 15.1.9, the sequence $\mathbf{a}$ is a shift-and-add sequence of period $|F|^{\deg(q)} - 1$ that is equidistributed to order k. Applying the window construction to this sequence with windows of size $j \le k$ gives a balanced shift-and-add sequence with elements in $(F[x]/(\pi))^j$ and period $|F|^{\deg(q)} - 1$.

12.4 Further properties of shift and add sequences

Let V be a vector space of dimension e over $\mathbb{F}_p$. We consider periodic sequences of period $T = p^n - 1$ with entries in V. Let $\alpha \in \mathbb{F}_{p^n}$ be a primitive element and let $L : \mathbb{F}_{p^n} \to V$ be a set-theoretic

mapping that is not identically 0. Let $A = (a_0, a_1, \ldots)$ be the sequence given by $a_i = L(\alpha^i)$.

We say that L is *balanced* if $n \geq e$ and if the set $L^{-1}(a)$ contains the same number, p^{n-e}, of elements, for every $a \in V$. If L is balanced, then it is surjective. If L is linear over $\mathbb{F}_p$, denote by $K = \mathrm{Ker(L)}$ the kernel of L. If $u \in \mathbb{F}_{p^n}$ then denote

$$uK = \{ux \in \mathbb{F}_{p^n} : x \in K\} = \{ux \in \mathbb{F}_{p^n} : L(x) = 0\}$$

the translate of this subspace by the action of multiplication by u. We say that L has the *kernel property* if

$$L \text{ is linear, } k|n, \text{ and } \bigcap_{i=0}^{k-1} \alpha^{-i} K = \{0\}. \tag{12.1}$$

If in doubt, we refer to equation (12.1) as the kernel property for L with respect to α^{-1}. If L has the kernel property then L is surjective (see the proof of part (3) of Theorem 12.4.1 in Section 12.5). See also the remark at the end of Section 12.5.

In Theorem 12.4.1, we show that properties of the mapping $L : \mathbb{F}_{p^n} \to V$ give rise to properties of the resulting sequence $\mathbf{a}$ according to Table 12.1.

Properties of L	Properties of $\mathbf{a}$
$\mathbb{F}_p$-linear	shift and add
balanced	ideal autocorrelations
kernel property	de Bruijn

Table 12.1: Properties of L versus properties of $\mathbf{a}$

Recall that two periodic sequences of the same period are shift distinct if the second sequence cannot be realized as a shift of the first sequence. In the following theorem, ϕ denotes Euler's function.

Let V be a vector space of dimension e over $\mathbb{F}_p$. Let $L : \mathbb{F}_{p^n} \to V$ be a set-theoretic mapping that is not identically 0. Let $\alpha \in \mathbb{F}_{p^n}$ be a primitive element. Let

$$\mathbf{a}(L, \alpha) = a_0, a_1, \cdots$$

denote the sequence with $a_i = L(\alpha^i)$.

Theorem 12.4.1. *The following statements hold.*

(1) The mapping L is $\mathbb{F}_p$-linear if and only if the sequence $\mathbf{a}(L, \alpha)$ is a shift-and-add sequence for any primitive $\alpha \in \mathbb{F}_{p^n}$, and in this case its (minimum) period is $p^n - 1$. There are

$$\frac{(p^{ne} - 1)\phi(p^n - 1)}{n(p^n - 1)}$$

shift distinct nonzero sequences (of elements in V) with (minimum) period $p^n - 1$ which satisfy the shift and add property. Each of these arises from a pair (L, α), where $\alpha \in \mathbb{F}_{p^n}$ is primitive and $L : \mathbb{F}_{p^n} \to V$ is $\mathbb{F}_p$-linear.

(2) Suppose the mapping L is $\mathbb{F}_p$-linear. Then $L : \mathbb{F}_{p^n} \to V$ is surjective if and only if it is balanced, which holds if and only if the sequence $\mathbf{a}(L, \alpha)$ has ideal autocorrelations for any primitive $\alpha \in \mathbb{F}_{p^n}$. There are

$$\frac{(p^n - p)(p^n - p^2) \cdots (p^n - p^{e-1})\phi(p^n - 1)}{n}$$

shift distinct shift-and-add sequences (of elements in V) with ideal autocorrelations and minimal period $p^n - 1$. Each of these arises from such a pair (L, α), where $\alpha \in \mathbb{F}_{p^n}$ is primitive and $L : \mathbb{F}_{p^n} \to V$ is $\mathbb{F}_p$-linear and balanced.

(3) Suppose $L : \mathbb{F}_{p^n} \to V$ is $\mathbb{F}_p$-linear and surjective. Then for any primitive $\alpha \in \mathbb{F}_{p^n}$, $\mathbf{a}(L, \alpha)$ is a shift and add punctured de Bruijn sequence of rank k if and only if $n = ek$ and L has the kernel property.

(4) Let $\mathbf{a}$ and $\mathbf{b}$ be two shift and add punctured de Bruijn sequences with the same period $p^n - 1$, with symbols drawn from V. Let $L, M : \mathbb{F}_{p^n} \to V$ be surjective linear maps and let

$$\alpha, \beta \in \mathbb{F}_{p^n}$$

be primitive elements such that $\mathbf{a} = \mathbf{a}(L, \alpha)$ and $\mathbf{b} = \mathbf{a}(M, \beta)$. If $\mathbf{a}$ and $\mathbf{b}$ are isomorphic (possibly after a shift) then the primitive elements $\alpha, \beta \in \mathbb{F}_{p^n}$ are Galois conjugate.

(5) Suppose $L : \mathbb{F}_{p^n} \to V$ is $\mathbb{F}_p$-linear and suppose that $n = ek$ for some k. Choose a basis for V and write

$$L(x) = (L_1(x), L_2(x), \cdots, L_e(x)) \tag{12.2}$$

for the resulting coordinates of $L(x)$. Each $L_j : \mathbb{F}_{p^n} \to \mathbb{F}_p$ is $\mathbb{F}_p$-linear so there exist (cf. Theorem 3.2.14 part (7)) unique elements $u_j \in \mathbb{F}_{p^n}$ such that

$$L_j(x) = Tr_{\mathbb{F}_p}^{\mathbb{F}_{p^n}}(u_j x). \tag{12.3}$$

Then L has the kernel property if and only if the following collection of ek elements

$$\{u_j \alpha^i : 1 \leq j \leq e,\ 0 \leq i \leq k - 1\}$$

forms a basis for $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$.

We need a lemma before giving the proof of Theorem 12.4.1.

Lemma 12.4.2. *Let d be a positive integer and let p be prime. Suppose $\alpha, \beta \in \mathbb{F}_{p^d}$ are primitive elements. Suppose $A : \mathbb{F}_{p^d} \to \mathbb{F}_p$ is a nonzero $\mathbb{F}_p$-linear mapping. Define the mapping $B : \mathbb{F}_{p^d} \to \mathbb{F}_p$ by $B(0) = 0$ and*

$$B(\beta^i) = A(\alpha^i)$$

for $0 \leq i \leq p^d - 2$. Then B is $\mathbb{F}_p$-linear if and only if α and β are Galois conjugates.

Proof. There exists an integer t such that $\alpha = \beta^t$. Therefore

$$B(\beta^i) = A(\alpha^i) = A(\beta^{ti})$$

so $B(x) = A(x^t)$ for all $x \in F_{p^d}$. If α and β are Galois conjugates then t is a power of p by Theorem 3.2.9, so the mapping $x \mapsto x^t$ is $\mathbb{F}_p$-linear. Therefore B is $\mathbb{F}_p$-linear. Conversely, suppose B is $\mathbb{F}_p$-linear. Let

$$f(x) = \sum_{i=0}^{d-1} a_i x^i$$

be an irreducible polynomial with coefficients $a_i \in \mathbb{F}_p$ such that $f(\alpha) = 0$. We claim that $f(\beta) = 0$. To show this it suffices to show that $B(\beta^t f(\beta)) = 0$ for all $t \geq 0$. But

$$B(\beta^t f(\beta)) = \sum_{i=0}^{d-1} a_i B(\beta^{t+i}) = \sum_{i=0}^{d-1} a_i A(\alpha^{t+i}) = A(\alpha^t p(\alpha)) = 0.$$

So by Theorem 3.2.9, α and β are Galois conjugate. □

12.5 Proof of Theorem 12.4.1

Proof of part (1). Suppose L is $\mathbb{F}_p$ linear and $\alpha \in \mathbb{F}_{p^n}$ is primitive. If τ is an integer with $0 \leq \tau < p^n - 1$, then

$$a_i + a_{i+\tau} = L(\alpha^i + \alpha^{i+\tau}) = L((1 + \alpha^\tau)\alpha^i).$$

Since α is primitive, there exists θ with

$$1 + \alpha^\tau = \alpha^\theta.$$

Therefore, for all i we have

$$a_i + a_{i+\tau} = a_{i+\theta},$$

so **a** is a shift and add sequence. Let $R : V \to \mathbb{F}_p$ be a nonzero $\mathbb{F}_p$-linear mapping. Then the composition

$$R \circ L : \alpha \in \mathbb{F}_{p^n} \to \mathbb{F}_p$$

is $\mathbb{F}_p$-linear so the sequence $R \circ L(\alpha^i) = R(a_i) \in \mathbb{F}_p$ is an m-sequence and has minimum period $p^n - 1$. Hence the sequence $\mathbf{a}$ has minimum period $p^n - 1$ also.

The converse is a slight modification of Blackburn's theorem. If $\alpha \in \mathbb{F}_{p^n}$ is any primitive element, then every sequence with period $2^n - 1$ and elements in V is of the form $\mathbf{a} = \mathbf{a}(L, \alpha)$ for some set theoretic map $L : \mathbb{F}_{p^n} \to V$ since the $p^n - 1$ powers of α are distinct. Suppose the sequence $\mathbf{a}$ is a shift and add sequence. We claim that then $L : \mathbb{F}_{p^n} \to V$ is $\mathbb{F}_p$-linear. Let $W \subset V$ be the image of $L(x)$. Since $\mathbf{a}$ is a shift and add sequence, the set W is a vector subspace of V of some dimension, say d. A choice of basis for W gives an isomorphism

$$f = (f_1, f_2, \cdots, f_d) : W \to (\mathbb{F}_p)^d$$

where each $f_j : W \to \mathbb{F}_p$ is linear and surjective. So each sequence $f_j(\mathbf{a}) = \mathbf{a}(f_j \circ T, \alpha)$ is a shift and add sequence of elements in $\mathbb{F}_p$ which, by Zierler's theorem (Theorem 12.2.1) is therefore an m-sequence. In other words, there exists an $\mathbb{F}_p$-linear mapping

$$\varphi_j : \mathbb{F}_{p^n} \to \mathbb{F}_p$$

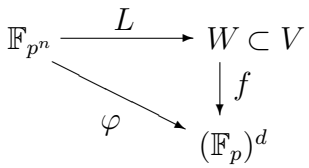
such that

$$\varphi_j(x) = f_j(L(x))$$

for all $x \in \mathbb{F}_{p^n}$. Putting these together gives a mapping

$$\varphi = (\varphi_1, \cdots, \varphi_d) : \mathbb{F}_{p^n} \to (\mathbb{F}_p)^d$$

so that the diagram



commutes. It follows that $L = f^{-1} \circ \varphi : \mathbb{F}_{p^n} \to W \to V$ is linear, as claimed.

To count the number of shift and add sequences with (minimal) period $p^n - 1$ we use Blackburn's theorem (Theorem 12.2.2): first count the number of pairs (L, α) where $\alpha \in \mathbb{F}_{p^n}$ is a primitive element and $L : \mathbb{F}_{p^n} \to V$ is $\mathbb{F}_p$-linear. Then determine when two such pairs define the same sequence.

The number of primitive elements $\alpha \in \mathbb{F}_{p^n}$ is $\phi(p^n - 1)$. To count the number of $\mathbb{F}_p$-linear mappings $L : \mathbb{F}_{p^n} \to V$ choose bases for both, as vector spaces of dimension n and e respectively, over $\mathbb{F}_p$. Each linear mapping L then corresponds to a unique $n \times e$ matrix with entries in $\mathbb{F}_p$, and there are p^{ne} such matrices. So there are $p^{ne} - 1$ nonzero linear mappings L.

Now consider decomposing the collection of pairs (L, α) into equivalence classes, with two pairs belonging to the same class if the resulting sequences are the same. In the next paragraph we show that each class contains exactly n pairs by showing that (L, α) and (M, β) belong to the same class if and only if α and β are Galois conjugates. Hence $\beta = \alpha^{p^t}$ for some t, and $M(x^{p^t}) = L(x)$. In other words, the mapping L is uniquely determined by M, α, and β.

So suppose $\mathbf{a}(L, \alpha) = \mathbf{a}(M, \beta)$. Then

$$M(\beta^i) = L(\alpha^i) \tag{12.4}$$

for all i. In particular the images of M and L coincide. Let $v \neq 0 \in V$ be in the image of M and L. Choose any linear mapping R from V to F_p such that $R(v) \neq 0$. Then R is surjective and both compositions $R \circ L$ and $R \circ M$ are nonzero.

Now we have the equation $R \circ L(\alpha^i) = R \circ M(\beta^i)$, for all i. Since both $R \circ T$ and $R \circ S$ are nonzero $\mathbb{F}_p$-linear mappings, Lemma 12.4.2 implies that α and β are Galois conjugates with $\beta = \alpha^{p^t}$ for some t. Therefore $M(x^{p^t}) = L(x)$ by equation (12.4).

Conversely, given (L, α) let $\beta \in \mathbb{F}_{p^n}$ be a Galois conjugate to α with $\beta = \alpha^{p^t}$. Let

$$M(x^{p^t}) = L(x).$$

Then $M : \mathbb{F}_{p^n} \to V$ is $\mathbb{F}_p$-linear and $M(\beta^i) = L(\alpha^i)$ for all i. Consequently $\mathbf{a}(L, \alpha) = \mathbf{a}(M, \beta)$.

To summarize, there are $p^{ne} - 1$ choices for L and $\phi(p^n - 1)$ choices for α. Such a pair determines a class consisting of the n Galois conjugates β of α, and uniquely determined $\mathbb{F}_p$-linear mappings M to go with them. This counts the total number of sequences; the shift distinct sequences are counted by dividing by the period, $p^n - 1$. This completes the proof of part (1) of Theorem 12.4.1.

Proof of part (2). If L is balanced then it is surjective. If L is surjective then it is balanced because $L^{-1}(a)$ is the set of solutions to a system of inhomogeneous linear equations, which is therefore a translate of the kernel $L^{-1}(0)$. Corollary 12.1.4 implies that $\mathbf{a}(L, \alpha)$ has ideal autocorrelations.

On the other hand, suppose that $\mathbf{a}(L, \alpha)$ has ideal autocorrelations, but L is not surjective. Let $I \subset V$ denote the image of the mapping $L : \mathbb{F}_{p^n} \to V$. Choose a complementary subspace $J \subset V$ so that $V \cong I \oplus J$. Then J has positive dimension. Let $\chi_1 : J \to \mathbb{C}^\times$ be any nontrivial character, and define $\chi : V \to \mathbb{C}^\times$ to by $\chi(a, b) = \chi_1(b)$. Then χ is a character of V, and $\chi(L(x)) = 1$ for all $x \in \mathbb{F}_{p^n}$. Let $\tau \neq 0$ be a nonzero shift. The autocorrelation of shift τ with respect to the character χ is

$$\sum_{i=0}^{p^n-2} \chi(L(\alpha^i - \alpha^{i+\tau})) = \sum_{i=0}^{p^n-2} 1 = p^n - 1$$

which is greater than 1. This is a contradiction, hence L is surjective.

To count the number of sequences with ideal autocorrelations, the same argument as in the proof of part (1) works, but we must count only those pairs (L, α) such that L has rank equal to e. Choosing bases for V and $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$, the mapping L may be represented as an $n \times e$ matrix of elements of $\mathbb{F}_p$. The matrices of rank e are counted by choosing the first row to be any nonzero vector ($p^n - 1$ choices), the second row to be any vector that is not in the span of the first vector ($p^n - p$ choices), the third row to be any vector that is not in the span of the first two vectors ($p^n - p^2$ choices), and so on. As in the proof of part (1) above, this counts the total number of sequences; the shift distinct sequences are counted by dividing this number by the period, $p^n - 1$. This completes the proof of part (2).

Proof of part (3). Let $\alpha \in \mathbb{F}_{p^n}$ be primitive. Suppose $\mathbf{a} = \mathbf{a}(L, \alpha)$ is a punctured de Bruijn sequence of rank k. Then the period of $\mathbf{a}$ is $|V|^k - 1 = p^n - 1$ so $n = ek$. Consider the mapping $\Phi : \mathbb{F}_{p^n} \to V^k$ given by $\Phi(x) = (L(x),\ L(\alpha x),\ \cdots,\ L(\alpha^{k-1}x))$. These k symbols form a block of the sequence $\mathbf{a}$. Therefore if $\mathbf{a}$ is a (punctured) de Bruijn sequence of rank k, then every nonzero k-tuple of vectors in V appears at some point in the sequence, so the mapping Φ is surjective. Since $|\mathbb{F}_{p^n}| = |V|^k$, the mapping Φ is surjective if and only if it is injective. But the kernel of Φ is exactly the intersection in equation (12.1). Therefore L has the kernel property. Conversely, if L has the kernel property and if $n = ek$, then Φ is surjective, so $\mathbf{a}$ is a punctured de Bruijn sequence of rank k.

Proof of part (4). Suppose there exists a shift τ and a set theoretic mapping $\psi : V \to V$ so that $\psi(\mathbf{a}) = \mathbf{b}^\tau$. We may assume that the shift $\tau = 0$ by replacing the mapping $L : \mathbb{F}_{p^n} \to V$ with a new mapping $L'(x) = L(ux)$ where $u = \alpha^\tau \in \mathbb{F}_{p^n}$. According to Lemma 12.1.5 the mapping ψ is automatically an $\mathbb{F}_p$-linear isomorphism of vector spaces. Let $\Phi : \mathbb{F}_{p^n} \to V^k$ and $\Psi : \mathbb{F}_{p^n} \to V^k$ be the vector space isomorphism considered above, that is,

$$\Phi(x) = \left(L(x), L(\alpha x), \cdots, L(\alpha^{k-1}x)\right),$$

and the analog for M,

$$\Psi(x) = \left(M(x), M(\beta x), \cdots, M(\beta^{k-1}x)\right),$$

Define $\Lambda : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ to be the mapping that makes the diagram

$$\begin{array}{ccc} \mathbb{F}_{p^n} & \xrightarrow{\Lambda} & \mathbb{F}_{p^n} \\ {\scriptstyle\Phi}\downarrow & & \downarrow{\scriptstyle\Psi} \\ V^k & \xrightarrow{\psi\times\cdots\times\psi} & V^k \end{array}$$

commute. Then $\Lambda : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ is an isomorphism of vector spaces. Moreover $\Lambda(\alpha x) = \beta\Lambda(x)$ for all $x \in \mathbb{F}_{p^n}$ because $\Phi(\alpha x)$ is obtained by "shifting" $\Phi(x) \in V^k$, and similarly for Ψ. Thus Λ

is almost a field isomorphism. However we do not know whether $\Lambda(1) = 1$. Let $u \in \mathbb{F}_{p^n}$ be the unique element such that $\Lambda(u) = 1$ and consider the diagram

$$\begin{array}{ccccc} \mathbb{F}_{p^n} & \xrightarrow{\cdot u} & \mathbb{F}_{p^n} & \xrightarrow{\Lambda} & \mathbb{F}_{p^n} \\ {\scriptstyle L'}\downarrow & & {\scriptstyle L}\downarrow & & \downarrow{\scriptstyle M} \\ V & = & V & \xrightarrow{\psi} & V \end{array}$$

where $L'(x) = L(ux)$. This diagram also commutes, and the composition $\Lambda' : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ across the top row is now a field isomorphism such that $\Lambda'(\alpha) = \beta$. This implies that α and β are Galois conjugate. This concludes the proof of part (4).

Proof of part (5). Having chosen a basis for V, the mapping $\Phi : \mathbb{F}_{p^n} \to V^k$ of the preceding paragraph may be expressed as a $k \times e$ matrix $\Phi(x) = [L_j(\alpha^i x)]$ with $0 \le i \le k-1$ and $1 \le j \le e$. Using equation (12.3) this becomes the matrix $\Phi(x) = [Tr_{\mathbb{F}_p}^{\mathbb{F}_{p^n}}(u_j \alpha^i x)]$. This mapping is an isomorphism if and only if the collection of linear functions

$$L_{ij}(x) = Tr_{\mathbb{F}_p}^{\mathbb{F}_{p^n}}(u_j \alpha^i x),$$

with $0 \le i \le k-1$ and $1 \le j \le e$, forms a basis of the dual space $\mathrm{Hom}_{\mathbb{F}_p}(\mathbb{F}_{p^n}, \mathbb{F}_p)$. However, the collection of vectors $\{u_j \alpha^i\}$ is linearly independent (and hence forms a basis of $\mathbb{F}_{p^n}$) if and only if the collection of linear functions L_{ij} is linearly independent. This completes the proof of Theorem 12.4.1. □

Remark. If $K = \mathrm{Ker(L)}$ is preserved under multiplication by elements from the sub-field $F = \mathbb{F}_{p^e} \subset \mathbb{F}_{p^n}$, then the kernel condition holds automatically. This occurs, for example, if $V = F = \mathbb{F}_{p^e}$ and if L is F-linear, in which case the resulting sequence is an m-sequence over F. More generally if $\sigma \in Gal(\mathbb{F}_{p^n}/\mathbb{F}_p)$ is an element of the Galois group and $\sigma(K)$ is preserved by multiplication by elements of F, then L has the kernel property. The kernel property is somewhat mysterious and we do not know of a simple method for counting the number of linear mappings L that have this property.

12.6 Arithmetic shift and add sequences

Following the theme of this book, it is natural to consider an arithmetic analog of the shift and add property. In this section we give some basic definitions and give a complete characterization of sequences with the arithmetic shift and add property.

Let $N \geq 2$ be a natural number and let $\mathbf{a} = a_0, a_1, \cdots$ be an infinite N-ary sequence. Let

$$a = \sum_{i=0}^{\infty} a_i N^i$$

be the N-adic number associated with $\mathbf{a}$. As usual, for any integer τ let $\mathbf{a}^\tau = a_\tau, a_{\tau+1}, \cdots$ be the left shift of $\mathbf{a}$ by τ positions and let $a^{(\tau)}$ be the N-adic number associated with $\mathbf{a}^\tau$. By the *N-adic sum* of $\mathbf{a}$ and $\mathbf{a}^\tau$, we mean the coefficient sequence of the N-adic number $a + a^{(\tau)}$; it is obtained from $\mathbf{a}$ and $\mathbf{a}^\tau$ using addition-with-carry.

A first attempt to define the arithmetic shift and add property would be to ask that the set of shifts of $\mathbf{a}$ be closed under N-adic addition, but this is too much to ask. Even if the sequence $\mathbf{a}$ is (strictly) periodic, the N-adic sum of $\mathbf{a}$ and $\mathbf{a}^\tau$ will only be eventually periodic, and in particular, repeated addition of $\mathbf{a}^\tau$ will result in infinitely many distinct (eventually periodic) sequences.

The solution is to consider only the periodic part of the sum of two sequences. This is consistent with how we have defined arithmetic correlations. In order to simplify the language, let us say that the *left shift of the N-adic number a* means the N-adic number associated with the left shift of the coefficient sequence of $\mathbf{a}$.

Definition 12.6.1. *The sequence* $\mathbf{a}$ *has the* arithmetic shift and add property *if for any shift $\tau \geq 0$, either*

1. *some left shift of $a + a^{(\tau)}$ is zero or*
2. *some left shift of $a + a^{(\tau)}$ equals a. That is, there is a $\tau' \geq 0$ so that $(a + a^{(\tau)})^{(\tau')} = a$.*

The left shift τ' allows us to ignore the effects of initial non-periodic parts of sequences. If we replace addition with carry in this definition by addition without carry, then we obtain a definition that is equivalent to the definition of the without-carry shift and add property. We may similarly define the shift and subtract property. The main result of this section is the following theorem.

Theorem 12.6.2. *An N-ary sequence has the arithmetic shift and add property if and only if it is the N-adic expansion of a rational number f/q where q is prime and N is primitive modulo q.*

These sequences, known as *ℓ-sequences*, are studied in Section 16.1. They are maximal period FCSR sequences: the arithmetic analogs of m-sequences.

Preliminaries to the proof. It is helpful here to use rational representations of sequences. If $\mathbf{a}$ is periodic with minimal period T, then for some $q, f \in \mathbb{Z}$ we have: $a = f/q$ with $-q \leq f \leq 0$ and $\gcd(q, f) = 1$. Then a shift $\mathbf{a}^\tau$ of $\mathbf{a}$ also corresponds to a rational number of this form, say $a^{(\tau)} = f_\tau/q$. Moreover, there is an integer c_τ so that

$$a^{(\tau)} = c_\tau + N^{T-\tau} a.$$

Thus $f_\tau = c_\tau q + N^{T-\tau} f$. It follows that

$$f_\tau \equiv N^{T-\tau} f \pmod{q}.$$

Therefore the set of numerators f_τ of the N-adic numbers associated with the left shifts of $\mathbf{a}$ is the fth cyclotomic coset

$$C_f = \left\{ N^i f \pmod{q} : \ i = 0, 1, 2, \cdots \right\} \subset \mathbb{Z}/(q)$$

relative to N (see Section 3.2.d). Now suppose that $(\mathbf{a} + \mathbf{a}^\tau)^{\tau'} = \mathbf{a}$. Let g/r be the rational representation of the N-adic number associated with $\mathbf{a} + \mathbf{a}^\tau$. Then

$$\frac{g}{r} = d + N^{\tau'} \frac{f}{q}$$

for some integer d. In particular, we can take $r = q$. Then $g = qd + N^{\tau'} f$. Thus $g \equiv N^{\tau'} f \pmod{q}$, so that g is in C_f.

Theorem 12.6.3. *Let q, f be relatively prime integers. Suppose the coefficient sequence $\mathbf{a}$ of the N-adic expansion of f/q is periodic. Then the sequence $\mathbf{a}$ has the arithmetic shift and add property if and only if $G = C_f \cup \{0\}$ is an additive subgroup of $\mathbb{Z}/(q)$.*

Proof. As above, let T be the minimal period of $\mathbf{a}$. It follows from the above discussion that if $\mathbf{a}$ has the arithmetic shift and add property, then $f + N^\tau f \pmod{q} \in G$ for every τ. Since G is closed under multiplication by N modulo $N^T - 1$, we also have $N^\sigma f + N^\tau f \pmod{q} \in G$ for every σ and τ. Since every nonzero element of G is of the form $N^\sigma f$ modulo q, it follows that G is closed under addition. Therefore it is a subgroup.

Conversely, suppose G is a subgroup of $\mathbb{Z}/(q)$. If $f = 0$ or $f = -q$, then $\mathbf{a}$ has the shift and add property, so we assume that $-q < f < 0$. For every τ there is a τ' so that $f + N^{T-\tau} f \equiv 0 \pmod{q}$ or $f + N^{T-\tau} f \equiv N^{\tau'} f \pmod{q}$. In the latter case

$$N^{T-\tau'}(f + N^{T-\tau} f) = N^T f + zq = f + wq,$$

where z and w are integers, the latter because q divides $N^T - 1$. Therefore

$$N^{T-\tau'} \left(\frac{f}{q} + c_\tau + N^{T-\tau} \frac{f}{q} \right) = \frac{f}{q} + y,$$

where y is an integer. It follows that

$$(a + a^{(\tau)})^{(\tau')} = a + u$$

for some integer u. But adding an integer to a periodic N-adic integer does not change the periodic part unless the periodic part is all zero or all $N-1$, and this would mean that $f = 0$ or $f = -q$. Thus for sufficiently large k we have

$$(a + a^{(\tau)})^{(kT+\tau')} = a.$$

The argument in the case when $f + N^{T-\tau} f \equiv 0 \pmod{q}$ is similar — a sufficiently large shift of $a + a^{(\tau)}$ is zero. This proves the theorem. □

Corollary 12.6.4. *A sequence has the arithmetic shift and add property if and only if it has the arithmetic shift and subtract property.*

Proof of Theorem 12.6.2. How can it be that $C_f \cup \{0\}$ is an additive subgroup of $\mathbb{Z}/(q)$? It implies, in particular, that $C_f \cup \{0\}$ is closed under multiplication by integers modulo q. We have taken q so that $\gcd(f, q) = 1$. Thus f is invertible $\pmod{q}$ so $C_f \cup \{0\} = \mathbb{Z}/(q)$, that is, $C_f = \mathbb{Z}/(q) - \{0\}$. In particular, every nonzero element of $\mathbb{Z}/(q)$ is of the form $N^i f$, hence is a unit. Thus q is prime, and N is a primitive element modulo q. □

12.7 Exercises

1. Let G be a finite Abelian group and let $\mathbf{a}$ be a periodic sequence of elements of G with period $T \equiv k \pmod{|G|}$ such that $\mathbf{a}$ is is equidistributed to order 1. What can be said about the autocorrelations of $\mathbf{a}$?

2. Let F be a finite field. Let $\mathbf{a}$ and $\mathbf{b}$ be a pair of period sequences over F (possibly with different periods). Let U be the set of F-linear combinations of shifts of $\mathbf{a}$ and $\mathbf{b}$,

$$U = \left\{ \sum_{\tau} x_\tau \mathbf{a}^\tau + \sum_{\sigma} y_\sigma \mathbf{b}^\sigma, x_\tau, y_\sigma \in F \right\}.$$

Let

$$V = \{\epsilon_1 \mathbf{a}^\tau + \epsilon_2 \mathbf{b}^\sigma, \epsilon_1, \epsilon_2 \in \{0, 1\}\}.$$

a. Prove that if $\mathbf{a}$ and $\mathbf{b}$ are m-sequences, then $U = V$.
b. Prove that if $U = V$, then $\mathbf{a}$ and $\mathbf{b}$ are m-sequences.
c. Generalize these statements to more than two sequences.

3. Let N and d be natural numbers so that $x^d - N$ is irreducible over $\mathbb{Z}$.

a. Define a notion of N-ary d-arithmetic shift and add sequences, where addition of sequences corresponds to addition of π-adic numbers.

b. Prove that an N-ary sequence has the d-arithmetic shift and add property if and only if it is the π-adic expansion of a rational element f/q where q is irreducible and N is primitive modulo q.

Chapter 13 M-Sequences

13.1 Basic properties of m-sequences

Let R be a finite commutative ring and let $\mathbf{a}$ be a periodic sequence of elements in R. Let T be the period of $\mathbf{a}$.

Definition 13.1.1. *The sequence* $\mathbf{a}$ *is an* m-sequence *(over the ring* R*) of rank* r *(or degree* r*) if it can be generated by a linear feedback shift register with* r *cells, and if every nonzero block of length* r *occurs exactly once in each period of* $\mathbf{a}$*.*

In other words, the sequence $\mathbf{a}$ is the output sequence of a LFSR that cycles through all possible nonzero states before it repeats. The second condition in the definition also says that $\mathbf{a}$ is a punctured de Bruijn sequence (see Section 11.2.d). In this section we recall standard results about m-sequences which have been known since [37] and which may be found in [53] and [54].

Proposition 13.1.2. *If* $\mathbf{a}$ *is an m-sequence of rank* r *over a (finite commutative) ring* R*, generated by a LFSR with connection polynomial* $q(x) \in R[x]$*, then*

1. *The ring* R *is a field and the connection polynomial* $q(x)$ *is a primitive polynomial with* $\deg(q) = r$*.*
2. *The polynomial* $q(x)$ *splits into* r *linear factors over the unique field extension* E *of* R *such that* $\deg(E/R) = r$*. If* $L : E \to R$ *is any surjective* R*-linear mapping and if* $\alpha \in E$ *is a root of* $q(x)$ *then there exists* $A \in E$ *such that for all* i*,*

$$a_i = L(A\alpha^i). \tag{13.1}$$

3. *The sequence* $\mathbf{a}$ *is a punctured de Bruijn sequence, so it is balanced, equidistributed to order* r*, and has the run property (see* Section *11.2.b).*
4. *The sequence* $\mathbf{a}$ *satisfies the shift and add condition, so it has ideal autocorrelation function.*
5. *In the case* $R = \mathbb{Z}/(2)$ *let* $C \subset (\mathbb{Z}/(2))^r$ *be the linear block code spanned by a single period of the sequence* $\mathbf{a}$ *and all of its left shifts, and by the complement of the sequence* $\mathbf{a}$ *and all of its left shifts. Then* C *is equivalent to the (punctured) first-order Reed-Muller code of wordlength* $2^r - 1$*.*

Proof. By Theorem 6.6.2 the states of the LFSR can be modelled on the ring $E = R[x]/(q)$ where $q(x)$ is the connection polynomial of the LFSR, in such a way that the state change operation

corresponds to multiplication by a fixed element $\alpha = x^{-1}$ in $E = R[x]/(q)$. Since the shift register cycles through all nonzero states before repeating, it follows that the powers $\{\alpha^i\}$ of this element account for all the nonzero elements in E. In particular, every element in E is invertible, hence E is a field and $\alpha \in E$ is a primitive element. The ring R may be identified as the subring of E consisting of polynomials of degree 0, so R must also be a field. Finally, the powers of α^{-1} also account for all the nonzero elements of E, but $\alpha^{-1} = x$ is a root of $q(x)$ in E so $q(x)$ is a primitive polynomial. Part (2) is just a restatement of Theorem 6.6.4.

The sequence $\mathbf{a}$ is a punctured de Bruijn sequence by definition, so part (3) is just a restatement of Proposition 11.2.4. Every LFSR sequence satisfies the shift and add condition, so part (4) is just a restatement of Proposition 11.2.4.

For part (5), recall that the first order Reed-Muller code (with wordlength 2^r) consists of codewords, each of which is the incidence vector of a hyperplane $H \subset (\mathbb{Z}/(2))^r$ (not necessarily through the origin), together with the all-zeroes and the all-ones codewords. Choose a $\mathbb{Z}/(2)$-linear isomorphism $\mathbb{F}_{2^r} \cong (\mathbb{Z}/(2))^r$. Then the nonzero elements in $(\mathbb{Z}/(2))^r$ become labeled by the elements $1, \alpha, \cdots, \alpha^{2^r-2}$ so the incidence vector of a hyperplane H is the vector $L(1), L(\alpha), L(\alpha^2), \cdots, L(\alpha^{2^r-2})$ where $L : \mathbb{F}_2^r \to \mathbb{Z}/(2)$ is either a $\mathbb{Z}/(2)$ linear map (in which case the hyperplane H does not contain the origin) or $1 + L$ is a linear map (in which case H does contain the origin). In the first case, this incidence vector is a shift of the m-sequence $\mathbf{a}$ and in the second case it is the complement of a shift of $\mathbf{a}$. □

In particular, m-sequences satisfy Golomb's three randomness properties. Golomb made the following conjecture, which is still considered open.

Conjecture 13.1.3. *The only binary sequences (meaning $R = \mathbb{F}_2$) satisfying the three randomness postulates are m-sequences.*

There exist non-binary sequences that are not m-sequences but which still satisfy Golomb's three randomness properties. See Chapter 15 and [56].

13.2 Decimations

Let G be an Abelian group and let $\mathbf{a}, \mathbf{b}$ be periodic sequences of elements from G, with the same period T. The sequences $\mathbf{a}$ and $\mathbf{b}$ are said to be *shift distinct* if $\mathbf{a}$ differs from every left shift of $\mathbf{b}$. If $d \geq 1$ is an integer, the d-fold *decimation* of $\mathbf{a} = (a_0, a_1, \cdots)$ is the sequence $\mathbf{b} = (b_0, b_1, \cdots)$ where $b_i = a_{di}$, which consists of every d-th element from $\mathbf{a}$. In this case we write $\mathbf{b} = \mathbf{a}[d]$.

Proposition 13.2.1. *Let F be a finite field with $|F|$ elements, and let $\mathbf{a}$ be an m-sequence of degree r with values in F. Then the following statements hold.*

1. *Every m-sequence of degree r is a left shift of a decimation of $\mathbf{a}$.*

2. *The decimation* $\mathbf{a}[d]$ *is again an m-sequence if and only if* d *is relatively prime to* $|F|^r - 1$.
3. *The decimation* $\mathbf{a}[d]$ *is a left shift of* $\mathbf{a}$ *if and only if* d *is a power of* $|F|$.
4. *Hence there are* $\phi(|F|^r - 1)/r$ *shift distinct m-sequences of degree* r *with values in* F*, and they are in one to one correspondence with the set of degree* r *monic primitive polynomials over* F.

Proof. Let E be the field with $|F|^r$ elements and let $L : E \to F$ be a surjective F-linear homomorphism, for example the trace Tr_F^E. Then there exists a primitive element $\alpha \in E$ and a nonzero $A \in E$ such that $a_i = L(A\alpha^i)$ (for all i). Let $\mathbf{b} = (b_0, b_1, \cdots)$ be a second m-sequence of degree r, so $b_j = L(B\beta^j)$ for some primitive element $\beta \in E$ and some nonzero element $B \in E$. There is a unique integer d $(1 \le d \le |L| - 1)$ so that $\beta = \alpha^d$ and there is a unique integer t $(0 \le t \le |L| - 1)$ so that $B = A\beta^t$. Then the t-shift of the d-fold decimation of the sequence $\mathbf{a}$ is the sequence $\mathbf{c} = (c_0, c_1, \cdots)$ where

$$c_k = L(A\alpha^{d(k+t)}) = L(A\beta^k\beta^t) = L(B\beta^k) = b_k.$$

This proves (1). The element $\beta = \alpha^d$ is primitive if and only if d is relatively prime to $|L| - 1$ which proves (2). Now let $q(x) = q_0 + q_1x + \cdots + q_rx^r$ be the minimal polynomial of α, with $q_i \in F$. Then $q(x)$ is also the connection polynomial of a LFSR that generates the sequence $\mathbf{a}$. If $s = |F|$ then $q_i^s = q_i$ so $(q(\alpha))^s = q(\alpha^s) = 0$. It follows that the roots of q are the elements $\alpha, \alpha^s, \alpha^{s^2}, \cdots, \alpha^{s^{r-1}}$. If β is any one of these, then α and β have the same minimal polynomial. Hence the the m-sequence generated by α coincides (up to a shift) with the m-sequence generated by β because they are LFSR sequences with the same connection polynomial. On the other hand, if β is not one of these, then α and β have different minimal polynomials, so the corresponding m-sequences correspond to different LFSRs so they are different. This proves (3), and (4) follows. □

13.3 Interleaved structure

Let $F = \mathbb{F}_q \subseteq E = \mathbb{F}_{q^m} \subset K = \mathbb{F}_{q^n}$ be finite fields, let $\alpha \in K$ be a primitive element and let $\omega \in K$. Let $\mathbf{a} = (a_0, a_1, \cdots)$ be the resulting m-sequence with $a_i = \mathsf{Tr}_F^K(\omega\alpha^i)$. Consider the decimation $\mathbf{b} = \mathbf{a}[d]$ with $d = (|K| - 1)/(|E| - 1)$, that is, $b_i = \mathsf{Tr}(\omega\alpha^{di})$. Note that $\beta = \alpha^d$ is a primitive element of E (since $\beta^{|E|-1} = 1$). We obtain a (shorter) m-sequence $\mathbf{c}$ with $c_i = \mathsf{Tr}_F^E(\beta^i)$. Let $z = (q^{n-m} - 1)/(q - 1)$.

Proposition 13.3.1. *IF* $d = (|K| - 1)/(|E| - 1)$*, then depending on* ω*, the decimation* $\mathbf{b} = \mathbf{a}[d]$ *is either a shift of the m-sequence* $\mathbf{c}$ *(for* $d - z$ *values of* ω*) or the zero sequence of length* $|E| - 1$ *(for* $d - z$ *values of* ω*). Therefore the m-sequence* $\mathbf{a}$ *may be realized by interleaving* $d - z$ *different periods of the m-sequence* $\mathbf{c}$ *together with* z *copies of the zero sequence of length* $|E| - 1$.

Proof. Think of writing a single period of the m-sequence $\mathbf{a}$ in a series of rows, each of length d, to obtain an array $A_{st} = \mathsf{Tr}_F^L(\omega\alpha^{sd+t})$ with $0 \le t \le d-1$ and $0 \le s \le |E|-2$. Then

$$A_{st} = \mathsf{Tr}_F^E\mathsf{Tr}_E^L(\omega\alpha^t\beta^s) = \mathsf{Tr}_F^E(h_t\beta^s)$$

where $h_t = \mathsf{Tr}_E^L(\omega\alpha^t)$. The sequence $\mathbf{b}$ is obtained by reading down the columns, one after another. Reading down the t-th column, if $h_t \neq 0$ then we find the m-sequence $\mathsf{Tr}_F^E(h_t\beta^s)$ as s varies from 0 to $|E|-2$, which is a shift of the m-sequence $\mathbf{c}$. If $h_t = 0$ then we find a column consisting entirely of $|E|-1$ zeroes. The number of such columns is calculated from the fact that if $h_t \neq 0$ then the t-th column contains $q^{m-1}-1$ zeroes, and there are $q^{n-1}-1$ zeroes in all. □

If $E = F$ then $A_{st} = \beta^s h_t$ so in this case each row is a constant multiple, β^s of the first row, and each column is either identically zero, or it is a cyclic permutation of the sequence $1, \beta, \beta^2, \cdots, \beta^{r-2}$.

13.4 Fourier transforms and m-sequences

Let p be a prime number and $F = \mathbb{Z}/(p)$ be the field with p elements. Let E be the field with p^r elements. Fix a surjective F-linear mapping $L : E \to F$. Fix a nontrivial additive character $\varphi : F \to \mathbb{C}^\times$, for example, $\varphi(a) = \zeta^a$ where $\zeta = e^{2\pi i/p}$. Then $\psi = \varphi \circ L : E \to \mathbb{C}^\times$ is a nontrivial additive character of E. This data allows us to enumerate all the $\mathbb{C}$-valued additive characters of E (including the trivial character 1) by setting

$$\chi_y(x) = \varphi(L(yx))$$

for any $y \in E$. (See Section 2.3.a and Section 3.2.g.) If $B : E \to F$ is any function then the composition $\varphi \circ B : E \to \mathbb{C}$ has a Fourier transform (with respect to the additive group structure on E),

$$\widehat{\varphi B}(\chi_y) = \sum_{x\in E} \varphi(B(x))\overline{\chi}_y(x) = \sum_{x\in E} \varphi(B(x) - L(yx)).$$

(When $F = \mathbb{F}_2$ and $\varphi(y) = (-1)^y$ it is customary to drop the mention of φ and to write this equation as $\widehat{B}(y) = (-1)^{B(x)+L(xy)}$.) Now let $\alpha \in E$ be a primitive element and let $\mathbf{a} = (L(\alpha^0), L(\alpha^1), \cdots)$ be the resulting m-sequence.

Proposition 13.4.1. *Let $\mathbf{b} = (b_0, b_1, \cdots)$ be a periodic sequence, with period $|E|-1$, of elements in F. Define $B : E \to F$ by $B(\alpha^i) = b_i$ and $B(0) = 0$. Then the cross-correlation $\mathcal{C}_{\mathbf{b},\mathbf{a}}(t)$ (with respect to the character φ) is given by the (complex) Fourier transform of the function $\varphi \circ B : E \to \mathbb{C}$. More precisely, if $t \ge 0$ then*

$$\widehat{\varphi B}(\chi_{\alpha^t}) = \mathcal{C}_{\mathbf{b},\mathbf{a}}(t) + 1.$$

Proof. Compute the Fourier transform (as in Section 2.3.b),

$$\begin{aligned}\widehat{\varphi B}(\chi_{\alpha^t}) &= \sum_{x\in E}\varphi B(x)\overline{\chi}_{\alpha^t}(x) = \sum_{x\in E}\varphi(B(x))\overline{\varphi}L(\alpha^t x))\\ &= \sum_{i=0}^{p^r-2}\varphi(B(\alpha^i))\overline{\varphi}L(\alpha^{t+i}) + \varphi(B(0))\overline{\varphi}(L(0))\\ &= \sum_{i=0}^{p^r-2}\varphi(b_i)\overline{\varphi}(a_{t+i}) + \varphi(B(0)) = \mathcal{C}_{\mathbf{b},\mathbf{a}}(t) + \varphi(B(0)) \quad \square\end{aligned}$$

If $\mathbf{b} = (b_0, b_1, \cdots)$ is a periodic sequence of elements in some field F (with period T), and if $\varphi : F \to \mathbb{C}$ is a nontrivial character, then recall from Definition 11.2.1 that the *imbalance* of $\mathbf{b}$ with respect to φ is the sum

$$Z(\mathbf{b}) = \sum_{i=0}^{T-1}\varphi(b_i).$$

In the case $F = \mathbb{Z}/(2)$ the imbalance is the number of zeroes minus the number of ones in a single period of $\mathbf{b}$.

Corollary 13.4.2. *Let* $\mathbf{a}$ *be an m-sequence of period* $p^r - 1$*, with entries in the field* $F = \mathbb{Z}/(p)$*. Let* $\mathbf{b}$ *be any periodic sequence with the same period, and let* $\varphi : F \to \mathbb{C}^\times$ *be a nontrivial character. Then the cross-correlation* $\mathcal{C}_{\mathbf{b},\mathbf{a}}(t)$ *with respect to the character* φ *satisfies*

$$\sum_{t=0}^{p^r-2}|\mathcal{C}_{\mathbf{b},\mathbf{a}} + 1|^2 = p^{2r} - |Z(\mathbf{b}) + 1|^2.$$

Proof. We may assume $a_i = L(\alpha^i)$ where α is a primitive element in the field $E = \mathbb{F}_{p^r}$ and where $L : E \to F$ is a surjective F-linear mapping such as the trace. Define $B : E \to F$ by $B(\alpha^i) = b_i$ and $B(0) = 0$. By Parseval's formula (equation (2.15)),

$$\sum_{\chi}|\widehat{\varphi B}(\chi)|^2 = |E|\sum_{x\in E}|\varphi(B(x))|^2 = p^r \cdot p^r.$$

The sum on the left is over all additive characters χ of E, including the trivial character $\chi_0 = 1$. By Proposition 13.4.1 this sum is

$$\sum_{t=0}^{p^r-2}|\mathcal{C}_{\mathbf{b},\mathbf{a}} + 1|^2 + |\widehat{\varphi B}(\chi_0)|^2.$$

But $\widehat{\varphi B}(\chi_0) = \sum_{x\in E}\varphi(B(x))\overline{\chi_0}(x) = \sum_{i=0}^{p^r-2}\varphi(b_i) + 1 = Z(\mathbf{b}) + 1$. $\square$

The "simplest" function $B : E \to F$ is an F-linear function. For such a function, the sequence $\mathbf{b}$ is a shift of the m-sequence $\mathbf{a}$ by some amount, say t_0. Then the Fourier coefficient $\widehat{\varphi B}(\chi_t)$ is 0 for $t \neq t_0$ and it is p^r if $t = t_0$. Thus, the Fourier coefficients "detect" the function B. Because of the existence of the fast Fourier transform, such a linear function B should be considered insecure. A function $B : E \to F$ might be considered to be invisible from the point of view of Fourier analysis if the magnitude $|\widehat{\varphi B}(\chi)|$ of the Fourier coefficients of φB are all equal, in which case (by Corollary 13.4.2) they must equal $\sqrt{|E|}$. This leads to the following definition ([166], [115]).

Definition 13.4.3. *Fix a nontrivial character $\varphi : F \to \mathbb{C}^\times$. A function $B : E \to F$ is* bent *(with respect to φ) if the magnitude $|\widehat{\varphi B}(\chi_y)|$ of every Fourier coefficient of $\varphi \circ B$ is equal to $\sqrt{|E|}$. The sequence $\mathbf{b} = (B(\alpha^0), B(\alpha^1), \cdots)$ is a* bent sequence *if B is a bent function and $\alpha \in E$ is a primitive element.*

The theory of bent functions and related concepts is a subject of considerable research that is largely outside the scope of this book. We just mention one result that shows the notion of bentness depends only on the function B, not on the character φ.

Proposition 13.4.4. *([115]) If $B : E \to F$ is bent with respect to one nontrivial character $\varphi : F \to \mathbb{C}^\times$ then it is bent with respect to every nontrivial character $\psi : F \to \mathbb{C}^\times$.*

Proof. Since $F = \mathbb{F}_p \cong \mathbb{Z}/(p)$ is a prime field, $\psi = \varphi^s$ for some integer s. Then

$$|\widehat{\varphi B}(\chi_y)|^2 = \left| \sum_{x \in E} \varphi(B(x) - L(yx)) \right|^2 = \left| \sum_{x \in E} \zeta^{B(x) - L(yx)} \right|^2 = |E|$$

and

$$|\widehat{\psi B}(\chi_y)|^2 = \left| \sum_{x \in E} \psi(B(x) - L(yx)) \right|^2 = \left| \sum_{x \in E} \zeta^{s(B(x) - L(yx))} \right|^2.$$

The equality between these is an immediate consequence of Lemma 3.2.12. □

Proposition 13.4.1 says that for any nontrivial character $\varphi : F \to \mathbb{C}^\times$, and for any choice of primitive element $\alpha \in E$ the cross-correlation (with respect to φ) between the bent sequence $\mathbf{b}$ and the m-sequence $\mathbf{a} = (L(\alpha^0), L(\alpha^2), \cdots)$ satisfies $\mathcal{C}_{\mathbf{b},\mathbf{a}}(t) = \sqrt{|E|} - \varphi(B(0))$ for all t.

Bent functions have a number of interesting combinatorial properties. A bent function may be considered as a difference set in an elementary Abelian 2-group [144, 38, 154]. Correlation properties of bent functions have been studied [115, 152]. In Theorem 18.5.3 we present bounds on the linear span of bent functions which were derived in [114].

13.5 Cross-correlation of an m-sequence and its decimation

In this section we consider the cross-correlation between an m-sequences and its decimations. This is a very difficult subject which has been studied for many years by a variety of techniques, and about which little is known. Often, other questions concerning pseudorandom sequences can be reduced to understanding the cross-correlation of m-sequences.

Throughout this section we fix a finite field $F = \mathbb{F}_q$ a nontrivial character $\chi : F \to \mathbb{C}^\times$ and an m-sequence $\mathbf{a} = (a_0, a_1, \cdots)$of rank r over F. Let $E = \mathbb{F}_{q^n}$ be the unique field extension of F of degree n. Let $\mathsf{Tr} : E \to F$ denote the trace Tr_F^E. According to Proposition 13.1.2, by possibly replacing $\mathbf{a}$ with a shift of itself, we may assume there is a primitive element $\alpha \in E$ such that $a_i = \mathsf{Tr}(\alpha^i)$. Let $\mathbf{b} = (b_0, b_1, \cdots)$ be a second sequence that is a d-fold decimation of $\mathbf{a}$ for some d. That is, $b_i = \mathsf{Tr}_F^E(\alpha^{di})$. We want to find the cross-correlations of $\mathbf{a}$ and $\mathbf{b}$. According to Proposition 13.2.1 the sequence $\mathbf{b}$ is a m-sequence if and only if d is relatively prime to $|E| - 1$. Conversely, every m-sequence is a shift of such a decimation of $\mathbf{a}$.

13.5.a Two basic computations

The following proposition gives two interpretations of the cross-correlation of the sequences $\mathbf{a}, \mathbf{b}$, one as a character sum (about which we say more in Section 13.5.d), and the other in terms of counting the number of points in the intersection of two algebraic varieties.

Proposition 13.5.1. *Let $\mathbf{a}$ be an m-sequence and let $\mathbf{b}$ be its d-fold decimation. The cross-correlation of the sequences $\mathbf{a}, \mathbf{b}$ with respect to the character χ is given by*

$$\mathcal{C}_{\mathbf{b},\mathbf{a}}(t) = \sum_{x \in E} \chi(\mathsf{Tr}(x^d - Ax)) - 1 \tag{13.2}$$

$$= \sum_{u \in F} \sum_{v \in F} |Q_u \cap A^{-1} H_v| \chi(u) \overline{\chi}(v) - 1 \tag{13.3}$$

where $A = \alpha^t$, and where $|Q_u \cap H_v|$ denotes the number of elements of E in the intersection of the hypersurface

$$Q_u = \left\{ x \in E |\ \mathsf{Tr}(x^d) = u \right\}$$

and the hyperplane

$$H_v = \left\{ x \in E :\ \mathsf{Tr}(x) = v \right\}.$$

Proof. Let $\beta = \alpha^d$. Then the cross-correlation with respect to the character χ is

$$\mathcal{C}_{\mathbf{b},\mathbf{a}}(t) = \sum_{i=0}^{q^n - 2} \chi(b_i) \overline{\chi}(a_{i+t}) \tag{13.4}$$

$$= \sum_{i=0}^{q^n-2} \chi(\mathsf{Tr}(\beta^i))\overline{\chi}(\mathsf{Tr}(\alpha^{i+t})) \tag{13.5}$$

$$= \sum_{0\neq x\in E} \chi(\mathsf{Tr}(x^d))\overline{\chi}(\mathsf{Tr}(Ax)) \tag{13.6}$$

$$= \sum_{x\in E} \chi(\mathsf{Tr}(x^d - Ax)) - 1. \tag{13.7}$$

which proves equation (13.2). Equation (13.3) is immediate from the definitions. □

We need to use the following simple lemma. Recall that two hyperplanes in a vector space are *parallel* if one is a translate of the other.

Lemma 13.5.2. *Let $T : E \to F$ be an F-linear surjective mapping (such as the trace). For $u \in F$ set $H_u = \{x \in E : \ T(x) = u\}$. Let $b \in E$ and $u, v \in F$. Then bH_u is parallel to H_v if and only if $b \in F$.*

Proof. We remark that H_a is parallel to H_b for every $a, b \in F$, and that $x \in bH_u$ if and only if $T(b^{-1}x) = u$.

If $b \in F$, then $bH_u = H_{bu}$ which is parallel to H_u, proving the "if" part.

Now assume that $b \in E$ and bH_u is parallel to H_v and hence also to H_0. Then for any fixed $x_0 \in bH_u$ we have $bH_u = H_0 + x_0$. Now

$$\begin{aligned} H_0 &= bH_u - x_0 \\ &= \{bx - x_0 : \ T(x) = u\} \\ &= \{y \in E : \ T(b^{-1}y + b^{-1}x_0) = u\} \\ &= \{y \in E : \ T(b^{-1}y) = 0\} \\ &= bH_0. \end{aligned}$$

Choose any $z \in E - H_0$. Since H_0 is a codimension one subspace of E the addition of this one more linearly independent element will span E as a vector space over F. Therefore we may write $bz = az + h$ for some $a \in F$ and some $h \in H_0$, so $z = h/(b-a)$ unless $b = a$. But multiplication by b preserves H_0, as does multiplication by a, hence so does multiplication by $(b-a)$ and therefore also multiplication by $(b-a)^{-1}$. This implies $z \in H_0$ which is a contradiction. Thus we conclude that $b = a \in F$. □

13.5.b Linearly decimated sequences

In this section we assume that d is a power of $q = |F|$. This d is relatively prime to $q^n - 1$, so $\mathbf{b}$ is an m-sequence. Now the function $\mathsf{Tr}(x^d)$ is actually a linear function, so we refer to the sequence

b as a *linear decimation* of the sequence **a**. In fact, $\mathsf{Tr}(x^d) = \mathsf{Tr}(x)$ so the cross-correlation with respect to the character χ is

$$\mathcal{C}_{\mathbf{b},\mathbf{a}}(t) = \sum_{u\in F}\sum_{v\in F} |H_u \cap A^{-1}H_v|\chi(u)\overline{\chi}(v) - 1$$

where $A \in E$ are determined by the shift t.

Proposition 13.5.3. *Let* **a** *be an m-sequence and let* **b** *be its d-fold decimation where d is a power of q. Then d is relatively prime to* $q^n - 1$ *so* **b** *is also an m-sequence, and the cross-correlation of the sequences* $\mathbf{b}, \mathbf{a}$ *is*

$$\mathcal{C}_{\mathbf{b},\mathbf{a}}(t) = \begin{cases} -1 & \text{if } t \neq 0 \\ q^n - 1 & \text{if } t = 0 \end{cases}$$

Proof. First consider the case when $A \notin F$. By Lemma 13.5.2 this implies that H_u and $A^{-1}H_v$ are not parallel, so their intersection contains q^{n-2} elements and in particular it is independent of u and v. But $\sum_{u\in F}\chi(u) = 0$ so we conclude that $\mathcal{C}_{\mathbf{b},\mathbf{a}}(t) = -1$ in this case. Now suppose $A \in F$. Then H_u and $A^{-1}H_v$ are parallel: they coincide if $u = A^{-1}v$, otherwise their intersection is empty. That is, they coincide if and only if $v = Au$. Thus the cross-correlation is

$$\begin{aligned} \mathcal{C}_{\mathbf{b},\mathbf{a}} &= q^{n-1}\sum_{u\in F}\chi(u)\chi^{-1}(Au) - 1 \\ &= q^{n-1}\sum_{u\in F}\chi((1-A)u) - 1 \\ &= \begin{cases} -1 & \text{if } A = 1 \\ q^n - 1 & \text{otherwise.} \end{cases} \end{aligned}$$

□

In fact, this answer (see also Proposition 12.1.3) is just the autocorrelation function of **a** (since **b** is a shift of **a**), but we have computed it using a technique that works in many other situations, as we see below.

13.5.c Quadratically decimated sequences

If $d = q^i + q^j$ is a sum of two powers of q, then the function $\mathsf{Tr}(x^d)$ is a quadratic form, and the set Q_u is a *quadric hypersurface*. By Section 3.3 the problem of computing the cross-correlation when $d = q^i + q^j$ is equivalent to that of computing the cross-correlation for $d = 1 + q^i$. First we need the following technical result whose proof will appear at the end of this section.

Theorem 13.5.4. *Let $Q : E = \mathbb{F}_{q^n} \to F = \mathbb{F}_q$ be a quadratic form of rank m, and Type I, II, or III (see Theorem 3.3.1). Let $L : E \to F$ be a linear function, let $\chi : F \to \mathbb{C}^\times$ be a nontrivial character, and for any $w \in F$ set*

$$N_w = |\{x \in E : \ Q(x) + L(x) = w\}| \tag{13.8}$$

$$S(Q, L, \chi) = \sum_{w \in F} N_w \chi(w). \tag{13.9}$$

Then the value of $S(Q, L, \chi)$ and the number of times that value occurs, as L varies over all linear maps $E \to F$ is given in Table 13.1.

q **even**	$S(Q, L, \chi)$	Number of occurrences
Type I	0	$q^n - q^m$
	$-q^{n-m/2}$	$\frac{1}{2}(q^m - q^{m/2})$
	$q^{n-m/2}$	$\frac{1}{2}(q^m + q^{m/2}) - 1$
Type II	0	$q^n - q^{m-1} - 1$
	$-q^{n-(m-1)/2}$	$\frac{1}{2}\left(q^{m-1} - q^{(m-1)/2}\right)$
	$q^{n-(m-1)/2}$	$\frac{1}{2}\left(q^{m-1} + q^{(m-1)/2}\right)$
Type III	0	$q^n - q^m$
	$-q^{n-m/2}$	$\frac{1}{2}(q^m + q^{m/2}) - 1$
	$q^{n-m/2}$	$\frac{1}{2}(q^m - q^{m/2})$
q **odd**	$\lvert S(Q, L, \chi)\rvert$	Number of occurrences
	0	$q^n - q^m$
	$q^{n-m/2}$	q^m

Table 13.1: Cross-correlation of quadratically decimated sequences

Theorem 13.5.5. *Let $\alpha \in E$ be a primitive element and let $\mathbf{a}$ be the m-sequence $a_k = \mathsf{Tr}_F^E(\alpha^k)$. Let $d = 1 + q^i$ and let $\mathbf{b}$ be the d-fold decimation of $\mathbf{a}$. For each shift t, $(0 \le t \le |E| - 2)$ let $L : E \to F$ be the linear mapping $L(x) = \mathsf{Tr}_F^E(\alpha^t x)$. Let $\chi : F \to \mathbb{C}^\times$ be a character. Let $m = \mathrm{rank}(Q)$ where Q is the quadratic form $Q(x) = \mathsf{Tr}_F^E(x^d)$. Then the rank m and the type (I, II, or III) of Q is given in Theorem 3.3.5 and the cross-correlation of the sequences $\mathbf{b}, \mathbf{a}$ is the number*

$$\mathcal{C}_{\mathbf{b},\mathbf{a}}(t) = S(Q, L, \chi) - 1$$

as determined in Table 13.1.

The proof of Theorems 13.5.4 and 13.5.5 will occupy the rest of this section. To compute the cross-correlation $\mathcal{C}_{\mathbf{b},\mathbf{a}}(t)$ we compute the sum (13.3) for all possible values of A. As t varies $(0 \le t \le |E| - 2)$ the element $A = \alpha^t \in E$ varies over all nonzero elements so $L_A(x) = -\mathsf{Tr}_F^E(Ax)$ varies over all nonzero F-linear maps $L : E \to F$. (See Theorem 3.2.14; the minus sign has been inserted for convenience in later calculations.) For each nonzero value of $A \in E$ we have $A^{-1}H_v = \{x \in E : \ L_A(x) = -v\}$. Rather than compute $\mathcal{C}_{\mathbf{b},\mathbf{a}}(t)$ for each t separately we determine the possible values of $\mathcal{C}_{\mathbf{b},\mathbf{a}}(t)$ and the number of values of t for which each value occurs. Thus it suffices to compute the value of

$$\begin{aligned}
\mathcal{C}_{\mathbf{b},\mathbf{a}}(L) + 1 &= \sum_{u \in F} \sum_{v \in F} |\{x \in E : \ Q(x) = u \text{ and } L(x) = -v\}| \, \chi(u - v) \\
&= \sum_{w \in F} |\{x \in E : \ Q(x) + L(x) = w\}| \, \chi(w) \\
&= \sum_{w \in F} N_w \chi(w) = S(Q, L, \chi)
\end{aligned}$$

and the number of times each value occurs, as $L : E \to F$ is allowed to vary over all nonzero F-linear mappings. Hence, Theorem 13.5.5 reduces to Theorem 13.5.4, which we now prove.

The numbers N_w were determined in Theorem 3.3.4. Fix a basis $\{x_1, x_2, \cdots, x_n\}$ of E over F and let $L(x) = c_1x_1 + \cdots c_nx_n = c \cdot x$ where $c_i \in F$. Choose a maximal subspace $V \subset E$ on which Q is non-degenerate. As vector spaces over F we have an isomorphism, $E \cong V \oplus \mathrm{Ker(Q)}$. Let Q_1, L_1 denote the restrictions of Q, L to V and let L_2 be the restriction of L to Ker(Q). Then $L(a, b) = L_1(a) + L_2(b)$ for $a \in V$ and $b \in \mathrm{Ker(Q)}$. According to Proposition 3.3.3 if $\mathrm{Ker(Q)} \not\subset \mathrm{Ker(L)}$, then for any $w \in F$, $N_w = q^{n-1}$ so $S = q^{n-1} \sum_{w \in F} \chi(w) = 0$. Such an S occurs for $q^n - q^m$ different (nonzero) values of L for the following reason. The condition $\mathrm{Ker(Q)} \subset \mathrm{Ker(L)}$ means that $L_2 = 0$ and L_1 is arbitrary so there are $q^m - 1$ possible nonzero choices for L satisfying $\mathrm{Ker(Q)} \subset \mathrm{Ker(L)}$.

It remains to evaluate S when $\mathrm{Ker(Q)} \subset \mathrm{Ker(L)}$. In this case (by Proposition 3.3.3), $N_w = q^{n-m}N'_w$ where N'_w is the number of solutions $x \in V$ to the equation $Q_1(x) + L_1(x) = w$. By a linear change of variables, we may assume that $V = F^m$ and that $Q_1(x)$ is a quadratic form of rank m on V. Then

$$S(Q, L, \chi) = \begin{cases} 0 & \text{for} \quad q^n - q^m \quad \text{choices of } L \\ q^{n-m} S_1(Q_1, L_1, \chi) & \text{for} \quad q^m - 1 \quad \text{choices of } L \end{cases}$$

where $S_1(Q_1, L_1, \chi) = \sum_{w \in F} N'_w \chi(w)$. For each value of $S_1(Q_1, L_1, \chi)$ we need to calculate the number of choices of $L_1 : V \to F$ for which this value occurs (recalling that $L_2 = 0$). The calculation of $S_1(Q_1, L_1, \chi)$ and the number of times it occurs is rather messy.

Case that q is even. In this case every additive character $\chi : F \to \mathbb{C}^\times$ takes values in $\{\pm 1\}$ and so $\chi^{-1} = \chi$. Let $Q_1, L_1 : F^m \to F$ with $L_1(x) = c_1x_1 + \cdots + c_mx_m = c \cdot x$. First assume Q_1 is a Type I form of rank m on F^m. Then m is even, and for all $w \in F$, $N_w = q^{m-1} + \nu(w + Q_1(c))q^{m/2-1}$ from Theorem 3.3.4. Therefore

$$\begin{aligned} S_1(Q_1, L_1, \chi) &= \sum_{w \in F} q^{m-1}\chi(w) + q^{m/2-1} \sum_{w \in F} \nu(w + Q_1(c))\chi(w) \\ &= q^{m/2-1} \sum_{v \in F} \nu(v)\chi(v)\chi^{-1}(Q_1(c)) \\ &= q^{m/2-1}\chi^{-1}(Q_1(c))((q-1)\chi(0) - \sum_{v \neq 0} \chi(v)) \\ &= q^{m/2-1}\chi^{-1}(Q_1(c))((q-1)\chi(0) + \chi(0)) \\ &= q^{m/2}\chi^{-1}(Q_1(c)). \end{aligned}$$

This occurs $q^m - 1$ times, representing the different possible values of c. The character value $\chi^{-1}(h) \in \{\pm 1\}$ is -1 for $q/2$ values of $h \neq 0$; it is $+1$ for $q/2 - 1$ values of $h \neq 0$, and it is $+1$ when $h = 0$. Counting the number of c for which $Q_1(c) = h$ we conclude that $\chi^{-1}(Q_1(c)) = -1$ for $(q^m - q^{m/2})/2$ values of c, otherwise it is $+1$. In summary,

$$S_1(Q_1, L_1, \chi) = \begin{cases} q^{m/2} & \text{occurs for } \frac{1}{2}(q^m + q^{m/2}) - 1 \text{ values of } 0 \neq c \in V \\ -q^{m/2} & \text{occurs for } \frac{1}{2}(q^m - q^{m/2}) \text{ values of } 0 \neq c \in V. \end{cases}$$

Consequently,

$$S(Q, L, \chi) = \begin{cases} 0 & \text{for } q^n - q^m \text{ values of } L \\ q^{n-m/2} & \text{for } \frac{1}{2}(q^m + q^{m/2}) - 1 \text{ values of } L \\ -q^{n-m/2} & \text{for } \frac{1}{2}(q^m - q^{m/2}) \text{ values of } L. \end{cases}$$

The same calculation for Type III gives

$$S(Q, L, \chi) = \begin{cases} 0 & \text{for } q^n - q^m \text{ values of } L \\ -q^{n-m/2} & \text{for } \frac{1}{2}(q^m + q^{m/2}) - 1 \text{ values of } L \\ q^{n-m/2} & \text{for } \frac{1}{2}(q^m - q^{m/2}) \text{ values of } L. \end{cases}$$

Now assume that $Q_1 : F^m \to F$ is a Type II quadratic form of rank m. Then m is odd. Let $\psi : F \to \mathbb{C}^\times$ be the trace character,

$$\psi(w) = (-1)^{\mathsf{Tr}_{\mathbb{F}_2}^F(w)}.$$

Let $L_1(x) = c_1x_1 + \cdots + c_mx_m$. The number N_w of solutions to the equation $Q_1(x) + L_1(x) = w$ is

$$N'_w = q^{m-1} + \tau(c, w)q^{(m-1)/2},$$

where $\tau(c, w)$ is given by equation (3.12). That is,

$$\tau(c, w) = \begin{cases} 0 & \text{if } c_m = 0 \\ \psi(B_{m-1}(c)/c_m^2)\psi(w/c_m^2) & \text{otherwise.} \end{cases}$$

We wish to compute

$$S_1(Q_1, L_1, \chi) = q^{m-1} \sum_{w \in F} \chi(w) + q^{(m-1)/2} \sum_{w \in F} \tau(c, w)\chi(w).$$

The first term in the above sum vanishes. If $c_m = 0$ (which holds for $q^{m-1} - 1$ possible nonzero values of $(c_1, \cdots, c_{m-1})$) then $\tau(c, w) = 0$ so $S_1 = 0$. As c_m varies among the remaining $q - 1$ nonzero possibilities the character $\psi(w/c_m^2)$ varies over all nonzero characters of F, one of which is χ. So there are $q - 2$ values of c_m such that $\psi(w/c_m^2)\chi(w)$ is nontrivial, which again gives $S_1(Q_1, L_1, \chi) = 0$. This accounts for $(q-2)q^{m-1}$ values of c. Thus, $S_1(Q_1, L_1, \chi) = 0$ for a total of $(q-1)(q^{m-1})-1$ nonzero values of c. If c_m equals the one remaining value where $\psi(w/c_m^2) \equiv \chi(w)$ then

$$S_1(Q_1, L_1, \chi) = q^{(m-1)/2}\psi(B_{m-1}(c)/c_m^2) \sum_{w \in F} 1 = \pm q^{(m+1)/2}.$$

To determine the number of times each sign occurs, consider the equation $B_{m-1}(c)/c_m^2 = u \in F$. Then $\psi(u) = +1$ for $q/2$ values of u (including $u = 0$) and $\psi(u) = -1$ for $q/2$ nonzero values of u. Since B_{m-1} is a Type I quadratic form, by Theorem 3.3.2 there will therefore be $(q/2)(q^{m-2} - q^{(m-1)/2-1})$ choices of $(c_1, \cdots, c_{m-1})$ which give $\psi(B_{m-1}(c)/c_m^2) = -1$. Similarly there will be $q^{m-2} + (q-1)q^{(m-1)/2-1} + (q/2-1)(q^{m-2} - q^{(m-1)/2-1}) = (q/2)(q^{m-2} + q^{(m-1)/2-1})$ choices that give the sign $+1$, counting $u = 0$ and $u \neq 0$ separately. In summary,

$$S_1(Q_1, L_1, \chi) = \begin{cases} 0 & \text{for } q^m - q^{m-1} - 1 & \text{choices of } c \\ -q^{(m+1)/2} & \text{for } \frac{1}{2}\left(q^{m-1} - q^{(m-1)/2}\right) & \text{choices of } c \\ q^{(m+1)/2} & \text{for } \frac{1}{2}\left(q^{m-1} + q^{(m-1)/2}\right) & \text{choices of } c. \end{cases}$$

Therefore

$$S(Q, L, \chi) = \begin{cases} 0 & \text{for } q^n - q^{m-1} - 1 & \text{choices of } c \\ -q^{n-(m-1)/2} & \text{for } \frac{1}{2}\left(q^{m-1} - q^{(m-1)/2}\right) & \text{choices of } c \\ q^{n-(m-1)/2} & \text{for } \frac{1}{2}\left(q^{m-1} + q^{(m-1)/2}\right) & \text{choices of } c. \end{cases}$$

Case that q is odd. As above, assume $Q_1 : F^m \to F$ is a quadratic form of rank m, $L(x) = c_1x_1 + \cdots + c_mx_m$ is a linear mapping, and $\chi : F \to \mathbb{C}^\times$ is a nontrivial character. Then

$$S_1(Q_1, L_1, \chi) = \sum_{w \in F} N'_w \chi(w)$$

where N'_w is the number of solutions to the equation $Q(x) + L(x) = w$ which by Theorem 3.3.4 is

$$N'_w = \begin{cases} q^{m-1} + \eta(w+R)\eta(\Delta')q^{(m-1)/2} & \text{if } m \text{ is odd} \\ q^{m-1} + \nu(w+R)\eta(\Delta')q^{m/2-1} & \text{if } m \text{ is even} \end{cases}$$

where $R = R(Q, c)$. First suppose that m is odd. Setting $x = w + R$, using $\eta(0) = 0$ and $\sum_{x \in F} \chi(x) = 0$ gives

$$S_1(Q_1, L_1, \chi) = \eta(\Delta')q^{(m-1)/2}\chi(-R) \sum_{x \neq 0} \eta(x)\chi(x) + 0.$$

Using Theorem 3.2.17 for the Gauss sum, we conclude that $|S_1(Q_1, L_1, \chi)| = q^{(m-1)/2}q^{1/2} = q^{m/2}$. Next suppose that m is even. Using $\nu(0) = q - 1$ and $\nu(x) = -1$ for $x \neq 0$ gives

$$\begin{aligned} S_1(Q_1, L_1, \chi) &= \eta(\Delta')q^{m/2-1}\chi(-R) \left(\sum_{x \neq 0} \nu(x)\chi(x) + (q-1) \right) \\ &= \eta(\Delta')q^{m/2-1}\chi(-R) \left(-\sum_{x \in F} \chi(x) + \chi(0) + q - 1 \right) \end{aligned}$$

from which we conclude that $|S_1(Q_1, L_1, \chi)| = q^{m/2-1}q = q^{m/2}$. Since $S(Q, L, \chi) = q^{n-m}S_1(Q_1, L_1, \chi)$ we obtain

$$|S(Q, L, \chi)| = \begin{cases} 0 & \text{for } q^n - q^m \text{ values of } c \\ q^{n-m/2} & \text{for } q^m \text{ values of } c \end{cases}$$

whether m is even or odd. This completes the proof of Theorem 13.5.5. □

13.5.d Other decimations, especially $d = -1$

In this section we suppose the ground field $F = \mathbb{F}_2 = \mathbb{Z}/(2)$ is the field with two elements. In a remarkable paper, P. Lachaud and J. Wolfman made a tremendous advance in the computation of cross-correlations [118]. Their result is based on the following simple observation. In this setting E is degree n extension of the field $\mathbb{F}_2$, so that $|E| = 2^n$.

Consider again equation (13.2) for the cross-correlation. The unique nontrivial character χ on F is $\chi(u) = (-1)^u$. That is, $\chi(0) = 1$ and $\chi(1) = -1$. Consider the set

$$V = \left\{(x, y) \in E \times E : y^2 - y = x^d - Ax\right\}. \tag{13.10}$$

So V is the set of solutions to the following "magic" equation: $y^2 - y = x^d - Ax$, which turns out to be intimately related to the cross-correlation, as we now explain.

If we fix $x \in E$ and if $\mathsf{Tr}(x^d - Ax) = 0$, then by Theorem 3.2.15 there are two solutions $y \in E$ to the magic equation, that is, we obtain two points in V. If $\mathsf{Tr}(x^d - Ax) \neq 0$, then there are no solutions to the magic equation. So the number of points in V is $2N$, where N is the number of 1's that occur in the sum (13.2). Moreover the number of -1's that occur in the sum in equation (13.2) is $2^n - N$. We conclude that

$$\mathcal{C}_{\mathbf{b},\mathbf{a}}(t) = N - (2^n - N) - 1 = |V| - 2^n - 1$$

where $|V|$ is the number of points in the set V, or equivalently, the number of solutions to the magic equation. This converts the problem of determining (or estimating) the cross-correlation into a "geometric" problem: that of counting (or estimating) the number of points in V.

In some cases, this number can be explicitly computed. Lachaud and Wolfman considered the case $d = -1$, in which case equation (13.2) is known as a *Kloosterman sum*, and the set V in equation (13.10) is an *elliptic curve*. The number of points in such an elliptic curve is known, by deep number theoretic results of Honda and Tate [79],[181]. This gives explicit values for the cross-correlation: numbers that had been earlier found in some low dimensional cases by laborious computer experimentation.

13.6 The Diaconis mind-reader

The magician enters the room where his audience is seated in rows facing the stage. He has a deck of cards bound with an elastic band, and he tosses the deck to someone in the audience. "Take off the elastic band and cut the deck a few times," he says, "then pass it to the person on your left and have him cut the deck a few times." The cardinal rule in magic, never to let your audience handle the props, has already been broken. "Now take the card on top, pass the deck to the person on your left, have him take the card on top, and so on. When the deck gets to the end of the row, pass it forward to the next row and have each person in that row take a card, until they're all gone. You may look at your card but don't show it to the person beside you." There is some rustling while the card dealing progresses. "Now I need five volunteers, who won't be upset if I read their minds," he says. Lots of people raise their hands. The the magician picks five (consecutive) people, and asks them to come to the front of the room. "I want you to think about the card you're holding and I'll tell you what it is. Concentrate on the card." Silence. A long

pause. "Think harder!" Silence. "OK, this isn't working. Look at your card again, concentrate on it, hold it against your forehead, and concentrate on it." The magician's face contorts with the effort of trying to read the minds of these people, with no apparent success. "OK," he says again, "I'm getting some interference from the black and red cards. You'll have to work with me on this. Will the people holding red cards please just take one step forward, closer to me and then we'll try again." The red cards move one step forward. "Mmmmmm," says the magician, "Now I think I'm getting it. Let's see. You have the three of spades. You have the seven of clubs..." And sure enough, he gets them all correct!

This wonderful magic trick, based on m-sequences, was invented by the famous magician, mathematician, and statistician, Persi Diaconis. Here is how it works.

The trick can be performed with a "deck" of either 31 cards or 63 cards. In the latter case it is necessary to borrow 11 cards from a second, identical deck. We will describe the 31 card trick. The cards are pre-arranged according to a chosen m-sequence of period 31. The m-sequence is generated by a LFSR with 5 cells. Each "state" of the shift register gets translated into a single card, for example, by having the rightmost two bits determine the suit, and the leftmost three bits determine the face value. If a red card represents a 1 and a black card represents a 0, then choose the suits so that 01 and 11 are translated into red suits, while 00 and 10 are translated into black suits. The leftmost 3 bits give a face value, 0 to 7, perhaps with a value of 0 representing a Queen. When the five consecutive audience members stand at the front of the room and give away the suit colors by stepping forward, the magician sees the "state" of the shift register and he can calculate from this the value of the rightmost card. Then, he runs the shift register in his mind to generate the next bit, and this allows him to calculate the second card from the right, and so on. Cutting the deck creates a left shift of the m-sequence. It doesn't matter how many times it is done, the cards remain in the same cyclic order.

Here is a possible implementation of this scheme. For the suit bits (rightmost two bits) use: $00 = \clubsuit$, $01 = \heartsuit$, $10 = \spadesuit$, $11 = \diamondsuit$. So a row of five people with the second and fifth stepping forward corresponds to 01001 or

010 2	01 ♡

which would be assigned to the rightmost person.

Choose a primitive element in $\mathbb{F}_{32}$, for example, a solution to $x^5 = x^2 + 1$. Here is the resulting m-sequence and its translation into playing cards.

0	0	0	0	1	0	1	0	1	1	1	0	1	1	0	0	0	} continued
				Q	Q	A	2	5	2	5	3	7	6	5	3	6	
				♡	♠	♡	♠	♡	♢	♢	♠	♡	♢	♠	♣	♣	

$$\left\{\begin{array}{cccccccccccccccccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & & & & \\ 4 & Q & A & 3 & 7 & 7 & 7 & 6 & 4 & A & 3 & 6 & 5 & 2 & 4 & A & 2 & 4 \\ \heartsuit & \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & \spadesuit & \clubsuit & \heartsuit & \diamondsuit & \spadesuit & \heartsuit & \spadesuit & \clubsuit & \heartsuit & \spadesuit & \clubsuit & \clubsuit & \clubsuit \end{array}\right.$$

The 01001 sequence considered above appears at the end of the second row. Having decoded this as a 2♡ for the rightmost person, it is a simple matter to run the shift register one step (adding the rightmost 1 and the middle 0) to obtain the next bit, a 1. So the next card corresponds to 10100 or 5♣, and so on.

The disadvantage of this scheme is the requirement that the deck have $2^N - 1$ cards. However, the number 2 is a primitive root modulo 53, so it is possible to do the same trick with a deck of 52 cards, by replacing the (binary) LFSR with a (binary) FCSR having connection integer $q = 53$. In this case, $q + 1 = 110110_2$ so the magician needs to know the colors for 6 consecutive cards, and the shift register has 4 nonzero connections and an integer memory, making it a bit harder to calculate.

13.7 Exercises

1. Let $n = \prod_{i=1}^{k} p_i$ be the product of distinct odd primes p_i. Let $R = \mathbb{Z}/(n)$. Let $q(x) \in R[x]$ with $q(0) = -1$ and $\deg(q) \geq 1$. Find an upper bound on the linear complexity of an LFSR sequence with connection polynomial $q(x)$ over R in terms of the multiplicative orders of x modulo $q(x)$ over the various $\mathbb{Z}/(p_i)$. Under what circumstances is the upper bound reached?

2. What is the maximum cross-correlation between an m-sequence of period 255 and a proper nonlinear decimation?

Chapter 14 Related Sequences and their Correlations

In this chapter we briefly review some of the sequences that are related to m-sequences, many of which have found applications in communications. See also Section 18.5, where the linear span of sequences derived from m-sequences is discussed.

14.1 Welch bound

For applications to spread spectrum communications, one attempts to find a collection of shift distinct sequences with low pairwise cross-correlation values. For a given (maximal) cross-correlation, there are theoretical limitations on the number of sequences in such a collection, the simplest of which is the *Welch bound*.

Suppose we have a collection of n periodic sequences of elements in a finite field F, each with the same period T. We can expand this set to include all the shifts of these sequences. If T is the minimal period of each sequence and the sequences are pairwise shift distinct, then we obtain a set of $N = Tn$ vectors, commonly referred to as a *signal set*. Let us denote these vectors $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \cdots, \mathbf{a}^{(N)}$. Let $\chi : F \to \mathbb{C}^\times$ be a nontrivial character. Then the cross-correlation with respect to χ of the vectors $\mathbf{a}^{(u)}$ and $\mathbf{a}^{(v)}$ $(1 \le u, v \le N)$ is the number

$$\mathcal{C}_{uv} = \sum_{k=1}^{T} \chi(\mathbf{a}_k^{(u)})\overline{\chi}(\mathbf{a}_k^{(v)}).$$

Then $\mathcal{C}_{uu} = T$. A good signal set is one for which the cross-correlations $\mathcal{C}_{uv}$ are small, for $u \neq v$.

Theorem 14.1.1. *Let $\mathbf{a}^{(1)}, \cdots, \mathbf{a}^{(N)}$ be a collection of N vectors in F^T. Let $\chi : F \to \mathbb{C}^\times$ be a nontrivial character. Let $\mathcal{C}_{uv} = \mathcal{C}_{\mathbf{a}^{(u)},\mathbf{a}^{(v)}}$ and set $\mathcal{C}_{\max} = \max\{|\mathcal{C}_{uv}| : u \neq v\}$. Then*

$$\mathcal{C}_{\max}^2 \ge \frac{T(N-T)}{N-1}.$$

Proof. For notational convenience write $x_k^{(u)} = \chi(\mathbf{a}_k^{(u)})$ for $1 \le u \le N$ and $1 \le k \le T$, so that $\mathcal{C}_{uv} = \sum_{k=1}^{T} x_k^{(u)} \bar{x}_k^{(v)}$. Summing the squares of the $\mathcal{C}_{uv}$ gives

$$N(N-1)\mathcal{C}_{\max}^2 + NT^2 \;\; \ge \;\; \sum_{u=1}^{N}\sum_{v=1}^{N} |\mathcal{C}_{uv}|^2$$

$$
\begin{aligned}
&= \sum_{u=1}^{N}\sum_{v=1}^{N}\sum_{i=1}^{T}\sum_{j=1}^{T} x_i^{(u)}\bar{x}_i^{(v)}\bar{x}_j^{(u)}x_j^{(v)} \\
&= \sum_{i=1}^{T}\sum_{j=1}^{T}\left(\sum_{u=1}^{N} x_i^{(u)}\bar{x}_j^{(u)}\right)^2 \\
&\geq \sum_{i=1}^{T}\left(\sum_{u=1}^{N} x_i^{(u)}\bar{x}_i^{(u)}\right)^2 = TN^2
\end{aligned}
$$

by dropping the terms with $i \neq j$. The result follows. □

If the N sequences in the preceding theorem consist of all possible shifts of a collection of n periodic sequences (each with period T) then $N = nT$ so

$$\mathcal{C}_{\max}^2 \geq \frac{T^2(n-1)}{nT-1} \sim \frac{nT^2}{nT} = T.$$

Thus, if n is even moderately large the signal set will contain pairs whose cross-correlation is lower-bounded by approximately $\sqrt{T}$.

14.2 Families derived from a decimation

Let $F = \mathbb{F}_q \subset E = \mathbb{F}_{q^n}$ be finite fields, let $\alpha \in E$ be a primitive element and let $2 \leq d \leq |E| - 1$ be a decimation. It is possible to construct a family of pseudo-random sequences from the m-sequence $\mathbf{a} = (\mathsf{Tr}_F^E(\alpha^0), \mathsf{Tr}_F^E(\alpha^1), \mathsf{Tr}_F^E(\alpha^2), \cdots)$ and its d-fold decimation $\mathbf{b} = (\mathsf{Tr}_F^E(\alpha^0), \mathsf{Tr}_F^E(\alpha^d), \mathsf{Tr}_F^E(\alpha^{2d}), \cdots)$ by definining, for $A, B \in E$ the sequence $S(A, B)$ whose i-th term is

$$S(A,B)_i = \mathsf{Tr}_F^E(A\alpha^i + B\alpha^{di}) = \mathsf{Tr}_F^E(A\alpha^i) + \mathsf{Tr}_F^E(B\alpha^i). \tag{14.1}$$

The sequence $S(A, B)$ can evidently be generated as the termwise sum of the output sequences from two shift registers, having connection polynomials $h_1(x) = -1 + h_1^{(1)}x + \cdots + h_n^{(1)}x^n$ and $h_2(x) = -1 + h_1^{(2)}x^2 \cdots + h_r^{(2)}x^r$ respectively, such that α is a root of $h_1(x)$ and $\beta = \alpha^d$ is a root of $h_2(x)$. Here $1 \leq r \leq n$, with $r = n$ if and only if $\gcd(d, q^n - 1) = 1$. Figure 14.1 shows such an arrangement. By Proposition 6.5.1 the same sequence can be generated by a single shift register with connection polynomial $h_1(x)h_2(x)$.

As A and B vary in E the resulting sequences $S(A, B)$ are not all shift distinct. In Section 14.3 we see how to choose a collection of shift distinct sequences among the sequences $S(A, B)$.

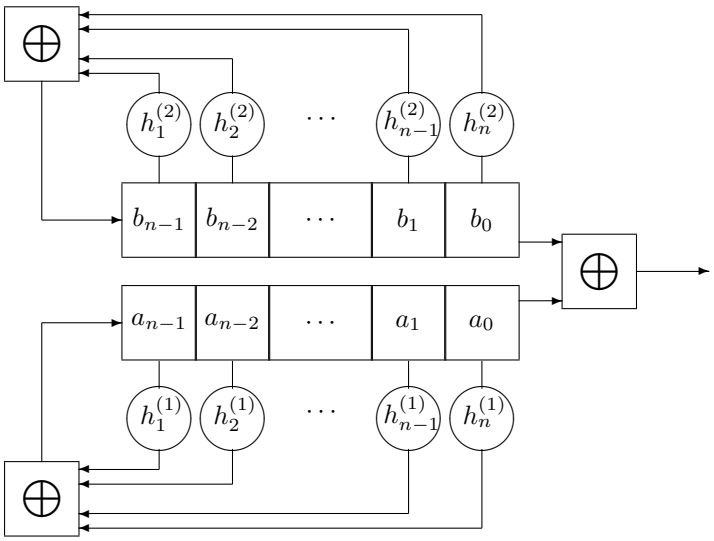


Figure 14.1: Gold sequence generator

14.3 Gold sequences

The family of (generalized) Gold sequences amounts to the construction in Section 14.2 in the case of a quadratic decimation $d = q^i + q^j$ (or equivalently, $d = 1 + q^{j-i}$). References for this section include [51, 94].

Lemma 14.3.1. *Suppose $d = 1 + q^i$ is a quadratic decimation. If n is even then assume further that $i \neq n/2$. Then the sequences in the set*

$$\mathcal{S}_1 = \{S(1, B) : \ B \in E\} \cup \{S(0, 1)\} \tag{14.2}$$

are shift distinct. If $d = 1 + q^i$ is relatively prime to $|E| - 1$ then this coincides with the set

$$\mathcal{S}_2 = \{S(A, 1) : \ A \in E\} \cup \{S(1, 0)\}$$

Proof. If $S(A, B)$ and $S(A', B')$ are sequences of the form in equaton (14.1) and if one is a left shift of the other, then setting $S(A, B)_i = S(A', B')_{i+\tau}$ gives

$$\mathsf{Tr}_F^E(x(A - CA') + x^d(B - C^d B')) = 0 \tag{14.3}$$

for all $x = \alpha^i \in E$, where $C = \alpha^\tau$. If $d = 1+q^i$, then equation (14.3) may be written $L(x)+Q(x) \equiv 0$ where $L(x) = \mathsf{Tr}_F^E(x(A - CA')$ is a linear function and $Q(x) = \mathsf{Tr}_F^E(x^d(B - C^d B')$ is a quadratic form. If, moreover, $i \neq n/2$, then by Corollary 3.3.7, $B - C^d B' = 0$ and $A - CA' = 0$. It follows

that $S(1,B)$ and $S(1,B')$ are shift distinct whenever $B \neq B'$. Similarly, $S(1,B)$ is shift distinct from $S(0,1)$.

If d is relatively prime to q^n-1, then $\mathcal{S}_1 = \mathcal{S}_2$ because both families contain $S(0,1)$ and $S(1,0)$ and if $B \neq 0$ then $S(1,B)$ is a left shift of $S(B^{-d},1)$. □

Let $\chi : F \to \mathbb{C}^\times$ be a nontrivial character. We wish to find the cross-correlation (with shift t) with respect to χ of two sequences $\mathbf{a} = S(1,A)$ and $\mathbf{b} = S(1,B)$ in $\mathcal{S}_1$. It is

$$\begin{aligned} \mathcal{C}_{\mathbf{b},\mathbf{a}}(t) + 1 &= \sum_{t=0}^{q^n-2} \chi\left(\mathsf{Tr}_F^E\left(\alpha^k - \alpha^{k+t} + A\alpha^{dk} - B\alpha^{d(k+t)}\right)\right) \\ &= \sum_{x \in E} \chi\left(\mathsf{Tr}_F^E\left(x(1-C) + x^d(A - BC^d)\right)\right) \end{aligned}$$

where $C = \alpha^t$. Let $Q(x) = \mathsf{Tr}_F^E((A - BC^d)x^d)$ and $L(x) = \mathsf{Tr}_F^E(x(1-C))$. Then

$$\mathcal{C}_{\mathbf{b},\mathbf{a}}(t) + 1 = \sum_{w \in F} N_w \chi(w) \tag{14.4}$$

where $N_w = |\{x \in E : \ L(x) + Q(x) = w\}|$. The numbers N_w are not known in general, but if $d = 1 + q^i$ is a quadratic decimation then the numbers N_w were calculated in Theorem 3.3.2. Equation (14.4) and Theorem 13.5.4 then give the following result.

Theorem 14.3.2. *Let $F = \mathbb{F}_q \subset E = \mathbb{F}_{q^n}$ be finite fields. Let $d = 1 + q^i$ be a quadratic decimation. Let $A, B \in E$ and consider the Gold sequences $\mathbf{a}, \mathbf{b}$ given by $a_k = \mathsf{Tr}_F^E(\alpha^k + A\alpha^{dk})$ and $b_k = \mathsf{Tr}_F^E(\alpha^k + B\alpha^{dk})$. Let $\chi : F \to \mathbb{C}^\times$ be a nontrivial character. Then the possible values for the cross-correlation $\mathcal{C}_{\mathbf{b},\mathbf{a}}(t)$ of these sequences with respect to the character χ is given in Table 14.1, where $g = \gcd(n,i)$ and $e = 1 + q^g$, where $H = A - B\alpha^{dt}$ and where $a = s^d$ means that $a \in E$ is the d-th power of some element $s \in E$.* □

14.4 Kasami sequences, small set

Let $F = \mathbb{F}_q \subset E = \mathbb{F}_{q^m} \subset L = \mathbb{F}_{q^{2m}}$ be finite fields, so that L is a degree two extension of E. Let $\alpha \in F$ be a primitive element and let

$$d = (|L| - 1)/(|E| - 1) = 1 + q^m.$$

Recall from Section 3.2.e that $x^d = \mathsf{N}_E^L(x) \in E$ is the norm of x (for any $x \in L$). Consequently, an element $H \in L$ is actually contained in E if and only if it is the d-th power of some element $s \in L$.

q **even**					
Conditions			Type	Rank	Values of $\mathcal{C}_{\mathbf{b},\mathbf{a}}+1$
n/g even	$n/2g$ odd	$H=s^d$	I	$n-2g$	$\{0,\pm q^{\frac{n}{2}+g}\}$
		$H\neq s^d$	III	n	$\{0,\pm q^{n/2}\}$
	$n/2g$ even	$H=s^d$	III	$n-2g$	$\{0,\pm q^{\frac{n}{2}+g}\}$
		$H\neq s^d$	I	n	$\{0,\pm q^{n/2}\}$
n/g odd			II	$n-g+1$	$\{0,\pm q^{(n+g)/2}\}$

q **odd**					
Conditions			Type	Rank	Values of $\lvert\mathcal{C}_{\mathbf{b},\mathbf{a}}+1\rvert$
n/g even	$n/2g$ odd	$H^2=s^e$ $H\neq s^e$	$\eta=-1$	$n-2g$	$\{0,q^{\frac{n}{2}-g}\}$
		otherwise	$\eta=1$	n	$\{0,q^{n/2}\}$
	$n/2g$ even	$H=s^e$	$\eta=1$	$n-2g$	$\{0,q^{\frac{n}{2}-g}\}$
		$H\neq s^e$	$\eta=-1$	n	$\{0,q^{n/2}\}$
n/g odd				n	$\{0,q^{n/2}\}$

Table 14.1: Cross-correlation for Gold sequences.

The *small set of Kasami sequences* is the collection of sequences $\mathcal{K}(A)$ as $A\in L$ varies, where

$$\mathcal{K}(A)_i=\mathsf{Tr}_F^E\left(\mathsf{Tr}_E^L(\alpha^i)+A\alpha^{di})\right)=\mathsf{Tr}_F^E\left(\mathsf{Tr}_E^L(\alpha^i)+A\mathsf{N}_E^L(\alpha^i)\right)$$

for the various possible choices $A\in L$. For any $A\in L$ there exists $\widehat{A}\in L$ such that $\mathsf{Tr}_E^L(\widehat{A})=A$. Consequently the Kasami sequence can also be written

$$\mathcal{K}(A)_i=\mathsf{Tr}_F^L(\alpha^i+\widehat{A}\alpha^{di})$$

which is therefore a Gold sequence $S(1,\widehat{A})$, with $n=2m$, $g=m$, and $e=d$. The cross-correlation between any two such sequences therefore has values as described in Table 14.1 (where $n/g=2$ is even and $n/2g=1$ is odd). However Lemma 14.3.1 does not apply in this situation because $i=n/2$. Instead, we have the following result.

Lemma 14.4.1. *If $A,A'\in E$ and $A\neq A'$ then the Kasami sequences $\mathcal{K}(A)$ and $\mathcal{K}(A')$ are shift distinct.*

Proof. Choose $\widehat{A}, \widehat{A}' \in L$ so that $\mathsf{Tr}_E^L(\widehat{A}) = A$ and $\mathsf{Tr}_E^L(\widehat{A}') = A'$. If $\mathcal{K}(A)$ coincides with the shift by t of the sequence $\mathcal{K}(A')$ then setting $C = \alpha^t$ gives

$$\mathsf{Tr}_F^L(\alpha^i(1-C) + \widehat{A}\alpha^{di} - \widehat{A}'C^d\alpha^{di}) = 0$$

for all i. Therefore the function $F(x) = L(x)+Q(x)$ is identically zero, where $L(x) = \mathsf{Tr}_F^L(x(1-C))$ and $Q(x) = \mathsf{Tr}_F^L(x^d(\widehat{A} - \widehat{A}'C^d))$. By Corollary 3.3.7 this implies that $C = 1$, that $i = n/2$ (which we already know to be satisfied), and that $\mathsf{Tr}_E^L(\widehat{A} - \widehat{A}'C^d) = 0$. But this implies that $A = A'$. □

14.5 Geometric sequences

In this section we fix a prime number $p \in \mathbb{Z}$ and consider field extensions

$$F = \mathbb{F}_p \subset L = \mathbb{F}_q \subset K = \mathbb{F}_{q^m}$$

where $q = p^e$ is some power of p. Let $\mathsf{Tr}_L^K : K \to L$ and $\mathsf{Tr}_F^L : L \to F$ denote the trace mappings. Fix a primitive element $\alpha \in K$. This gives an m-sequence whose i-th term is $\mathsf{Tr}_F^K(\alpha^i) = \mathsf{Tr}_F^L\left(\mathsf{Tr}_L^K(\alpha^i)\right)$. A *geometric sequence* [99] **a** is obtained by replacing the second function Tr_F^L by some other, possibly nonlinear function $f : L \to K$, that is,

$$a_i = f\left(\mathsf{Tr}_L^K(\alpha^i)\right). \tag{14.5}$$

This sequence may be interpreted as the result of applying a "nonlinear filter" to the output of a maximal length shift register (of size m) over L. In other words, let $q(x) \in L[x]$ be a primitive polynomial of degree m. Then the output sequence of the n-stage LFSR with connection polynomial $q(x)$ and entries in L is the sequence $\mathsf{Tr}_L^K(\alpha^i)$ where $\alpha \in K$ is a root of $q(x)$. Then apply the function $f : L \to F$ to this output, as in Figure 14.2.

We wish to calculate the autocorrelation and cross-correlations of such sequences. If $g : L \to F$ is another (possibly nonlinear) function then a second geometric sequence **b** may be obtained by starting with a (possibly different) primitive element $\beta \in K$ and setting $b_i = g(\mathsf{Tr}_L^K(\beta^i))$. Since $\beta = \alpha^d$ (for some integer d, relatively prime to $q^m - 1$) we may write $b_i = g(\mathsf{Tr}_L^K(\alpha^{di}))$, cf. Proposition 13.2.1. The underlying sequences $\mathsf{Tr}_L^K(\alpha^i)$ and $\mathsf{Tr}_L^K(\alpha^{di})$ are related by the decimation d. Recall from Section 13.5.b that d is a *linear* decimation of it is a power of q (meaning that the function $x \mapsto x^d$ is linear over L) and that d is a *quadratic* decimation if $d = q^s + q^t$ is a sum of two powers of q (meaning that the function $x \mapsto x^d$ is a quadratic function as x varies in L.)

The auto- and cross-correlation functions of the sequences **a** and **b** will be written in terms of two auxiliary quantities determined by the feed forward functions f and g. Let $\chi : F = \mathbb{F}_p \to \mathbb{C}^\times$ be a nontrivial character. Recall (Section 11.3.a) that the cross-correlation of the functions f, g with respect to the character χ is the function

$$\mathcal{C}_{g,f}(A) = \sum_{u \in L} \chi(g(u))\overline{\chi}(f(Au))$$

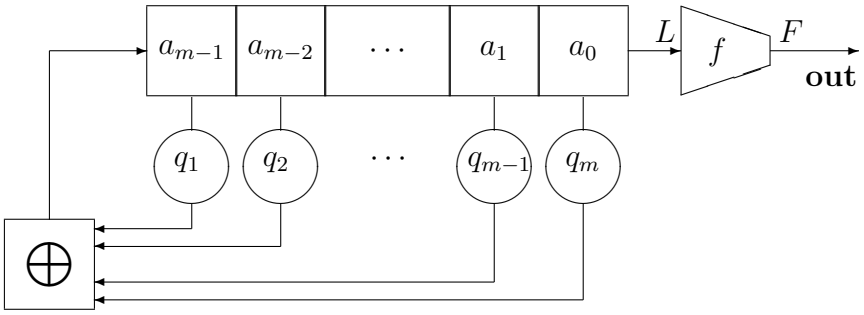


Figure 14.2: Geometric sequence generator

for any $A \in F$. Also define the imbalance of the function f to be $Z_\chi(f) = \sum_{u \in L} \chi(f(u))$. Set $C_0 = \chi(f(0))\overline{\chi}(g(0))$.

Theorem 14.5.1. *(cf. [25]) Let $a_i = f(\mathsf{Tr}_L^K(\alpha^i))$ and $b_i = g(\mathsf{Tr}_L^K(\alpha^{di}))$ where $d = q^s$ is a "linear" decimation. Let $\chi : F \to \mathbb{C}^\times$ be a non-trivial additive character. Then the autocorrelation function (with respect to χ) of $\mathbf{a}$ is*

$$\mathcal{A}_{\mathbf{a}}(\tau) = \begin{cases} q^{m-2}|Z_\chi(f)|^2 - 1 & \text{if } A \notin L \\ q^{m-1}\mathcal{C}_{f,f}(A) - 1 & \text{if } A \in L \end{cases}$$

where $A = \alpha^\tau$. The cross-correlation (with respect to χ) of $\mathbf{a}$ and $\mathbf{b}$ is

$$\mathcal{C}_{\mathbf{b},\mathbf{a}}(\tau) = \begin{cases} q^{m-2}Z_\chi(g)\overline{Z_\chi}(f) - C_0 & \text{if } A \notin L \\ q^{m-1}\mathcal{C}_{g,f}(A) - C_0 & \text{if } A \in L \end{cases}$$

We remark that $A = \alpha^t \in L$ if and only if τ is a multiple of $(|K|-1)/(|L|-1) = (q^m-1)/(q-1)$. Since the cross-correlation of the sequences $\mathbf{a}, \mathbf{b}$ is ultimately described in terms of the cross-correlation of the functions f, g this procedure can be iterated: the function $f : L \to F$ could itself be taken to be of the form $f(x) = f'(\mathsf{Tr}_E^L(x))$ for some intermediate field E and some function $f' : E \to F$. This approach was used in the construction of *cascaded GMW sequences* [99]. A similar approach may be applied to quadratically decimated geometric sequences [25].

Proof. We use the same technique that was used in Section 13.5.b. For $u, v \in L$ let

$$\begin{aligned} Q_u &= \left\{x \in K : \ \mathsf{Tr}_L^K(x^d) = v\right\} \\ H_v &= \left\{x \in K : \ \mathsf{Tr}_L^K(x) = u\right\} \end{aligned}$$

Then H_v is an (affine) hyperplane in K and Q_u is a hypersurface in K. Each $A \in K$ acts by multiplication so

$$A^{-1}H_v = \left\{A^{-1}x \in K : \ \mathsf{Tr}_L^K(x) = v\right\} = \left\{y \in K : \ \mathsf{Tr}_L^K(Ay) = v\right\}.$$

The cross-correlation with shift τ is:

$$\begin{aligned}
\mathcal{C}_{\mathbf{b},\mathbf{a}}(\tau) &= \sum_{i=0}^{|K|-2} \chi(g(\mathsf{Tr}_L^K(\alpha^{di})))\overline{\chi}(f(\mathsf{Tr}_L^K(\alpha^{i+\tau}))) \\
&= \sum_{x\in K} \chi(g(\mathsf{Tr}_L^K(x^d)))\overline{\chi}(f(\mathsf{Tr}_L^K(Ax))) - C_0 \\
&= \sum_{v\in L}\sum_{u\in L} \chi(g(u))\overline{\chi}(f(v))|Q_u \cap A^{-1}H_v| - C_0
\end{aligned}$$

where $A = \alpha^\tau$. Here we have used the fact that if i varies over $0 \le i \le |K| - 2$, then α^i varies over all nonzero elements of K. Since we assume d is a power of q we have: $\mathsf{Tr}_L^K(x^d) = \mathsf{Tr}_L^K(x)$ so the hypersurface $Q_u = H_u$ is a hyperplane. According to Lemma 13.5.2 the hyperplanes H_u and $A^{-1}H_v$ are parallel if and only if $A \in L$. If $A \notin L$ the intersection $H_u \cap H_{A^{-1}v}$ is a subspace of dimension $m-2$ in K which gives $\mathcal{C}_{\mathbf{b},\mathbf{a}} = Z_\chi(g)\overline{Z_\chi}(f)q^{m-2} - C_0$. If $A \in L$ then $A^{-1}H_v = H_{A^{-1}v}$ and its intersection with H_u is empty unless $v = Au$ in which case the cross-correlation is

$$\mathcal{C}_{\mathbf{b},\mathbf{a}}(\tau) = \sum_{u\in L} \chi(g(u))\overline{\chi}(f(Au))r^{m-1} - C_0. \qquad \square$$

14.6 GMW sequences

As in Section 14.5, consider fields

$$F = \mathbb{F}_p \subset L = \mathbb{F}_q \subset K = \mathbb{F}_{q^m}$$

and let $\alpha \in K$ be a primitive element. Fix an integer h, relatively prime to $q-1$. The *GMW sequence* $\mathbf{a}$ (cf. [57, 175, 100]) based on the data (α, h) is a geometric sequence in which the function $f : L \to F$ is $f(x) = \mathsf{Tr}_F^L(x^h)$, that is,

$$a_i = \mathsf{Tr}_F^L\left(\mathsf{Tr}_L^K(\alpha^i)^h\right). \tag{14.6}$$

Theorem 14.6.1. *The GMW sequence in equation (14.6) has full period, $q^m - 1$, and perfect autocorrelation function. It is balanced with respect to any character. It is equidistributed to order $[K : L] = m$. If h is a power of p then it is a shift-and-add sequence.*

Proof. Let $\chi : F \to \mathbb{C}^\times$ be a nontrivial (additive) character. According to Theorem 14.5.1 the autocorrelation of $\mathbf{a}$ is determined by the imbalance and autocorrelation functions of f. The imbalance is $Z_\chi(f) = \sum_{u\in L} \chi(f(u)) = \sum_{u\in L} \chi(u^h)$. Since h is relatively prime to $q-1$, the elements u^h exactly vary over the elements of L as u varies over the elements of L so $Z_\chi(f) = 0$. The autocorrelation function of f is

$$\begin{aligned} \mathcal{C}_{f,f}(A) &= \sum_{u\in L} \chi(f(u))\overline{\chi}(f(Au)) \\ &= \sum_{x\in F}\sum_{y\in F} \chi(x)\overline{\chi}(y)|H_x \cap A^{-1}H_y| \end{aligned}$$

where $H_x = \left\{u \in L : \ \mathsf{Tr}_F^L(u) = x\right\}$ and similarly for H_y. Using Lemma 13.5.2 as we did in Section 14.5, these are non-parallel hyperplanes unless $A \in F$. If $A \notin F$ this gives

$$\mathcal{C}_{f,f}(A) = p^{e-2}\sum_{x\in F} \chi(x) \sum_{y\in F} \overline{\chi}(y) = 0$$

while if $A \in F$ this gives 0 (because parallel hyperplanes do not intersect) unless $y = Ax$, in which case it gives

$$\mathcal{C}_{f,f}(A) = p^{e-1}\sum_{x\in F} \chi(x)\overline{\chi}(Ax).$$

The function $\eta(x) = \chi(x)\overline{\chi}(Ax)$ is an additive character $\eta : F \to \mathbb{C}^\times$, which is non-trivial unless $A = 1$. Hence $\sum_{x\in F} \eta(x) = 0$, which completes the proof that the autocorrelation function is perfect.

To see that the GMW sequence $\mathbf{a}$ is equidistributed to order $[K : L]$, we first observe that the sequence $y_i = \mathsf{Tr}_L^K(\alpha^i) \in L$ is equidistributed to order $[K : L]$ since it is an m-sequence. It follows that the sequence $\mathbf{y} = (y_0^h, y_1^h, \cdots)$ is also equidistributed to order $[K : L]$ since the function $y \mapsto y^h$ is a permutation of the set L (and it takes 0 to 0). In fact, each block $(z_1, \cdots, z_m)$ of size $m = [K : L]$ occurs exactly once in a single period of $\mathbf{y}$ (except for the all-zero block). Now let $b = (b_1, b_2, \cdots, b_m)$ with $b_i \in F$, be a block of size $[K : L]$. For each $b \in F$ there are $|L|/|F| = p^{e-1}$ elements $z \in L$ such that $\mathsf{Tr}_F^L(z) = b$. Consequently there are $p^{m(e-1)}$ blocks $(z_1, z_2, \cdots, z_m)$ of size $m = [K : L]$, with $z_i \in L$ such that $(\mathsf{Tr}_F^L(z_1), \mathsf{Tr}_F^L(z_2), \cdots, \mathsf{Tr}_F^L(z_m)) = (b_1, b_2, \cdots, b_m)$. But each such block $(z_1, z_2, \cdots, z_m)$ occurs once in the sequence $\mathbf{y}$, unless it is the all-zero block. Hence the block $(b_1, b_2, \cdots, b_m)$ occurs $p^{m(e-1)}$ times in the sequence $\mathbf{a}$, except for the all-zero block, which occurs $p^{m(e-1)} - 1$ times.

Now suppose h is a power of p. To see that the sequence $\mathbf{a}$ is a shift-and-add sequence, consider a shift τ and let $A = \alpha^\tau \in K$. Assuming $A \neq -1$, there is a unique s $(0 \le s \le p^m - 2)$ such that

$1 + A = \alpha^s$. The sum of **a** plus its τ-shift is the sequence $\mathbf{b} = (b_0, b_1, \cdots)$ where

$$\begin{aligned} b_i = a_i + a_{i+\tau} &= \mathsf{Tr}_F^L\left((\mathsf{Tr}_L^K(\alpha^i))^h\right) + \mathsf{Tr}_F^L\left((\mathsf{Tr}_L^K(\alpha^{i+\tau}))^h\right) \\ &= \mathsf{Tr}_F^L\left((\mathsf{Tr}_L^K(\alpha^i) + \mathsf{Tr}_L^K(A\alpha^i))^h\right) \\ &= \mathsf{Tr}_F^L\left((\mathsf{Tr}_L^K(\alpha^s\alpha^i))^h\right) \\ &= a_{i+s}. \end{aligned}$$

□

14.7 d-form sequences

In this section we restrict our attention to characteristic 2. Let $m, e \geq 1$ and $q = 2^e$. Let

$$F = \mathbb{F}_2 \subset L = \mathbb{F}_q \subset K = \mathbb{F}_{q^m}$$

Let $d \geq 1$. Then a *d-form* on $\mathbb{F}_r$ is a function $H : K \to L$ satisfying

$$H(ax) = a^d H(x)$$

for every $a \in L$ and $x \in K$. For example, $Tr_L^K(x^d)$ is a d-form.

Let $H(x)$ be a d-form and let $\alpha \in K$ be a primitive element. Let $k \geq 1$. Then the sequence $A = a_0, a_1, \cdots$ defined by

$$a_i = Tr_F^L(H(\alpha^i)^k)$$

is called a *d-form sequence*. The proof of the following theorem is outlined in the exercises.

Theorem 14.7.1. *([93]) Let H_1 and H_2 be two d-forms and let $A^j = a_0^j, a_1^j, \cdots$, $j = 1, 2$, be the d-form sequences they define using the same k. For any τ let*

$$z_\tau = |\{x \neq 0 \in K : H_1(x) + H_2(\alpha^\tau x) = 0\}|.$$

Then the cross correlation (with shift τ) of A^1, A^2 is:

$$C_{A^1,A^2}(\tau) = \frac{qz_\tau - (q^m - 1)}{q - 1}. \tag{14.7}$$

14.8 Legendre and Dirichlet sequences

Although they are not directly related to shift registers, the Legendre and Dirichlet sequences are included here because they have ideal autocorrelation functions. Let p be a prime number. A *Dirichlet character* modulo p is a (multiplicative) homomorphism $\psi : \mathbb{Z}/(p)^\times \to \mathbb{C}^\times$. Since

$x^{p-1} = 1$ for all $x \in \mathbb{Z}/(p)^\times$, such a character takes values in the set μ_{p-1} of roots of unity so $\overline{\psi}(x) = \psi(x^{-1})$. An example is the *Legendre symbol* or *quadratic character*

$$\psi(x) = \left(\frac{x}{p}\right) = \begin{cases} +1 & \text{if } x \text{ is a square} \\ -1 & \text{otherwise.} \end{cases}$$

It is customary to extend each Dirichlet character to all of $\mathbb{Z}/(p)$ by setting $\psi(0) = 0$, but there is no "correct" value for $\psi(0)$ and engineers often take $\psi(0) = 1$ to obtain a sequence of ± 1 values. The *Dirichlet sequence* corresponding to a Dirichlet character ψ is the periodic sequence $\psi(0), \psi(1), \psi(2), \cdots$. It is called a *Legendre sequence* or *quadratic residue sequence* if ψ is the quadratic character (in which case we set $\psi(0) = 1$). In either case, let us consider the autocorrelation $\mathcal{A}(\tau)$ with shift $\tau \not\equiv 0 \pmod{p}$. (Note that $\mathcal{A}(0) = p$.)

Proposition 14.8.1. *Let ψ be a non-trivial Dirichlet character. Let $\tau \not\equiv 0 \pmod{p}$. Set*

$$R = \psi(0)\left(\psi(\tau) + \overline{\psi}(-\tau)\right) = \begin{cases} 0 & \text{if } \psi(0) = 0 \\ 2\mathrm{Re}\left(\psi(\tau)\right) & \text{if } \psi(0) = 1, \psi(-1) = 1 \\ 2i\mathrm{Im}\left(\psi(\tau)\right) & \text{if } \psi(0) = 1, \psi(-1) = -1. \end{cases}$$

Then the autocorrelation with shift τ of the resulting Dirichlet sequence is

$$\mathcal{A}(\tau) = -1 + R.$$

Proof. Let $\tau \neq 0$. Assuming $\psi(0)$ is real, the autocorrelation is

$$\begin{aligned} \mathcal{A}(\tau) = \sum_{x=0}^{p-1} \psi(x+\tau)\overline{\psi}(x) &= \sum_{x \neq 0, -\tau} \psi\left((x+\tau)x^{-1}\right) + \psi(\tau)\overline{\psi}(0) + \psi(0)\overline{\psi}(-\tau) \\ &= \sum_{x \neq 0, -\tau} \psi(1 + x^{-1}\tau) + \psi(0)\left(\psi(\tau) + \overline{\psi}(-\tau)\right). \end{aligned}$$

As x varies within $\mathbb{Z}/(p)^\times$ the quantity $y = 1 + x^{-1}\tau$ takes on all values except $y = 1$. It takes the value $y = 0$ when $x = -\tau$. Hence,

$$\mathcal{A}(\tau) = \sum_{y \neq 0,1} \psi(y) + R = \sum_{y \neq 0} \psi(y) - \psi(1) + R = -1 + R$$

because for any group G and any nontrivial character ψ, $\sum_{g \in G} \psi(g) = 0$ and $\psi(1) = 1$. □

We remark that the quadratic character satisfies $\psi(-1) = -1$ if and only if $p \equiv 3 \pmod{4}$, in which case $R = 0$. This case was first described in [157].

14.9 Frequency hopping sequences

Frequency hopping as a spread-spectrum technique is described in Section 11.3.e. The key requirement is the generation of pseudorandom sequences with small Hamming cross-correlation. In this section we describe several methods of generating such sequences.

14.9.a The Lempel-Greenberger method

In [121], Lempel and Greenberger showed how to use an m-sequence to construct an optimal family of sequences. Their method has since been applied in a variety of different situations, and it may be described in general terms as follows.

Proposition 14.9.1. *Let V be a vector space over a finite field and let $\mathbf{a} = (a_0, a_1, \cdots)$ be a shift and add sequence (see Chapter 12) with period T and with entries in V, which do not all lie in a proper subspace of V. For any $v \in V$ let $\mathbf{a} + v$ be the translated sequence $(a_0 + v, a_1 + v, \cdots)$. Then any two such sequences $\mathbf{a} + v, \mathbf{a} + w$ are shift distinct (provided $v \neq w$) and their Hamming cross-correlation is*

$$\mathcal{C}^{\mathrm{Ham}}(\mathbf{a} + v, \mathbf{a} + w) = \begin{cases} \lceil T/|V| \rceil & \text{if} \quad v \neq w \\ \lfloor T/|V| \rfloor & \text{if} \quad v = w \end{cases}$$

Consequently the family of sequences $\mathbf{a} + v$ (for different choices of $v \in V$) constitute an optimal frequency hopping family, with $|V|$ shift distinct members.

Proof. If the finite field in question has characteristic p, then V is also a vector space over the prime field $\mathbb{F}_p$. By Theorem 12.2.2, the period of $\mathbf{a}$ is $p^n - 1$ for some n. Moreover, there exists an $\mathbb{F}_p$-linear mapping $T : \mathbb{F}_{p^n} \to V$ and there exists a primitive element $\beta \in \mathbb{F}_{p^n}$ such that $a_i = T(\beta^i)$. If $v, w \in V$ and $v \neq w$, then the sequences $\mathbf{a} + v, \mathbf{a} + w$ are shift distinct. (Otherwise there exists a shift τ such that $T(\beta^i - \beta^{i+\tau}) = w - v$ for all i which implies that T is a constant mapping, so it is zero.)

Let $v, w \in V$. The Hamming cross-correlation $\mathcal{C}^{\mathrm{Ham}}_{\mathbf{a}+v,\mathbf{a}+w}(\tau)$ with shift τ is therefore equal to the number of values of i $(0 \leq i \leq p^n - 2)$ such that $T(\beta^i - \beta^{i+\tau}) = w - v$. This is the number of nonzero $x \in \mathbb{F}_{p^n}$ such that $T(Ax) = w - v$ where $A = 1 - \beta^\tau$. Since T is linear and the values a_i do not lie in a proper subspace, the mapping T is surjective. Consequently, if $\tau \neq 0$ then this number is $p^n/|V| = \lceil T/|V| \rceil$ unless $w - v = 0$ in which case it is $p^n/|V| - 1 = \lfloor T/|V| \rfloor$. □

14.9.b Examples of FH sequences

From Section 12.3 we obtain the following families with optimal Hamming correlations.

1. Let q be a prime power and let $1 \leq j \leq r$ be integers. By taking windows of size j from an m-sequence of degree m over $\mathbb{F}_q$ we obtain a balanced shift and add sequence of period $q^m - 1$ over

the alphabet $V = \mathbb{F}_q^j$, as described in Section 12.3.b. Applying Proposition 14.9.1 gives a family of balanced FH sequences (over the same alphabet F with the same period $q^m - 1$), which contains $|V| = q^j$ shift distinct members. This family was originally described in [121].

2. Let $F \subset L \subset K$ be finite fields of characteristic p and let h be a power of p. Let $1 \leq j \leq [K : L]$. By taking windows of size j in the GMW sequence

$$a_i = \mathsf{Tr}_F^L \left(\left(\mathsf{Tr}_L^K(\alpha^i) \right)^h \right)$$

(as in Section 12.3.c) we obtain a balanced shift and add sequence of period $|K| - 1$ over the alphabet $V = F^j$. Applying Proposition 14.9.1 gives a family of balanced FH sequences (also over V) which contains $|F|^j$ shift distinct members.

3. Let F be a finite field, let $\pi(x), q(x) \in F[x]$ be irreducible polynomials with $\deg(q) = k \deg(\pi)$ for some $k \geq 2$ and assume that $\pi(x)$ is primitive modulo $q(x)$. This gives rise to a maximal length shift and add *function field sequence* (see Section 15.1.c) of period $|F|^{\deg(g)} - 1$ with symbols in the field $F(x)/(\pi(x))$. As in Section 12.3.d taking windows of size $j \leq k$ in this sequence gives a balanced shift and add sequence of period $|F|^{deg(q)} - 1$ whose symbols lie in $V = (F(x)/(\pi))^j$. Applying Proposition 14.9.1 gives a family of sequences (with the same period, over the same alphabet) containing $|V| = |F|^{j \deg(\pi)}$ shift distinct members.

4. Families of frequency hopping sequences may also be constructed from m-sequences over finite local rings (see Section 14.10.c) as described in [185].

14.10 Maximal sequences over a finite local ring

Let R be a finite local ring (see Section 4.1) with maximal ideal $\mathfrak{m}$ and residue field $\mu : R \to F = R/\mathfrak{m}$ having $q = |F|$ elements. Consider a LFSR over R with (irreducible) connection polynomial $g(x) \in R[x]$ of degree d. We are interested in the output sequences of such a shift register or equivalently, sequences (of elements in R) that satisfy the linear recurrence defined by $g(x)$. There does not appear to be universal agreement as to the meaning of "m-sequence", or "MP-sequence" (maximal period) over R because the possible multiplicative orders of elements in finite rings are not always known. We will consider two classes of sequences. The first, referred to as "m-sequences" in [185] have period $q^d - 1$. However, longer periods are possible. Sequences with the greatest possible period (given the ring R and the degree d) are referred to as "MP-sequences" in [117] and as "ML-sequences" in [32]. References for this section include [32, 116, 184, 185, 187].

14.10.a Generalities on LFSR sequences over a finite local ring

Let us assume that $g(x) \in R[x]$ is a *basic irreducible polynomial* (meaning that $\mu(g)$ is irreducible) whose leading term is invertible in R (so that $\deg(g) = \deg(\mu(g))$). Assume $g(0) = -1$, and consider a LFSR over R with connection polynomial $g(x)$. Let $d = \deg(g)$ and let S be the unique degree d Galois extension of R (see Theorem 4.4.1). Then, as in equation (4.6), we have a diagram

$$\begin{array}{ccc} S & \xrightarrow{\nu} & K{=}S/\mathfrak{M} \\ \Big\downarrow{\scriptstyle Tr_{S/R}} & & \Big\downarrow{\scriptstyle Tr_{K/F}} \\ R & \xrightarrow{\mu} & F{=}R/\mathfrak{m}. \end{array}$$

Let $q = |F|$ and $Q = q^d = |K|$. By equation (4.2), there exists j such that $|R| = q^j$ and $|\mathfrak{m}| = q^{j-1}$, hence $|S| = |R|^d = Q^j$ and $|\mathfrak{M}| = |\mathfrak{m}|^d = Q^{j-1}$.

Let $\alpha \in S$ be a root of $g(x)$. According to Proposition 6.6.5, for any initial state of the shift register there exists a unique $A \in S$ such that the output sequence $\mathbf{a} = (a_0, a_1, \cdots)$ is given by

$$a_n = \mathsf{Tr}_R^S(A\alpha^{-n}). \tag{14.8}$$

Since the mapping $\psi : S \to R^m = \Sigma$ of Section 6.6.b is a bijection, we see that

1. The period of the sequence $\mathbf{a}$ is equal to the (multiplicative) order of α and
2. $A \in \mathfrak{M}$ if and only if $a_n \in \mathfrak{m}$ for all $n \geq 0$.

It follows from point (1) that the greatest possible period for the sequence $\mathbf{a}$ is obtained when α is a generator of the largest cyclic subgroup of the (multiplicative) group $S^\times$ of invertible elements.

By Proposition 4.1.3, the group $S^\times$ is the product of two subgroups,

$$S^\times \cong K^\times \times (1 + \mathfrak{M}), \tag{14.9}$$

where $K^\times$ is cyclic of order $Q - 1$, and $1 + \mathfrak{M}$ has order $Q^{j-1} = q^{(j-1)d}$. Recall from Section 4.1 that the isomorphism in equation (14.9) is induced by a splitting $\iota : K^\times \to S^\times$ of the short exact sequence

$$1 \to 1 + \mathfrak{M} \to S^\times \to K^\times \to 1.$$

Let $Z \subset 1+\mathfrak{M}$ be a cyclic subgroup of largest order. Then its order $|Z|$ divides Q^j so it is relatively prime to $Q-1$. Hence $K^\times \times Z$ is a cyclic subgroup of $S^\times$ of largest order. By taking $\alpha \in K^\times \times Z$ to be a generator, we will obtain a sequence $\mathbf{a}$ with period $(Q-1)|Z|$, which is the largest possible. On the other hand, sequences with period $Q - 1$, described in the next section, are sometimes referred to as "m-sequences".

14.10.b m-sequences

If the reduction $\mu(g) \in F[x]$ is a primitive polynomial then $\mu(\alpha) \in K$ is a primitive element so its order in $K^\times$ is $Q-1$. It follows that the sequence a_n of equation (14.8) has a period which is a multiple of $Q-1$. However if $\alpha \in \iota(K)$ then $\alpha^{Q-1} = 1$ and we conclude that the sequence a_n has period $Q-1$.

Definition 14.10.1. *A (generalized)* m-sequence *over R is a sequence of the form $a_n = Tr_R^S(A\alpha^{-n})$ where $\alpha \in \iota(K)$ is a primitive element (*Section *4.4.c) and where $A \in S$ is a unit.*

By Lemma 4.4.7, such an element α is a root of a monic basic irreducible polynomial $g(x) \in R[x]$ such that $g(x)$ divides $x^{Q-1} - 1$. So the m-sequences of degree d over R are precisely the output sequences of linear feedback shift registers over R whose connection polynomial is primitive and divides $x^{Q-1} - 1$.

14.10.c m-squences over polynomials rings

As in the preceding section, assume that R is a finite local ring with residue field $F = \mathbb{F}_q$ and S is its Galois extension of degree d and residue field $K = \mathbb{F}_Q$. Identify F with its image $\iota(F) \subset R$ consisting of all elements $u \in R$ such that $u^q = u$, and similarly $\iota(K) = \left\{v \in K : \ v^Q = v\right\}$.

In this section we suppose that R is itself a polynomial residue class ring, say, $R = L[\xi]/(w(\xi)^k)$ where $L = \mathbb{F}_\ell$ is a finite field and $w(\xi) \in L[\xi]$ is an irreducible polynomial of degree m. By Proposition 4.2.3, the ring R is local, with $F \cong L[\xi]/(w(\xi))$, and $|F| = q = \ell^m$. The maximal ideal $\mathfrak{m} \subset R$ is principal and is generated by $w(\xi)$. That is, $\mathfrak{m} = (w(\xi))$.

Lemma 14.10.2. *If R and S are polynomial residue class rings as in the preceding paragraph, then the trace $\mathsf{Tr}_R^S : S \to R$ takes $\iota(K)$ to $\iota(F)$ and the resulting mapping is $\mathsf{Tr}_F^K : \iota(K) \to \iota(F)$.*

Proof. Let $v \in \iota(K)$. We need to show that $(\mathsf{Tr}_S^R(v))^q = \mathsf{Tr}_S^R(v)$. By Lemma 4.4.7 there exists a primitive polynomial $f(x) \in R[x]$ whose roots are primitive elements of $\iota(K)$ which are permuted by elements of the Galois group $\mathrm{Gal}(S/R)$. Let α be such a root. Hence $v = \alpha^t$ for some t. Moreover, α^q is also a root of f so there exists $\tau \in \mathrm{Gal}$ such that $\alpha^q = \tau(\alpha)$. Consequently

$$\begin{aligned}
\left(\mathsf{Tr}_S^R(v)\right)^q &= \left(\sum_\sigma \sigma(\alpha^t)\right)^q \\
&= \sum_\sigma \sigma(\alpha^q)^t \\
&= \sum_\sigma \sigma(\tau(\alpha))^t \\
&= \mathsf{Tr}_S^R(v)
\end{aligned}$$

where the sum is over $\sigma \in \mathrm{Gal}(S/R)$. This shows that the trace takes $\iota(K)$ to $\iota(F)$ and Lemma 4.4.3 shows that it agrees with the usual trace. □

As in equation (14.8) the output sequence of the LFSR is given by $a_i = \mathsf{Tr}_R^S(A\alpha^{-i})$ and Lemma 14.10.2 shows that if $A \in \iota(K)$ then $a_i \in \iota(F)$ forms an m-sequence in the usual sense. However, by choosing $A \in 1+\mathfrak{M} \subset S^\times$ we obtain $|1+\mathfrak{M}| = |\mathfrak{m}|^d$ shift distinct sequences with $a_i \in R$. These sequences and their translates are proposed for applications in frequency hopping in [185], where their Hamming cross-correlations are calculated.

14.10.d ML sequences over Galois rings

We continue to assume, as in Section 14.10.a that R is a finite local ring with residue field $F = \mathbb{F}_q$ and S is its Galois extension of degree d and residue field $K = \mathbb{F}_{q^d}$. In this section we consider the case that $R = GR(p^m, e)$ is the degree e Galois extension of $\mathbb{Z}/(p^m)$ where p is prime, see Section 4.5. Then $|R| = p^{me}$ and $F = \mathbb{F}_{p^e}$. It follows that $S = GR(p^m, ed)$ and so, by Section 4.5 the multiplicative group $1+\mathfrak{M}$ contains cyclic subgroups of order p^{d-1}. Therefore, by appropriate choice of α (or equivalently, by appropriate choice of the linear recursion $g(x)$ and of the initial loading), the sequence $\mathbf{a}$ will have (minimal) period $(Q-1)p^{d-1} = (p^{ed}-1)p^{d-1}$. Such sequences are referred to as "MP-sequences" or "ML-sequences" (maximal period, resp. maximal length) in the literature. Techniques for finding such a polynomial $g(x)$ are described in [187, 32] (for the case $e = 1$), and [116, 117, 184] (in general).

14.11 Exercises

1. The purpose of this exercise is to prove Theorem 14.7.1. Assume the hypotheses of that theorem.

   a. Let $t = (q^m-1)/(q-1)$, and for $0 \le i < q^m - 1$, let $i = i_0 + ti_1$ with $0 \le i_0 < t$ and $0 \le i_1 < q-1$. Let
   $$f(i) = (H_1(\alpha^i))^k + (H_2(\alpha^{i+\tau}))^k.$$
   Show that
   $$a_i^1 + a_{i+\tau}^2 = Tr_2^q(\alpha^{dki_1t} f(i_0)).$$

   b. The cross-correlation is the sum over all $0 \le i < q^m - 1$ of
   $$(-1)^{a_i^1 + a_{i+\tau}^2}.$$
   We can sum separately over i_0 and i_1. First fix i_0 with $f(i_0) \neq 0$. What is the contribution as i_1 varies? Next fix i_0 with $f(i_0) = 0$. What is the contribution from such terms?

   c. Let $w = |\{i_0 : f(i_0) = 0\}|$. Using part (b), show that $C_{A^1,A^2}(\tau) = qw - t$.

   d. Show that $w = z_\tau/(q-1)$ and use this to complete the proof of equation (14.7).

e. Let $m = 2g$ and let $a \in \mathbb{F}_q$. Show that

$$H_a(x) = Tr_q^{q^g}(ax^{q^g+1} + Tr_{q^g}^{q^m}(x^2))$$

is a 2-form (also called a quadratic form).

f. Let A and B be two sequences defined by 2-forms H_a and H_b as in part (e). It can be shown that for any τ we have

$$z_\tau \in \{q^{2g-1} - 1, q^{2g-1} - q^g + q^{g-1} - 1, q^{2g-1} + q^g - q^{g-1} - 1\}.$$

Use this fact and equation (14.7) to show that $C_{A,B}(\tau) \in \{-q^g - 1, -1, q^g - 1\}$.

Chapter 15 Maximal Period Function Field Sequences

In this section we develop the theory of AFSRs based on function fields. Such an AFSR is a feedback-with-carry shift register, however the register contents and the multipliers are themselves polynomials. That is, they are elements of the ring $R = F[x]$, where F is a finite field. The ring R is an integral domain whose fraction field is the field $K = F(x)$ of *rational functions* $f(x)/g(x)$ (cf. Section 3.5 and 5.2.d), whence the name "function field AFSR". In Section 15.2 we consider a more general setting in which the register contents and multipliers are taken from a ring R whose fraction field is any *global function field*, that is, any finite extension of $F(x)$. Although function field AFSRs are considered in Sections 8.3.b and 8.3.c, the present chapter may be read independently.

15.1 The Rational function field AFSR

Let F be a finite field. Let $\pi \in F[x]$ be a polynomial. Define $S \subset F[x]$ to be the collection of all polynomials of degree less than $\deg(\pi)$. It follows from the division algorithm (Theorem 5.5.3) that the set S is a complete set of representatives for the quotient ring $F[x]/(\pi)$. The set S is closed under addition, but not under multiplication.

Let $q(x) \in F[x]$ be a polynomial, relatively prime to $\pi(x)$. For any $u \in F[x]$ let $\tilde{u} \in F[x]/(\pi)$ and $\bar{u} \in F[x]/(q)$ denote its image in the respective quotients. Recall from Theorem 2.2.15 that coprimality of the polynomials $\pi(x)$ and $q(x)$ is equivalent to the statement that $\tilde{q} \in F[x]/(\pi)$ is invertible, or that $\bar{\pi}(x) \in F[x]/(q)$ is invertible. By the division theorem, the polynomial $q(x)$ has a unique expansion,

$$q(x) = q_0(x) + q_1(x)\pi(x) + \cdots + q_m(x)\pi(x)^m \tag{15.1}$$

such that $q_i \in S$. Consequently $\tilde{q} = \tilde{q_0}$, and we will sometimes denote this by $q \pmod{\pi}$.

Definition 15.1.1. *Fix $\pi(x) \in F[x]$. A function field sequence $\mathbf{a} = (a_0, a_1, \cdots)$ (with $a_i \in F[x]/(\pi)$) is the output sequence of an AFSR based on $(F[x], \pi, S)$ with connection element $q(x) \in F[x]$, relatively prime to $\pi(x)$. (See Corollary 15.1.3 for an alternate characterization of function field sequences.)*

Note that if $\pi(x) = x$, then a function field AFSR is an LFSR.

Let us take a few paragraphs to repeat the salient properties of AFSRs in the case of function field AFSRs. The symbols a_i in the sequence are elements of $F[x]/(\pi)$ which may be represented by polynomials of degree less than $\deg(\pi)$, the collection of which is denoted by S. The sequence is

generated by the finite state machine illustrated in Figure 8.1. The "multipliers" $q_0, q_1, \cdots, q_m \in S$ are chosen so that $q_0(x)$ is relatively prime to $\pi(x)$, and the connection element is defined to be $q(x) = \sum_{i=0}^{m} q_i(x)\pi(x)^i$. The *state*, written $(a_0, a_1, \cdots, a_{m-1}; z_{m-1})$, consists of the *loading* $(a_0, a_1, \cdots, a_{m-1})$, with $a_i(x) \in S$, and the *memory* $z_{m-1} \in F[x]$ with z_{m-1} a polynomial of any degree.

Given the initial loading $(a_0, a_1, \cdots, a_{m-1})$ with initial memory $z_{m-1} \in F[x]$, the next state is computed from the linear recursion with carry (cf. equation (8.8)),

$$-q_0 a_m + \pi z_m = z_{m-1} + \sum_{i=1}^{m} q_i a_{m-i}. \tag{15.2}$$

In other words, set $\sigma(x) = z_{m-1}(x) + \sum_{i=1}^{m} q_i(x) a_{m-i}(x)$ (= memory + feedback) $\in F[x]$. Then, as described in Definition 8.1.1, the next state is $(a_1, a_2, \cdots, a_m; z_m)$ with

$$a_m \equiv \theta\sigma \pmod{\pi} \quad \text{and} \quad z_m = \sigma \ (\text{div } \pi) \tag{15.3}$$

where $\theta = -\tilde{q}_0^{-1} \in F[x]/(\pi)$, or by abuse of notation, $\theta = -q_0^{-1} \pmod{\pi}$.

To each state $(a_0, a_1, \cdots, a_{m-1}; z_{m-1})$ let us associate the polynomial

$$u(x) = \sum_{j=0}^{m-1} \sum_{i=0}^{j} q_i(x) a_{j-i}(x) \pi(x)^j - z_{m-1}(x)\pi(x)^m. \tag{15.4}$$

Theorem 15.1.2. *The association $\{states\} \to F[x]$ in equation (15.4) is a one to one correspondence. The output sequence* $\mathbf{a}$ *of the AFSR with initial loading $(a_0, a_1, \cdots, a_{m-1}; z_{m-1})$ is precisely* $\mathbf{seq}_\pi(u/q)$. *In other words, it is the π-adic expansion of the fraction*

$$\frac{u(x)}{q(x)} = \sum_{i=0}^{\infty} a_i \pi^i, \tag{15.5}$$

where $q(x)$ is given by equation (15.1) and $u(x)$ is given by equation (15.4). If $u(x) \in F[x]$ corresponds to a given state $(a_0, a_1, \cdots, a_{m-1}; z_{m-1})$ then the polynomial $u'(x) \in F[x]$ that corresponds to the succeeding state is

$$u' = (u - a_0 q)/\pi. \tag{15.6}$$

Proof. The mapping $\{\text{states}\} \to F[x]$ is a bijection by Lemma 8.2.1. The output sequence coincides with the π-adic expansion (15.5) by Theorem 8.2.2. If u/q is given by equation (15.5), then the corresponding fraction for the succeeding state is

$$\frac{u'}{q} = a_1 + a_2\pi + \cdots = \frac{u/q - a_0}{\pi}$$

□

Corollary 15.1.3. *Every function field sequence* $\mathbf{a} = (a_0, a_1, \cdots)$ *(with* $a_i \in F[x]/(\pi)$*) is the coefficient sequence* $\mathbf{seq}_\pi(u/q)$ *of the* π*-adic expansion of a fraction* $u(x)/q(x)$ *such that* $q(x) \in F[x]$ *is relatively prime to* $\pi(x)$*.*

15.1.a Periodicity

As in the preceding section, fix relatively prime polynomials $\pi(x), q(x) \in F[x]$.

Proposition 15.1.4. *Let* $u(x) \in F[x]$*. The sequence* $\mathbf{a} = \mathbf{seq}_\pi(u/q)$ *is eventually periodic. It is strictly periodic if and only if* $\deg(u) < \deg(q)$*. In this case the memory* z *is bounded:* $\deg(z) \leq \deg(\pi) - 1$*, and the (minimal) period of* $\mathbf{a}$ *is the multiplicative order of* π *modulo* q*. That is, it is the smallest positive integer* N *such that* $\pi^N \equiv 1 \pmod{q}$*. (Equivalently,* $\bar{\pi}^N = 1$ *in the finite (multiplicative) group* $(F[x]/(q))^\times$ *of invertible elements in* $F[x]/(q)$*).*

Proof. The sequence $\mathbf{a}$ is the output sequence of an AFSR with connection element q. If $u(x) \in F[x]$ represents a given state of the AFSR, then the succeeding state corresponds to the polynomial $u'(x)$ of equation (15.6) from which we see that

$$\begin{aligned} \deg(u') &\leq \max\{\deg(u), \deg(a_0) + \deg(q)\} - \deg(\pi) \\ &\leq \max\{\deg(u), \deg(\pi) - 1 + \deg(q)\} - \deg(\pi)\} \\ &= \max\{\deg(u) - \deg(\pi), \deg(q) - 1\}. \end{aligned}$$

This shows that the degree of $u(x)$ drops monotonically until it enters the range $\deg(u) < \deg(q)$ and it stays within that range forever after. So the sequence $\mathbf{a}$ is eventually periodic. In such a periodic state, the memory z must satisfy $\deg(z) < \deg(\pi)$, for suppose $\deg(z) \geq \deg(\pi)$. Let $e = \deg(\pi)$. Then the last term in equation (15.4) has degree $\geq (m+1)e$ while the double sum has degree $\leq (e-1) + (e-1) + (m-1)e = (m+1)e - 2$. Consequently the leading terms of these two polynomials cannot cancel, hence $\deg(u) = \deg(z\pi^m) \geq (m+1)e$. But this contradicts the fact that

$$\deg(u) < \deg(q) \leq \deg(q_m \pi^m) \leq e - 1 + me = (m+1)e - 1.$$

Now suppose that $\mathbf{a}$ is strictly periodic, say with period T. Summing the geometric series,

$$\frac{u}{q} = (\sum_{i=0}^{T-1} a_i \pi^i) \sum_{i=1}^{\infty} \pi^{Ti} = \frac{\sum_{i=0}^{T-1} a_i \pi^i}{1 - \pi^T}.$$

The degree of the numerator in this last expression is strictly less than Te, the degree of the denominator. Thus the degree of u is less than the degree of q. Moreover, the equation

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

implies that $\pi^T \equiv 1 \pmod{q}$, so the multiplicative order N of π divides the period T of $\mathbf{a}$.

Conversely, suppose that $\deg(u) < \deg(q)$. Let N denote the multiplicative order of π modulo q, so $1 - \pi^N = sq$ for some polynomial s. It follows that

$$\frac{u}{q} = \frac{su}{1 - \pi^N},$$

and $\deg(su) < Ne$. Thus we can write

$$su = \sum_{i=0}^{N-1} b_i \pi^i$$

with $b_i \in S$. It follows that $a_j = b_{j \bmod N}$ for all j, so $\mathbf{a}$ is strictly periodic, of period N. In particular, the minimal period of $\mathbf{a}$ divides N. □

15.1.b Algebraic model

As in the previous section we consider an AFSR based on polynomials $\pi(x), q(x) \in F[x]$ which are relatively prime to $q = \sum_{i=0}^{m} q_i \pi^i$ and $\deg(q_i) < \deg(\pi)$. If $u \in F[x]$ denote by $\bar{u} \in F[x]/(q)$ and $\tilde{u} \in F[x]/(\pi)$ its image in these two quotient rings. Let Σ denote the collection of all (strictly) periodic states of the AFSR. Let

$$\Delta = \{u(x) \in F[x] : \deg(u) < \deg(q)\}.$$

Then Δ is a *complete set of representatives* for the elements of $F[x]/(q)$. Proposition 15.1.4 says that the correspondence between states and polynomials induces a one to one correspondence

$$\Sigma \leftrightarrow \Delta \leftrightarrow F[x]/(q).$$

In the language of Section 8.5, the sets V_q and Δ coincide, and Δ satisfies the conditions (1) and (2) of Theorem 8.5.1. Define $\Phi : F[x]/(q) \to F[x]/(\pi)$ to be the composition around the following diagram.

$$\begin{array}{ccc} \Delta & \subset & F[x] \\ \cong\downarrow & & \downarrow \\ F[x]/(q) & \xrightarrow[\Phi]{} & F[x]/(\pi) \end{array}$$

Because the set Δ is closed under addition and multiplication by elements of F, the mapping Φ is a F-linear homomorphism ($\Phi(a\bar{u}_1 + b\bar{u}_2) = a\Phi(\bar{u}_1) + b\Phi(\bar{u}_2)$ for all $a, b \in F$ and all $u_1, u_2 \in$

$F[x]/(q)$) although it is not a multiplicative homomorphism. By abuse of notation we write $\Phi(\bar{u}) = \bar{u} \pmod{\pi}$ for any $\bar{u} \in F[x]/(q)$. Finally, define $\phi : F[x]/(q) \to F[x]/(\pi)$ by

$$\phi(\bar{u}) = \tilde{q}^{-1}\Phi(\bar{u}) = \tilde{q}_0^{-1}\Phi(\bar{u}).$$

Then ϕ is also F-linear.

Equation (15.6) says that the polynomial $u \in \Delta$ corresponding to a given state and the polynomial $u' \in \Delta$ corresponding to the succeeding state are related by: $\pi u' = u - a_0 q$. Reducing this equation modulo π and modulo q gives

$$\begin{aligned} \tilde{a}_0 &= \tilde{q}^{-1}\tilde{u} \in F[x]/(\pi) \\ \bar{u}' &= \bar{\pi}^{-1}\bar{u} \in F[x]/(q). \end{aligned}$$

In other words, under the correspondence between states and polynomials, we find that

- the change of state operation $\tau : \Sigma \to \Sigma$ corresponds to the mapping $\sigma : F[x]/(q) \to F[x]/(q)$ defined by $\sigma(\bar{u}) = \bar{\pi}^{-1}\bar{u}$, and
- the output mapping is given by $\phi(\bar{u}) = \tilde{q}^{-1}\bar{u} \pmod{\pi}$.

Thus we obtain an "algebraic model" for the AFSR, illustrated in Figure 15.1.

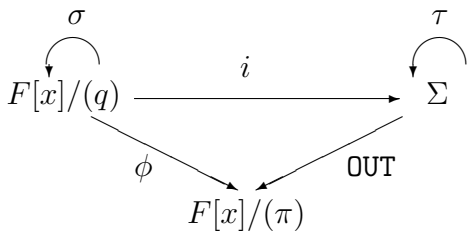


Figure 15.1: Algebraic model for AFSR

Here, $\texttt{OUT} : \Sigma \to F[x]/(\pi)$ denotes the output function which assigns to a state $(a_0, a_1, \cdots, a_{m-1}; z_{m-1})$ the contents a_0 of the rightmost cell. The mappings τ, σ, ϕ are described above. Finally the mapping $i : F[x]/(q) \to \Sigma$ is the correspondence given by equation (15.4). We immediately obtain the following consequence.

Corollary 15.1.5. *The coefficient a_i in the output sequence $\mathbf{a} = \mathbf{seq}_\pi(u/q)$ of the AFSR whose initial loading corresponds to $u(x) \in \Delta$ is given by*

$$\tilde{a}_i = \phi(\bar{u}\bar{\pi}^{-i}) = \tilde{q}_0^{-1}(\bar{u}\bar{\pi}^{-i} \pmod{q}) \pmod{\pi} \in F[x]/(\pi). \tag{15.7}$$

By abuse of notation this formula is sometimes written $a_i = q_0^{-1}(u\pi^{-i} \pmod{q}) \pmod{\pi}$.

Corollary 15.1.6. *Given any initial loading $(a_0, a_1, \cdots, a_{m-1})$ of the cells of the AFSR, there exists a value $z = z_{m-1}$ of the "memory" such that the output sequence is strictly periodic. If $q_m \in F$ (that is, if $\deg(q_m) = 0$) then this value of z is unique. In this case, $z = 0$ if and only if*

$$\deg\left(\sum_{i=0}^{m-1} a_i q_{m-i-1}\right) \leq \deg(\pi) - 1. \tag{15.8}$$

Proof. Given the initial loading $(a_0, a_1, \cdots, a_{m-1})$ let us consider the effects of different values z_{m-1} of the memory on the degree of the polynomial $u(x)$ in equation (15.4). Let $H(x)$ denote the double sum in equation (15.4) and set $e = \deg(\pi)$. By the division theorem for polynomials, there exists a unique polynomial $z \in F[x]$ such that

$$H(x) = z(x)\pi^m + J(x)$$

with $\deg(J) < \deg(\pi^m) = me \leq \deg(q)$ since $q_m \neq 0$. Taking this $z = z_{m-1}$ for the memory gives a state of the AFSR whose output sequence is the π-adic expansion of u/q, where $u = H - z\pi^m = J$ has degree $< \deg(q)$. So by Proposition 15.1.4 the output sequence is strictly periodic. This proves that such a z always exists.

Now suppose q_m has degree 0. Then $\deg(q) = em$. We wish to prove that z is unique. Given the initial loading $(a_0, a_1, \cdots, a_{m-1})$ suppose there are two values, $z \neq z'$ for the memory such that the output sequence is strictly periodic. Let u, u' be the corresponding polynomials from equation (15.4). Then $\deg(u), \deg(u') < em$ by Proposition 15.1.4. However $u - u' = (z' - z)\pi^m$ which has degree $\geq em$ and this is a contradiction.

Finally, suppose equation (15.8) holds. The terms of highest degree in the double sum H of equation (15.4) are

$$\pi^{m-1} \sum_{i=0}^{m-1} a_i q_{m-i-1}$$

which has degree

$$em - e + \deg\left(\sum_{i=0}^{m-1} a_i q_{m-i-1}\right) < em = \deg(q)$$

by assumption. However the term $z\pi^m$ has degree em. So any nonzero value for z will result in $\deg(u) \geq \deg(g)$ and the output sequence will fail to be strictly periodic, by Proposition 15.1.4. The converse is similar. □

15.1.c Long function field sequences

In this section we will show that there exist function field sequences **a** that are (punctured) de Bruijn sequences. As in Section 15.1, fix a finite (Galois) field F and a polynomial $\pi(x) \in F[x]$.

Definition 15.1.7. *A function field sequence* $\mathbf{a} = (a_0, a_1, \cdots)$ *(with* $a_i \in F[x]/(\pi)$*) with connection element* $q(x) \in F[x]$ *is a* (π, q)-adic ℓ-sequence *if it is a (strictly periodic) punctured de Bruijn sequence with period equal to* $|F[x]/(q)| - 1 = |F|^{\deg(q)} - 1$.

Proposition 15.1.8. *Let* $\pi, q \in F[x]$ *and let* $\mathbf{a} = \mathbf{seq}_\pi(u/q)$ *be the output sequence of the AFSR based on* π *with connection element*

$$q = q_0 + q_1\pi + \cdots + q_m\pi^m \tag{15.9}$$

where $\deg(q_i) < \deg(\pi)$*, and with initial loading corresponding to* $u(x) \in F[x]$*, cf. equation (15.4). Then* $\mathbf{a}$ *is a* (π, q)*-adic* ℓ*-sequence if and only if the following four conditions are satisfied.*

1. *The polynomial* $q(x) \in F[x]$ *is irreducible (so* $F[x]/(q)$ *is a field).*
2. *The element* $\bar{\pi} = \pi \pmod q \in F[x]/(q)$ *is a primitive element (which is not the same as being a primitive polynomial in* $F[x]$*).*
3. $\deg(q_m) = 0$ *or equivalently,* $q_m \in F$ *or equivalently,* $\deg(q) = m\deg(\pi)$
4. $u \in \Delta$*, that is,* $\deg(u) < \deg(q)$.

Assuming these conditions, then for any other nonzero $u' \in \Delta$ *the sequence* $\mathbf{a}' = \mathbf{seq}_\pi(u'/q)$ *is a shift of the sequence* $\mathbf{a}$.

Proof. First, suppose $\mathbf{a}$ is a (π, q)-adic ℓ-sequence. Then $\mathbf{a}$ is (strictly) periodic so condition (4) holds. According to Proposition 15.1.4 the period of the sequence $\mathbf{a}$ is the multiplicative order of π modulo q. This will be maximal if $F[x]/(q)$ is a field and if π is a primitive element in this field, whence conditions (1) and (2). To obtain a punctured de Bruijn sequence we will also need to have a sequence of period $|F[x]/(\pi)|^m - 1$ for some m, which implies that $|F[x]/(q)| = |F[x]/(\pi)|^m$ so $m\deg(\pi) = \deg(q)$. Consequently the π-adic expansion (15.9) of q has m terms and $\deg(q_m) = 0$, which is condition (3).

Conversely suppose π, q, and u are chosen so as to satisfy the above conditions. Then the resulting sequence $\mathbf{a}$ is periodic and has period $|F[x]/(q)| - 1 = |F|^{em} - 1$ where $e = \deg(\pi)$. We need to show that it is a punctured de Bruijn sequence. Suppose a block $b = (b_0, b_1, \cdots, b_{m-1})$ of length m occurs in $\mathbf{a}$ after some number of iterations. Consider the state of the AFSR at this point. The values $b_0, b_1, \cdots, b_{m-1}$ are the contents of the registers. By Corollary 15.1.6 there is a unique value z for the memory such that the output of the AFSR with this initial loading b and initial memory z is a strictly periodic sequence. Since the sequence is, in fact, periodic from this point, the memory must have this value z. It follows that the block b can occur at most once in any period of $\mathbf{a}$ — otherwise the sequence would repeat upon the next occurrence of b, and its period would be less than $|F|^{em} - 1$. However, there are $|F|^{em}$ possible blocks b, and the block $(0, 0, \cdots, 0)$ cannot occur in $\mathbf{a}$ (otherwise $\mathbf{a}$ would consist only of zeroes). Consequently every nonzero block b of length m occurs exactly once in a single period of $\mathbf{a}$.

Finally, each nonzero polynomial $u' \in \Delta$ corresponds to a (periodic) initial loading of the shift register. But Δ contains $|F|^m - 1$ elements, which coincides with the period of $\mathbf{a}$. Hence every nonzero initial stae gives rise to a shift of the same sequence. □

Remark In Section 15.1.b we described an algebraic model for the AFSR in terms of the mapping $\phi : F[x]/(q) \to F[x]/(\pi)$ which is the composition

$$F[x]/(q) \xrightarrow{\cong} \Delta \xrightarrow{(\mathrm{mod}\ \pi)} F[x]/(\pi)$$

Suppose, as above, that $q(x)$ is irreducible of degree $m\deg(\pi)$, and that π is primitive modulo q. Suppose the field $F = \mathbb{F}_{p^d}$ has characteristic p. We may consider $F[x]/(\pi)$ to be a vector space V over $\mathbb{F}_p$ and $F[x]/(q)$ to be a field L. Then $\phi : L \to V$ is linear over $F \supset \mathbb{F}_p$ (cf. Section 15.1.b); it is surjective; and it has the *kernel property* of Section 12.4 with respect to the primitive element π. For if $u(x) \in \Delta$ (that is, if $\deg(u) < \deg(q)$) and if $u \in \bigcap_{i=0}^{m-1} \pi^i \mathrm{Ker}(\phi)$ then $\phi(\pi^{-i}u) = 0$ for $0 \le i \le m-1$ which implies that the output sequence of the AFSR (whose initial loading corresponds to u) begins with m zeroes. But u corresponds to a periodic state, so the output is always zero, hence $u = 0$. It then follows from Theorem 12.4.1 that the output sequence is a punctured de Bruijn sequence, giving an alternate proof of Proposition 15.1.8.

Theorem 15.1.9. *Every (π, q)-adic ℓ-sequence $\mathbf{a}$ satisfies the shift-and-add property and has ideal autocorrelations.*

Proof. The sequence $\mathbf{a}$ is the coefficient sequence $\mathbf{a} = \mathbf{seq}_\pi(u/q)$ of the π-adic expansion of some fraction $u(x)/q(x)$ where $\deg(u) < \deg(q)$. According to the preceding proposition, for any integer τ, the τ-shift $\mathbf{a}_\tau$ of $\mathbf{a}$ is the coefficient sequence for some fraction $u'(x)/q(x)$ where $\deg(u') < \deg(q)$. To verify the shift-and-add property, let $c, d \in F$ and set $v(x) = cu(x) + du'(x)$. Since $\deg(v) < \deg(q)$ the coefficient sequence for $v(x)/q(x)$ (which is $c\mathbf{a} + d\mathbf{a}_\tau$) is a shift of the sequence $\mathbf{a}$. Therefore $\mathbf{a}$ satisfies the shift-and-add property. Since $\mathbf{a}$ is a de Bruijn sequence, it is also balanced. It follows from Proposition 12.1.3 that $\mathbf{a}$ has ideal autocorrelations. □

15.1.d Relation with m-sequences

The (π, q)-adic ℓ-sequences shares many of the properties of m-sequences. In this section we show that, except in trivial cases, such a sequence $\mathbf{a}$ is never an m-sequence, and we give sufficient conditions to guarantee that $\mathbf{a}$ cannot be obtained from an m-sequence by a linear change of variable.

Recall that the field $F = \mathbb{F}_{p^d}$ was fixed at the beginning of this chapter. As in the previous section we fix $\pi(x), q(x) \in F[x]$ relatively prime, of degrees e and $g = em$ respectively, with q irreducible and with π primitive modulo q. Let $\mathbf{a} = a_1, a_2, \cdots$ denote the resulting (π, q)-adic

ℓ-sequence. It has period $|F|^{em}-1$. Its entries are in the ring $F[x]/(\pi)$, which may be considered as a vector space over F with p^{de} elements. Finally, write $q=\sum_{i=0}^{m} q_i\pi^i$ (with $\deg(q_m)=0$). The sequence $\mathbf{a}$ is, up to a shift, given by $a_i=\phi(\pi^{-i})$ where $\phi: F[x]/(q)\to F[x]/(\pi)$ is the F-linear mapping described in Section 15.1.b.

If $\pi(x)=x$ then $F[x]/(\pi)=F$, the mapping $\phi: F[x]/(q)\to F$ is a F-linear surjective mapping, and the resulting sequence $\mathbf{a}$ is an m-sequence (over F). Every m-sequence occurs this way, as a (π,q)-adic ℓ-sequence by taking $q(x)$ to be the corresponding primitive polynomial, and $\pi(x)=x$. The following theorem is proven in [64].

Theorem 15.1.10. *Let $\hat{F}=\mathbb{F}_{p^{de}}$ be the field with $|\hat{F}|=|F[x]/(\pi)|$ and let $\hat{\mathbf{a}}=(\hat{a}_0,\hat{a}_1,\cdots)$ be an m-sequence with period $p^{dem}=|F[x]/(q)|-1$ of elements $\hat{a}_i\in\hat{F}$. If $\deg(\pi)\geq 2$ then there does not exist any set theoretic mapping $\psi: F[x]/(\pi)\to\hat{F}$ such that $\psi(\mathbf{a})=\hat{\mathbf{a}}$.*

In other words, the (π,q)-adic ℓ-sequence $\mathbf{a}$ and the m-sequence $\hat{\mathbf{a}}$ are not isomorphic, even after a possible left shift. They are shift distinct even after an isomorphism (cf. Section 5.1.b).

In [56], Gong, Di Porto, and Wolfowicz constructed pseudo-noise sequences by applying an invertible $\mathbb{F}_p$ linear map to each element in an m-sequence over $\mathbb{F}_{p^f}$. In particular, each such sequence is *isomorphic* to an m-sequence. According to Theorem 15.1.10, if $\deg(\pi)\geq 2$ then no (π,q)-adic ℓ-sequence can be obtained by the method of Gong et al.

15.1.e Existence

It is not immediately apparent that (π,q)-adic ℓ-sequences that are not m-sequences are abundant. In order to find such sequences we fix the field $F=\mathbb{F}_{p^d}$ and search for a pair of polynomials $\pi,q\in F[x]$ such that q is irreducible and π is primitive modulo q. In order to get a de Bruijn sequence we will also require that $g=\deg(q)$ is a multiple of $e=\deg(\pi)$. To guarantee that the resulting sequence is not an m-sequence we also require that $q=\sum_{i=0}^{m} q_i\pi^i$ with at least one q_i of degree greater than 0.

First recall the theorem of Pappalardi and Shparlinski [155]: Let $\overline{F}$ be an algebraic closure of F. Suppose π is not an mth power of a function $h\in\overline{F}[x]$, for any m which divides $|F|^g-1$. Then the number $N(\pi,F,g)$ of irreducible polynomials $q\in F[x]$ of degree g for which π is primitive satisfies

$$\left|N(\pi,F,g)-\frac{\phi(M-1)}{g}\right|\leq 3eg^{-1}2^{\nu(M-1)}\sqrt{M}$$

where $M=|F|^g$, ϕ denotes Euler's ϕ function and $\nu(n)$ denotes the number of distinct prime divisors of n. This implies the existence of many pairs (π,q) such that π is primitive mod q. For example, if $F=\mathbb{F}_2$ and $g=13$ it says that for any $e\leq 42$ there exists π with $\deg(\pi)=e$ and π primitive mod q. If $g\geq 75$, then for every divisor e of g there exist polynomials π of degree e that are primitive mod g.

In fact, primitive polynomial pairs (π, q) are considerably more abundant than the above estimates predict. By computer search we have found the following for $F = \mathbb{F}_2$: Fix $g \leq 22$. Suppose $\pi \in F[x]$ is a polynomial of degree $e < g$ and suppose π is not a power of a polynomial $\pi \neq h^n$ where n divides g. Then there exists an irreducible polynomial q of degree g such that π is primitive mod q unless $\pi = x^4 + x$ and $g = 6$. In other words, there is a single unacceptable pair (π, g) in this range! (In this case, the above estimate says $|N(\pi, F, g) - 6| \leq 64$ so $N = 0$ is, indeed, a possibility.)

A class of examples which may be easily analyzed is the following. Let $q(x) \in F[x]$ be a primitive polynomial of degree $g = me$. Let $\pi(x) = x^e$. Then π is primitive modulo q if and only if e is relatively prime to $|F[x]/(q)| - 1 = |F|^g - 1$. This is satisfied, for example, if g is relatively prime to $|F|^g - 1$. For example, if $F = \mathbb{F}_2$ and $\pi(x) = x^2$ we may take q to be any primitive polynomial of even degree. If such a q contains any terms of odd degree then some q_i has positive degree, so the resulting (π, q)-adic ℓ-sequence $\mathbf{a}$ is not an m-sequence. If $F = \mathbb{F}_2$ and $\pi(x) = x^3$ we may take q to be any primitive polynomial whose degree is an odd multiple of 3. If such a q contains any terms of degree not divisible by 3, then some q_i has positive degree, so the sequence $\mathbf{a}$ is not an m-sequence.

15.1.f Examples

In this section we let $p = 2$ and $d = 1$. If $\deg(\pi) = 1$, then we obtain m-sequences. The case $\pi(x) = x$ is the standard analysis of m-sequences by power series. The case $\pi(x) = x + 1$ is equivalent by a change of basis.

Suppose that π has degree 2. Then for any choice of q we obtain sequences with elements in $H = \mathbb{F}_2[x]/(\pi) = \{0, 1, x, x+1\}$. If $\pi(x) = x^2+x+1$, which is irreducible over $\mathbb{F}_2$, we have $H = \mathbb{F}_4$, but for all other πs of degree two the ring H is not a field. If we let $\pi(x) = x^2 + x + 1$ and use the connection element $q(x) = x^4 + x^3 + 1 = \pi^2 + x\pi + x$, then it can be shown that π is primitive modulo q and one period of the (π, q)-adic ℓ-sequence $\mathbf{a}$ we obtain is (possibly a left shift of)

$$1, 1, x, 0, 1, 0, x, x, x + 1, x + 1, 1, x + 1, 0, x + 1, x. \tag{15.10}$$

All other (π, q)-adic ℓ-sequences obtained by different choices of π of degree 2 and q of degree 4 with π primitive modulo q are obtained from the sequence (15.10) by some combination of shifts, reversals, and permutations of the alphabet $\{0, 1, x, x + 1\}$.

However, the sequence with one period equal to

$$1, 1, x, 1, 0, x + 1, x + 1, 1, x + 1, 0, x, x, x + 1, x, 0$$

is an m-sequence over $\mathbb{F}_4$, and all other m-sequences of span 2 over $\mathbb{F}_4$ are obtained from this sequence by some combination of shifts, reversals, and switching x and $x + 1$. This illustrates the

fact that the new set of sequences is disjoint from the set of m-sequences. In fact there is no set theoretic isomorphism $\phi : \mathbb{F}_4 \to \mathbb{F}_4$ so that $\phi(\mathbf{a})$ is an m-sequence, for the sequence $\mathbf{a}$ contains a string $x, 0, 1, 0, x$ and there is no string of the form a, b, c, b, a in any of these m-sequences.

15.2 Global function fields

15.2.a ℓ-Sequences and randomness

Recall from Section 8.3.c the setup for an AFSR based on a global function field. In this case the ring R is of the form $R = F[x_1, \cdots, x_n]/I$ where $F = \mathbb{F}_r$ is a finite field with $r = p^h$ elements, p prime and I is an ideal such that R has transcendence degree one over $\mathbb{F}_r$. Fix $\pi \in R$ and assume that $K = R/(\pi)$ is finite.

An AFSR based on (R, π, S) involves a choice $S \subset R$ of a complete set of representatives for $R/(\pi)$. Such a choice is also needed in order to obtain unique π-adic expansions of fractions

$$\frac{u}{q} = a_0 + a_1\pi + a_2\pi^2 + \cdots$$

where $a_i \in S$, $u, q \in R$ and, q and π are coprime. The coefficient sequence $a_0, a_1, \cdots$ is denoted $\mathbf{seq}_\pi(u/q)$. It is the output sequence of an AFSR with connection element q. Fix such an element $q \in R$ and let V_q denote the set of $u \in R$ such that $\mathbf{seq}_\pi(u/q)$ is (strictly) periodic.

In Section 8.3.c we considered three hypotheses on a complete set of representatives S for $R/(\pi)$ and connection element q which we now recall.

H1: The set S is closed under addition and contains F.
H2: Every $v \in R$ is a finite linear combination $v = v_0 + v_1\pi + \cdots + v_\ell\pi^\ell$ with $v_i \in S$.
H3: There exists k so that elements of V_q are distinct modulo π^k

A necessary condition for the sequence $\mathbf{a} = \mathbf{seq}_\pi(u/q)$ to have maximal length is that π should be *primitive* (mod q) and in particular that $R/(q)$ is a field. The rings $R/(\pi)$ and $R/(q)$ are also vector spaces over F of some dimension, say, e and g respectively. If π is primitive (mod q) then the period of $\mathbf{a}$ is $r^g - 1$. If $\mathbf{a}$ is also a punctured de Bruijn sequence of some span, m, then each nonzero sequence of length m occurs once in each period of $\mathbf{a}$, hence the period of a is $(r^e)^m - 1$. It follows that $g = em$.

Definition 15.2.1. *Let $\pi, q \in R$ be coprime, $q = \sum_{i=1}^{m} q_i\pi^i - q_0$ with $q_m \neq 0$, $|R/(\pi)| = r^e$, $S \subseteq R$, and suppose hypotheses H1 and H2 hold. Let $|R/(q)| = r^g$, and let $R/(q)$ be a field. Let $u \in R$ ($u \neq 0$) and let $\mathbf{a} = (a_0, a_1, \cdots) = \mathbf{seq}_\pi(u/q)$ be the nonzero periodic output sequence from an AFSR with connection element q. Then $\mathbf{a}$ is periodic with minimal period $r^g - 1$ if and only if π is primitive modulo q, in which case $\mathbf{a}$ is called a* (π, q)-adic ℓ-sequence.

Theorem 15.2.2. *Suppose* $\mathbf{a}$ *is a* (π, q)*-adic* ℓ*-sequence, hypotheses H1, H2, and H3 hold, and* $q_m \in F$*. Then the following hold.*

1. $\mathbf{a}$ *is a punctured de Bruijn sequence. Thus the number of occurrences of a sequence* $\mathbf{b}$ *of length* $t \leq m$ *in a period of* $\mathbf{a}$ *is* $r^{e(m-t)}$ *if* $\mathbf{b} \neq (0, \cdots, 0)$ *and is* $r^{e(m-t)} - 1$ *otherwise (see* Section *11.2.d).*
2. $\mathbf{a}$ *has the shift and add property (or SAA — see Chapter 12).*
3. $\mathbf{a}$ *is balanced.*
4. $\mathbf{a}$ *has the run property (see* Section *11.2.b).*
5. $\mathbf{a}$ *has ideal autocorrelations (see* Section *11.3.a).*

Proof. Properties (3), (4) and (5) follow from properties (1) and (2).

It suffice to prove property (1) when $t = m$. Since $q_m \in \mathbb{F}_{r^e}$ and Hypothesis H2 holds, we have

$$|R/(q)| = |S|^m = r^{em}.$$

Thus $\mathbf{a}$ has period $r^{em} - 1$. The various shifts of $\mathbf{a}$ plus the all-zero sequence give r^{em} periodic sequences corresponding to elements u/q. Thus,

$$|V_q| \geq r^{em}.$$

We have seen in the proof of Theorem 8.5.6 that the elements of V_q are distinct modulo q, so

$$|V_q| \leq r^{em}.$$

Thus,

$$|V_q| = r^{em} = |R/(\pi^m)|.$$

By Hypothesis H3, the elements of V_q are distinct modulo π^m, so V_q is a complete set of representatives modulo π^m.

The set of occurrences in $\mathbf{a}$ of a block $\mathbf{b}$ of m elements corresponds to the set of shifts of $\mathbf{a}$ that begin with $\mathbf{b}$. By the above, every nonzero u/q with $u \in V_q$ occurs as a shift of $\mathbf{a}$, so the set of occurrences in $\mathbf{a}$ of $\mathbf{b}$ corresponds to the set of nonzero u/q, $u \in V_q$, that begin with $\mathbf{b}$. We claim that the u/q are distinct modulo π^m, so that each nonzero $\mathbf{b}$ occurs once. Suppose not, so that $u/q \equiv v/q \pmod{\pi^m}$ for some $u \neq v \in V_q$. Then $u \equiv v \pmod{\pi^m}$ since q is invertible modulo r, and hence also modulo π^m. But by Hypothesis H3 the elements of V_q are distinct modulo π^m. It follows that $\mathbf{a}$ is a punctured de Bruijn sequence.

Furthermore, if $u, v \in V_q$, then $u + v \in V_q$. The shifts of $\mathbf{a}$ account for all the u/q with $u \neq 0 \in V_q$ so the SAA property follows. □

15.3 Exercises

1. Let $R = \mathbb{F}_2[x]$, $\pi = x^2$, and $S = \{a + bx : a, b \in \mathbb{F}_2\}$. Consider function field AFSRs in this environment. The sequences generated by these AFSRs have elements in $K = R/(\pi) = \{a + bx : a, b \in \mathbb{F}_2\}$, with multiplication in K defined by $(a + bx)(c + dx) = ac + (ad + bc)x$. Let $q = \pi^2 + x\pi + 1 = x^4 + x^3 + 1$ and let $\mathbf{a}$ be a sequence with minimal connection element q.

   a. Prove that $\mathbf{a}$ is an ℓ-sequence.
   b. Prove that there is no one to one correspondence between K and $\mathbb{F}_4$ that maps $\mathbf{a}$ into an m-sequence, into the sequence in equation 15.10, or into a decimation of the sequence in equation 15.10.

2. Let $R = \mathbb{F}_2[x, y]/(x^3 + y^2)$ (the *cuspidal cubic*). Let $\pi = x$ and $q = x^2y + x + 1 = y\pi^2 + \pi + 1$.

   a. Show that $R/(\pi) = \{a + by : a, b \in \mathbb{F}_2\}$, so $|R/(\pi)| = 4$ and we can take $S = \{a + by : a, b \in \mathbb{F}_2\} \subseteq R$.
   b. Determine $|R/(q)|$. (Hint: learn about Gröbner bases.)
   c. What is the period of a sequence with minimal connection element q? Is this sequence an ℓ-sequence?

Chapter 16 Maximal Period FCSR Sequences

A maximal period LFSR sequence is called an m-sequence. By analogy, we refer to a maximal period FCSR sequence as an ℓ-sequence. These sequences exhibit many of the same desirable statistical properties as m-sequences. As we see in Chapter 19, like m-sequences, they are cryptographically weak.

16.1 ℓ-Sequences

By Corollary 5.4.5, the maximum possible period for an N-ary FCSR with connection integer q is $T = q - 1$. This period is attained if and only if q is prime and N is a primitive root modulo q. In this case, for any initial loading of the register, the output sequence will either degenerate into all 0s or all $(N-1)$s, or else it will eventually drop into the big set of periodic states (see Section 7.3). An FCSR sequence with connection integer q corresponds to a cyclic subset of the group of invertible elements, or units, $\mathbb{Z}/(q)^\times \subset \mathbb{Z}/(q)$ (consisting of numbers relatively prime to q). So a necessary condition for the existence of a maximal length FCSR sequence is that $\mathbb{Z}/(q)^\times$ should be a cyclic group. Recall from Section 2.2.d that *the multiplicative group $\mathbb{Z}/(q)^\times$ is cyclic if and only if $q = p^m$, $2p^m$, 2, or 4*, where p is an odd prime.

If q is an odd prime power, then reduction modulo q gives an isomorphism $\mathbb{Z}/(2q)^\times \cong \mathbb{Z}/(q)^\times$ so we may restrict to the case that q is a power of an odd prime.

Definition 16.1.1. *An ℓ-sequence (or "long" sequence) is a periodic sequence* $\mathbf{a}$ *which is obtained from a FCSR with connection integer q such that $q = p^r$ is a power of an odd prime and the period of* $\mathbf{a}$ *is $\phi(q)$.*

According to Theorem 7.4.2, this occurs if and only if N is a generator of $\mathbb{Z}/(q)^\times$, that is, N is a *primitive root* modulo q. According to Theorem 7.4.1, if $\mathbf{a} = (a_0, a_1, \cdots)$, then there exists an integer h with $0 < h < q$ so that the equality

$$\frac{-h}{q} = a_0 + a_1 N + a_2 N^2 + \cdots$$

holds in $\mathbb{Z}_N$ (the N-adic numbers). That is, the sequence $\mathbf{a} = \mathbf{seq}_N(-h/q)$ is the coefficient sequence of the N-adic expansion of $-h/q$. It is explicitly given by

$$a_i = N^{-i} h \pmod{q} \pmod{N}.$$

Every N-ary ℓ-sequence with connection integer q is a left shift of this one. Since N is primitive (mod q), the elements $N^{-i}h \in \mathbb{Z}/(q)$ account for all the invertible elements in $\mathbb{Z}/(q)$. This gives a one to one correspondence between N-ary ℓ-sequences with connection integer q and elements of $\mathbb{Z}/(q)^{\times}$.

It is interesting to note that if q is the connection integer of a binary ℓ-sequence and $q \neq 3$, then the integer sequence of memory values for the FCSR with connection integer q has the same period as the ℓ-sequence. This was proved by Tian and Qi [180].

Arithmetic codes. By Theorem 5.4.4 and Corollary 5.4.5, the ℓ-sequence $\mathbf{a}$ is the reverse of the "decimal" expansion (with base N),

$$\frac{h}{q} = b_0 N^{-1} + b_1 N^{-2} + b_3 N^{-3} + \cdots$$

of the fraction h/q (see, for example, [111] Section 4.1, ex. 31). This binary expansion is called a $1/q$-*sequence* in [13], any single period of which is a codeword in the Barrows-Mandelbaum arithmetic code [5, 126]. They have been studied since the time of Gauss [49].

Shift and add. In Theorem 12.6.2 it was shown that an N-ary sequence satisfies the arithmetic shift and add property if and only if it is an ℓ-sequence with prime connection integer q, for which N is primitive.

Existence. The question of the existence, even of binary ℓ-sequences, with arbitrarily large period is subtle. Heilbronn (revising Artin's conjecture) conjectured, and Hooley [80] proved, that if an extension of the Riemann hypothesis to the Dedekind zeta function over certain Galois fields is true, then the number $N(n)$ of primes $q < n$ such that 2 is primitive mod q is:

$$N(n) = A \cdot \frac{n}{\ln(n)} + O\left(\frac{n \ln\ln(n)}{\ln^2(n)}\right),$$

where A (= .3739558136 to ten decimals) is Artin's constant. In other words, 37.4% of all prime numbers are conjectured to have this property. There are efficient techniques for finding large primes q for which N is a primitive root (see [27]) which are currently implemented in popular software systems such as Maple, Mathematica, and Pari. For example, an FCSR based on the prime number

$$q = 2^{128} + 2^5 + 2^4 + 2^2 - 1$$

needs only 2 bits of extra memory and has maximal period $T = q - 1$.

To construct an N-ary ℓ-sequences based on a connection integer $q = p^t$ (p an odd prime) it is necessary to find such numbers q for which N is a primitive root. The search for such pairs is not as difficult as it may seem at first glance because, according to Proposition 2.2.12, the number N is primitive mod $q = p^t$ if and only if it is primitive mod p^2.

One can ask about the abundance of primes p for which $N = 2$ is a primitive root modulo p^2. All of the primes p listed in Table 16.1 have this property. In fact Hardy and Wright pointed out that the condition that p^2 divides $2^{p-1} - 1$ holds for only two primes p less than $3 \cdot 10^7$ [75, p. 73], and by computer search E. Bombieri has extended this limit to $2 \cdot 10^{10}$ [15]. (The two primes are 1093 and 3511.) In both cases 2 is not primitive modulo p. Thus for a large number of primes, we need only check the primitivity of N modulo p. In fact, it is not known whether there are *any* primes p such that 2 is primitive modulo p but not modulo p^2, though there is no compelling reason to believe there are no such primes.

Other long sequences. There is another case in which large periods can be obtained. Suppose that q is a prime number such that $q = 2p + 1$ with p a prime number. Then sequences generated by an FCSR with connection integer q will have period $T \geq (q-1)/2$, which is half the maximum possible period. (This is because Fermat's congruence states that, if x is not a multiple of q, then $x^{q-1} \equiv 1 \pmod{q}$, so $\mathrm{ord}_q(2)$ divides $q-1 = 2p$ and hence is equal either to 2, which is impossible; to p; or to $q-1$.) It is apparently easier to check whether $(q-1)/2$ is prime than it is to determine whether 2 is a primitive root modulo q. It was conjectured by Hardy and Littlewood [74], and is widely believed by number theorists, that the number of primes $P(n)$ less than n of the form $2p + 1$, p prime, is asymptotically given by

$$P(n) \sim c_2 \cdot \frac{n}{\ln^2(n)},$$

where c_2 (= .0330080908 to ten decimal places) is a constant.

16.2 Distributional properties of ℓ-sequences

In this section we show that ℓ-sequences are close to being de Bruijn sequences: the number of occurrences of any two blocks can differ at most by 2. If the connection integer q is prime then this difference can be at most 1 so the ℓ-sequence is equidistributed to all orders.

Theorem 16.2.1. *Let* $\mathbf{a}$ *be an* N*-ary* ℓ*-sequence based on a connection integer* $q = p^t$ *with* p *an odd prime. Then the number* $n(\mathbf{b})$ *of occurrences of any block* $\mathbf{b}$ *of size* s *within a single period of* $\mathbf{a}$ *is*

$$\begin{aligned} n_1 \leq n(\mathbf{b}) \leq n_1 + 1 \qquad & \text{if } t = 1 \\ n_1 - n_2 - 1 \leq n(\mathbf{b}) \leq n_1 - n_2 + 1 \qquad & \text{if } t \geq 2 \end{aligned}$$

where

$$n_1 = \left\lfloor \frac{q}{N^s} \right\rfloor = \left\lfloor \frac{p^t}{N^s} \right\rfloor \quad \text{and} \quad n_2 = \left\lfloor \frac{p^{t-1}}{N^s} \right\rfloor.$$

Proof. The set of left shifts of $\mathbf{a}$ corresponds to the set of rational numbers $-u/q$, such that $0 < u < q$ and u is invertible mod q. Thus the number of occurrences of a length s block $\mathbf{b}$ in $\mathbf{a}$ equals the number of u with $0 < u < q$, p not dividing u, and the first s elements in the N-adic expansion of $-u/q$ equal to $\mathbf{b}$. Two rational numbers $-u_1/q$ and $-u_2/q$ have the same first s elements in their N-adic expansions if and only if $-u_1/q \equiv -u_2/q \pmod{N^s}$, which holds if and only if $u_1 \equiv u_2 \pmod{N^s}$ because $\gcd(q, N) = 1$. Thus we want to count the number of u with a given first s elements in their N-adic expansions, $0 < u < q$, and u not divisible by p.

Let $N^r < q < N^{r+1}$. If $s > r$, then there is either no such u or there is one such u, so the result follows. Thus we may assume $s \le r$.

We first count the number of u with the first s elements in their N-adic expansions fixed and $0 < u < q$, ignoring the divisibility condition. If $\mathbf{b} = b_0, \cdots, b_{s-1}$, let $b = \sum_{i=0}^{s-1} b_i N^i$. Let

$$q = \sum_{i=0}^{r} q_i N^i \quad \text{and} \quad q' = \sum_{i=0}^{s-1} q_i N^i.$$

If $b < q'$, then every choice of B with

$$0 \le N^s B \le \sum_{i=s}^{r} q_i N^i = N^s \left\lfloor \frac{q}{N^s} \right\rfloor$$

gives a unique $u = b + N^s B < q$ in the right range. There are $n_1 + 1$ such choices. If $b \ge q'$, then every choice of B with

$$0 \le N^s B < \sum_{i=s}^{r} q_i N^i = N^s \left\lfloor \frac{q}{N^s} \right\rfloor$$

gives a unique $u = b + N^s B < q$ in the right range. There are n_1 such choices. In summary,

$$|\{u \equiv b \pmod{N^s} : 0 \le u < p^t\}| \in \{n_1, n_1 + 1\}.$$

If q is prime this completes the proof.

If $q = p^t$, $t \ge 2$, is a prime power, consider those u for which $0 < u < q$ and p divides u. That is, $u = pv$ for some v, and $0 < v < q/p = p^{t-1}$. As above, $u_1 = pv_1$ and $u_2 = pv_2$ have the same first s elements in their N-adic expansions if and only if the same is true of v_1 and v_2. Therefore, if $r = p^{-1} \pmod{N}$ then the number of u values that are divisible by p is

$$|\{u \equiv b \pmod{N^s} : 0 \le u < p^t\}| - |\{v \equiv rb \pmod{N^s} : 0 \le v < p^{t-1}\}|$$

which is in the set

$$\{n_1, n_1+1\} - \{n_2, n_2+1\} = \{n_1 - n_2 - 1, n_1 - n_2, n_1 - n_2 + 1\}. \qquad \square$$

It is easy to find examples for which all three numbers occur.

Proposition 16.2.2. *If* $\mathbf{a}$ *is an* N*-ary* ℓ*-sequence based on a connection integer* $q = p^t$*, with* p *an odd prime, or if* $\mathbf{a}$ *is the* d*-fold decimation of such an* ℓ*-sequence, where* d *is odd, then*

$$a_{i+(q-1)/2} = q - 1 - a_i.$$

This holds even for AFSRs based on $R = \mathbb{Z}$ and $\pi = N$. That is, we may allow $q \not\equiv -1 \pmod{N}$.

Proof. First consider the case of an N-ary ℓ-sequence $\mathbf{a}$. Its period is $T = \phi(q) = p^t - p^{t-1}$, which is even. Since $N^{\phi(q)} \equiv 1 \pmod{q}$ we have:

$$p^t | (N^T - 1) = (N^{T/2} - 1)(N^{T/2} + 1).$$

The GCD of these two factors is 1 if N is even and is 2 if N is odd. Since p is odd, p^t divides one of the factors. By the primitivity of N modulo q, p^t cannot divide the first factor, hence it divides the second factor. It follows that $N^{T/2} \equiv -1 \pmod{q}$ so also

$$\sigma^{T/2} \equiv -1 \pmod{q}$$

where $\sigma = N^{-1} \pmod{q}$. By Theorem 8.5.1, here is an integer B such that $a_i = -q^{-1}(B \cdot \sigma^i \pmod{q}) \pmod{N}$. Thus

$$a_{i+T/2} = -q^{-1}(-B\sigma^i \pmod{q}) \pmod{N} = -q^{-1}(q - (B\sigma^i \pmod{q})) \pmod{N} = q - 1 - a_i.$$

Now let $\mathbf{b}$ be the d-fold decimation of $\mathbf{a}$, where d is odd. Then

$$b_i = a_{id} = q - 1 - a_{id+T/2} = q - 1 - a_{(i+T/2)d} = q - 1 - b_{i+T/2}. \qquad \square$$

If $\mathbf{a} = (a_0, a_1, \cdots)$ is a periodic N-ary sequence and if $\chi : \mathbb{Z}/(N) \to \mathbb{C}^\times$ is a character, then the *imbalance with respect to* χ was defined in Section 11.2.a to be the sum $Z(\mathbf{a}) = \sum_{i=0}^{T-1} \chi(a_i)$ where T is the period of $\mathbf{a}$. The sequence is *balanced* if $|Z(\mathbf{a})| \leq 1$.

In the special case of binary ℓ-sequences more is true (see also Lemma 19.2.4).

Corollary 16.2.3. *If* $\mathbf{a}$ *is a binary* ℓ*-sequence based on a connection integer* $q = p^t$*, with* p *an odd prime, or if* $\mathbf{a}$ *is the* d*-fold decimation of such an* ℓ*-sequence, where* d *is odd, then the second half of each period of* $\mathbf{a}$ *is the bitwise complement of the first half. In particular, the number of zeroes and the number of ones in a single period of* $\mathbf{a}$ *are equal, so* $\mathbf{a}$ *is balanced.*

Theorem 16.2.4. *Fix a primitive character $\chi : \mathbb{Z}/(N) \to \mathbb{C}^\times$. Let* $\mathbf{a}$ *be an N-ary ℓ-sequence based on a connection integer $q \equiv q_0 \pmod N$, $0 < q_0 < N$. If q is prime, then*

$$\begin{aligned} |Z(\mathbf{a})| &= \frac{|sin((q_0-1)\pi/N)|}{|sin(\pi/N)|} \\ &\leq \min(q_0 - 1, N - q_0 + 1) \\ &\leq \frac{N}{2}. \end{aligned}$$

If $q = p^t$ is a power of a prime number p with $t \geq 2$ and $q/p \equiv q_0' \pmod N$, $0 < q_0' < N$, then

$$\begin{aligned} |Z(\mathbf{a})| &\leq \frac{|sin((q_0-1)\pi/N)|}{|sin(\pi/N)|} + \frac{|sin((q_0'-1)\pi/N)|}{|sin(\pi/N)|} \\ &\leq \min(q_0 - 1, N - q_0 + 1) + \min(q_0' - 1, N - q_0' + 1) \\ &\leq N. \end{aligned}$$

If $N = 2$, then $Z(\mathbf{a}) = 0$.

Proof. In all cases, since q_0 is invertible mod N, there is a $b \in \{1, 2, \cdots, N-1\}$ so that $bN \equiv 1 \pmod q$. Since $\mathbf{a}$ is an ℓ-sequence, N is primitive modulo q.

Assume q is prime. Then in the algebraic model for the FCSR, the state goes through all $(1, b, b^2 \pmod q, \cdots, b^{q-2} \pmod q)$. By primitivity, this is a permutation of $(1, 2, 3, \cdots, q-1)$. So

$$\begin{aligned} Z(\mathbf{a}) &= \sum_{n=1}^{q-1} \zeta^n \qquad (16.1) \\ &= \frac{q - q_0}{N} \sum_{n=1}^{N} \zeta^n + \sum_{n=1}^{q_0-1} \zeta^n \\ &= \sum_{n=1}^{q_0-1} \zeta^n. \end{aligned}$$

This gives

$$|Z(\mathbf{a})| = \left| \zeta \frac{\zeta^{q_0-1} - 1}{\zeta - 1} \right| = \frac{|\zeta^{(q_0-1)/2} - \zeta^{-(q_0-1)/2}|}{|\zeta^{1/2} - \zeta^{-1/2}|} = \frac{|sin((q_0-1)\pi/N)|}{|sin(\pi/N)|}.$$

Let x be a positive real number. Then $sin(x\pi/N) = sin((N-x)\pi/N)$. By calculus we see that for any $x > 0$

$$\frac{|sin(x\pi/N)|}{|sin(\pi/N)|} \leq \min(x, N - x),$$

and the result follows.

Now suppose $q = p^t$, with $t \geq 2$ and p prime. Then the estimate on the sum in equation (16.1) still works, but to get $Z(\mathbf{a})$ we have to subtract off the terms with $p|n$. The same bound as above applies to this sum (with q replaced by q/p). Thus if $q/p \equiv q_0'$, with $0 < q_0' < N$, then

$$\begin{aligned} |Z(\mathbf{a})| &= \left| \sum_{n=1}^{q_0-1} \zeta^n - \sum_{n=1}^{q_0'-1} (\zeta^p)^n \right| \\ &\leq \left| \sum_{n=1}^{q_0-1} \zeta^n \right| + \left| \sum_{n=1}^{q_0'-1} (\zeta^p)^n \right| \\ &= \frac{|sin((q_0-1)\pi/N)|}{|sin(\pi/N)|} + \frac{|sin((q_0'-1)\pi/N)|}{|sin(\pi/N)|} \end{aligned}$$

(note that ζ^p is a primitive Nth root of one). The result follows from this.

If $N = 2$, then in any ℓ-sequence 0 and 1 occur equally often, so $|Z(\mathbf{a})| = 0$. □

In particular, this tells you how to make the imbalance small: pick q so that $\min(q_0-1, N-q_0+1)$ is small (in the prime case) or $\min(|q_0 - q_0'|, N - |q_0 - q_0'|)$ is small (in the prime power case).

Qi and Xu have also considered the balance of large segments of binary ℓ-sequences [161]. For any finite sequence $\mathbf{a}'$ of length T and $s \in \{0, 1\}$, let $N(\mathbf{a}', s)$ denote the number of times s occurs in $\mathbf{a}'$.

Theorem 16.2.5. *Let $\mathbf{a} = a_0, a_1, \cdots$ be a binary ℓ-sequence with connection integer $q = p^e$, p prime. Let $\mathbf{a}' = a_i, \cdots, a_{i+N-1}$. Then for $s \in \{0, 1\}$ we have*

$$\left| N(\mathbf{a}', s) - \frac{N}{2} \right| \leq \frac{4}{\pi} q^{1/2} \ln(q) \ln(p) + O(q^{1/2} \ln(q)).$$

If N is small this says nothing, but if N is close to the period $\phi(q)$, it says the deviation from the average is approximately the square root of N.

16.3 Arithmetic correlations

The ordinary autocorrelations of ℓ-sequences seems to be quite difficult to compute, although Xu and Qi have been able to do so for some special shifts [192]. Instead we consider the arithmetic version of the autocorrelation. Recall from Section 11.3.c that the *arithmetic cross-correlation* (with shift τ) of two N-ary sequences $\mathbf{a}, \mathbf{b}$ is

$$\mathcal{C}^A_{\mathbf{a},\mathbf{b}}(\tau) = Z(a - b^{(\tau)}),$$

where $a, b^{(\tau)} \in \mathbb{Z}_N$, $\mathbf{a} = \mathbf{seq}_N(a)$, $\mathbf{b}^\tau = \mathbf{seq}_N(b^{(\tau)})$, and Z is the imbalance (see Definition 11.2.1) with respect to a primitive character $\chi : \mathbb{Z}/(N) \to \mathbb{C}^\times$. Denote the *arithmetic autocorrelation* by

$$\mathcal{A}^A_{\mathbf{a}}(\tau) = \mathcal{C}^A_{\mathbf{a},\mathbf{a}}(\tau).$$

In this section we consider the arithmetic cross-correlations of an ℓ-sequence and its decimations. Throughout this section we assume that the connection integer $q = p^t$ is a power of an odd prime p such that N is a primitive root modulo q. Thus $\mathbf{a}$ is an ℓ-sequence and $\mathbf{a} = \mathbf{seq}_N(-u/q)$ is the (coefficient sequence of the) N-adic expansion of a fraction $-u/q$ with $0 < u < q$ and u not divisible by p.

Theorem 16.3.1. *Let* $\mathbf{a}$ *be an* N*-ary* ℓ*-sequence based on a prime connection integer* $q \equiv q_0 \pmod N$ *with* $0 \le q_0 < N$. *Let* τ *be an integer that is not a multiple of the period* $q-1$. *Then* $|\mathcal{A}^A_{\mathbf{a}}(\tau)| < \min(q_0 - 1, N - q_0 + 1) \le N/2$. *If* $N = 2$, *then* $\mathcal{A}^A_{\mathbf{a}}(\tau) = 0$.

Proof. Set $\mathbf{a} = \mathbf{seq}_N(-u/q)$ as above. By an argument similar to the one in Section 11.3.d, the arithmetic autocorrelation of $\mathbf{a}$ with shift τ is the imbalance of the (periodic part of the sequence of coefficients of the N-adic expansion of the) rational number

$$\frac{(u - N^{-\tau}u) \pmod q}{q} = \frac{(u - N^{T-\tau}u) \pmod q}{q},$$

where the reduction modulo q is taken in the range $[-(q-1), 0]$. Since q is prime, this is again the rational number corresponding to an ℓ-sequence. The result then follows from Theorem 16.2.4. □

For the remainder of the section, we treat only binary sequences, so $N = 2$. Recall from Section 13.2 that if $\mathbf{a}$ is a sequence and $d \neq 0$ is an integer, then the sequence, $\mathbf{b}$ is a d-fold decimation of $\mathbf{a}$ if for every i, we have $b_i = a_{di}$. Let $\mathcal{F}_{\mathbf{a}}$ be the family of all d-fold decimations of $\mathbf{a}$, where d is relatively prime to the period of $\mathbf{a}$.

The following is a remarkable fact about binary ℓ-sequences.

Theorem 16.3.2. *Suppose that* $\mathbf{a}$ *is a binary* ℓ*-sequence based on a connection number* $q = p^t$ *with* p *prime. If* $\mathbf{c}$ *and* $\mathbf{b}$ *are sequences in* $\mathcal{F}_{\mathbf{a}}$, *then the arithmetic cross-correlation of* $\mathbf{c}$ *and* $\mathbf{b}$ *with shift* τ *is zero unless* $\tau = 0$ *and* $\mathbf{b} = \mathbf{c}$.

The remainder of this section consists of a proof of Theorem 16.3.2. If $a \in \mathbb{Z}_2$, then denote by $\bar{a}$ the complementary 2-adic integer. That is, replace each 1 by 0 and each 0 by 1 in the 2-adic expansion of a. Then $a + \bar{a} = -1 \in \mathbb{Z}_2$.

Proposition 16.3.3. *Let* $d > 0$ *be relatively prime to the period* $T = p^t - p^{t-1}$ *of* $\mathbf{a}$. *Let* $\mathbf{b}$ *be a* d*-fold decimation of* $\mathbf{a}$. *Express* $\mathbf{b} = \mathbf{seq}_2(-g/q')$ *as the 2-adic expansion of a fraction* $-g/q'$, *with* g, q' *relatively prime. Then* q' *divides* $2^{T/2} + 1$.

Proof. Let $b \in \mathbb{Z}_2$ be the 2-adic number associated to $\mathbf{b}$. By Proposition 16.2.2,

$$-b - 1 = \bar{b} = x + 2^{T/2} b$$

for some ordinary integer x (in fact, $0 \leq x < 2^{T/2}$). It follows that

$$(2^{T/2} + 1)b = -(x + 1)$$

and therefore

$$(2^{T/2} + 1)g = q'(x + 1).$$

The result follows since g and q' are relatively prime. □

Proof of Theorem 16.3.2. Let $T = p^t - p^{t-1}$. Let $\mathbf{b}$ and $\mathbf{c}$ have associated 2-adic numbers $b = -g/r$ and $c = -h/s$, respectively, with $\gcd(g, r) = \gcd(h, s) = 1$. The shift of $\mathbf{c}$ by τ corresponds to a 2-adic integer $2^{T-\tau}c + x$ for some ordinary integer x. The arithmetic cross-correlation of $\mathbf{b}$ and $\mathbf{c}$ with shift τ is the number of zeros minus the number of ones in one length T period of

$$u = b - (2^{T-\tau}c + x) = \frac{-(gs - 2^{T-\tau}hr + xrs)}{rs} \in \mathbb{Z}_2.$$

If $\mathbf{b}$ is a shift of $\mathbf{c}$ with shift τ, then $u = 0$ and the result follows.

Suppose $\mathbf{c}$ is not a shift of $\mathbf{b}$ with shift τ. Let $\mathbf{u} = u_0, u_1, \cdots = \mathbf{seq}_2(u)$. It suffices to show that any period of $\mathbf{u}$ has the same numbers of zeros and ones. Let $u = -f'/q'$ with $\gcd(f', q') = 1$. Then $q' = \mathrm{lcm}(r, s)$, so by Proposition 16.3.3, q' divides $2^{T/2} + 1$. Moreover f' is nonzero. In a single period of $\mathbf{u}$ we have

$$u_i = (B \cdot 2^{-i} \pmod{q'}) \pmod 2$$

for some B. Thus

$$\begin{aligned} u_{i+T/2}, &= (B \cdot 2^{-(i+T/2)} \pmod r) \pmod 2 \\ &= (B \cdot 2^{-i} \cdot 2^{-T/2} \pmod r) \pmod 2 \\ &= (-B \cdot 2^{-i} \pmod r) \pmod 2. \end{aligned}$$

Since q' is odd, for any $y \neq 0$ the parity of $-y \pmod{q'}$ is the opposite of the parity of y. Thus $u_{i+T/2}$ is the complement of u_i. These elements occur in pairs in $\mathbf{u}$ (since T is a period of $\mathbf{u}$), so $\mathbf{u}$ is balanced. □

It is then interesting to ask whether all such decimations are distinct.

Conjecture 16.3.4. *If $q > 13$ is prime, 2 is primitive modulo q, and $\mathbf{a}$ is an ℓ-sequence based on q, then every pair of allowable decimations of $\mathbf{a}$ is shift distinct.*

This conjecture is false if we allow $q = 3$, 5, 11, or 13. If Conjecture 16.3.4 holds for a prime q, then the set of distinct decimations $\mathbf{a}[d]$ with d relatively prime to $q-1$ consists of $\phi(q-1)$ distinct elements with ideal arithmetic correlation The famous result of Hooley [80] implies that under the extended Riemann hypothesis, 2 is primitive for a set of primes of positive relative density, so there would be an abundance of large families with ideal arithmetic correlations.

The conjecture can be restated in very elementary terms as follows. Let $q > 13$ be a prime number such that 2 is primitive mod q. Let E be the set of even integers $0 \le e \le q-1$. Fix A with $1 \le A \le q-1$. Suppose the mapping $x \mapsto Ax^d \pmod q$ preserves (but permutes the elements within) the set E. Then $d = 1$ and $A = 1$. The equivalence between these two statements follows from the fact that $\mathbf{a}[d]$ and $\mathbf{a}[e]$ are shift distinct if and only if $\mathbf{a}$ and $\mathbf{a}[h]$ are shift distinct, where $h = de^{-1} \pmod{q-1}$.

Conjecture 16.3.4 has been verified for all primes $q < 8,000,000$, in fact even without assuming 2 is primitive modulo q. The conjecture has only been partially proven in general.

Theorem 16.3.5. *([59, 65]) Suppose either $d = -1$ (or, equivalently, $d = q-2$) or $q \equiv 1 \pmod 4$ and $d = (q+1)/2$. Then the decimation $\mathbf{a}[d]$ is shift distinct from $\mathbf{a}$.*

In 2004, Goresky, Klapper, Murty, and Shparlinski used exponential method to get some partial results – the conjecture holds for sufficiently large connection integers of a certain form [66]. This approach was improved on by Bourgain, Cochrane, Paulhus, and Pinner [19] using recent developments in methods of estimating exponential sums.

Theorem 16.3.6. *([19]) Conjecture 16.3.4 holds for any prime $q > 2.26 \cdot 10^{55}$.*

Finally, Xu and Qi have considered the generalization to the case when q is a nontrivial prime power, using purely algebraic methods.

Theorem 16.3.7. *([193]) Conjecture 16.3.4 holds if $q = p^e > 13$, $e \ge 2$, p prime, and 2 primitive modulo q.*

We omit the proofs of these theorems.

16.3.a Computing arithmetic cross-correlations

If $\mathbf{b} = \mathbf{seq}_N(b)$ and $\mathbf{c} = \mathbf{seq}_N(c)$ are two periodic sequences associated to $b, c \in \mathbb{Z}_N$, then the sequence $\mathbf{seq}_N(b-c)$ may not be strictly periodic (although it is eventually periodic). Thus, computing the arithmetic cross-correlation of two sequences appears to be problematic. How many coefficients of the difference must be computed before we reach the periodic part?

Proposition 16.3.8. *Let $\mathbf{b} = \mathbf{seq}_N(b)$ and $\mathbf{c} = \mathbf{seq}_N(c)$ be periodic N-ary sequences with period T, associated to $b, c \in \mathbb{Z}_N$. Then the sequence $\mathbf{d} = \mathbf{seq}_N(b-c)$ is strictly periodic at least from the Tth symbol.*

Proof. The strict periodicity of **b** and **c** implies that there are integers g, q, h and r such that $b = -g/q$, $c = -h/r$, $0 \le g \le q$, and $0 \le h \le r$. Thus

$$b - c = \frac{hq - gr}{qr}.$$

We have $-qr \le -gr \le hq - gr \le hq \le rq$. If $hq - gr \le 0$, then $b - c$ is strictly periodic and we are done. Otherwise

$$b - c = 1 + \frac{hq - gr - qr}{qr}$$

and $-qr < hq - gr - qr \le 0$. Therefore the sequence $\mathbf{u} = \mathbf{seq}_N(u)$ is strictly periodic, where

$$u = \frac{hq - gr - qr}{qr}.$$

If every element of **u** is $N - 1$, then $u = -1$ and $b - c = 0$. Otherwise, there is an $i < T$ such that $0 \le u_i \le N - 2$. When we carry out the addition $1 + u$, there is no carry beyond the ith position. It follows that the coefficients of $b - c = 1 + u$ are identical to those of u from the $i + 1$st coefficient on. This proves the proposition. □

Consequently, the arithmetic cross-correlation of **b** and **c** can be computed by computing the first $2T$ coefficients of the difference $b - c$, and finding the imbalance of the last T of these $2T$ coefficients. This is a linear time computation in T (although not easily parallelizable as is the case with standard cross-correlations). Furthermore, if we let $\bar{c}$ denote the N-adic number obtained by replacing each coefficient c_i of c by $N - 1 - c_i$, then $-c = \bar{c} + 1$. We can compute $b - c$ as $b + \bar{c} + 1$. If the carry from computing the first T coefficients of $b + \bar{c} + 1$ is 1, then $b + \bar{c} + 1$ is strictly periodic, so the first T coefficients suffice. Otherwise, the periodic part is exactly the first T coefficients of $b + \bar{c}$. Thus if we want to avoid storing b and c, we can simultaneously compute the arithmetic cross-correlation based on the first T coefficients of $b + \bar{c} + 1$ and $b + \bar{c}$ as the coefficients arrive, and use the former if there is a carry from the first T coefficients, and the latter otherwise.

16.4 Tables

We list in tabular form all prime connection integers q giving ℓ-sequences for $N = 2$ with $\log_2(q) \le 13$. For $N = 3, 5, 6, 7, 10$, we list all such q less than $10,000$. For $N = 4$, there are no prime integers q such that 4 is primitive — such a q must be odd, and, since 4 is a square, its order is at most $(q - 1)/2$. Similarly, $N = 8$ is not primitive unless $\gcd(3, q - 1) = 1$, when 8 is primitive if and only if 2 is primitive. Similarly, 9 is primitive modulo a prime q if and only if $q - 1$ is odd and 3 is primitive modulo q.

Length	Values of q giving binary ℓ-sequences
1	3
2	5
3	11, 13
4	19, 29
5	37, 53, 59, 61
6	67, 83, 101, 107
7	131, 139, 149, 163, 173, 179, 181, 197, 211, 227
8	269, 293, 317, 347, 349, 373, 379, 389, 419, 421, 443, 461, 467, 491, 509
9	523, 541, 547, 557, 563, 587, 613, 619, 653, 659, 661, 677, 701, 709, 757, 773, 787, 797, 821, 827, 829, 853, 859 , 877, 883 , 907, 941, 947, 1019
10	1061, 1091, 1109, 1117, 1123, 1171, 1187, 1213, 1229, 1237, 1259, 1277, 1283, 1291, 1301, 1307, 1373, 1381, 1427, 1451, 1453, 1483, 1493, 1499, 1523, 1531, 1549, 1571, 1619, 1621, 1637, 1667, 1669, 1693, 1733, 1741, 1747, 1787, 1861, 1867, 1877, 1901, 1907, 1931, 1949, 1973, 1979, 1987, 1997, 2027, 2029
11	2053, 2069, 2083, 2099, 2131, 2141, 2213, 2221, 2237, 2243, 2267, 2269, 2293, 2309, 2333, 2339, 2357, 2371, 2389, 2437, 2459, 2467, 2477, 2531, 2539, 2549, 2557, 2579, 2621, 2659, 2677, 2683, 2693, 2699, 2707, 2741, 2789, 2797, 2803, 2819, 2837, 2843, 2851, 2861, 2909, 2939, 2957, 2963, 3011, 3019, 3037, 3067, 3083, 3187, 3203, 3253, 3299, 3307, 3323, 3347, 3371, 3413, 3461, 3467, 3469, 3491, 3499, 3517, 3533, 3539, 3547, 3557, 3571, 3581, 3613, 3637, 3643, 3659, 3677, 3691, 3701, 3709, 3733, 3779, 3797, 3803, 3851, 3853, 3877, 3907, 3917, 3923, 3931, 3947, 3989, 4003, 4013, 4019, 4021, 4091, 4093

Table 16.1: Values of q Giving Rise to Binary ℓ-sequences for Length ≤ 11.

16.5 Exercises

1. Find an odd prime power q such that 2 is primitive modulo q and an integer s so that the numbers of occurrences of subsequences of length s in the binary ℓ-sequence with connection integer q vary by 2.

2. Let $\mathbf{a}$ be an N-ary ℓ-sequence with prime connection integer q. Assume that $q \equiv 1 \pmod{N}$ (so that now we are considering sequences generated by AFSRs).

a. Prove that $Z(\mathbf{a}) = 0$.
b. Use part (a) to prove that $\mathcal{A}_{\mathbf{a}}^{A}(\tau) = 0$ for all τ such that $q - 1$ does not divide τ.

Length	Values of q giving binary ℓ-sequences
12	4099, 4133, 4139, 4157, 4219, 4229, 4243, 4253, 4259, 4261, 4283, 4349, 4357, 4363, 4373, 4397, 4451, 4483, 4493, 4507, 4517, 4547, 4603, 4621, 4637, 4691, 4723, 4787, 4789, 4813, 4877, 4933, 4957, 4973, 4987, 5003, 5011, 5051, 5059, 5077, 5099, 5107, 5147, 5171, 5179, 5189, 5227, 5261, 5309, 5333, 5387, 5443, 5477, 5483, 5501, 5507, 5557, 5563, 5573, 5651, 5659, 5683, 5693, 5701, 5717, 5741, 5749, 5779, 5813, 5827, 5843, 5851, 5869, 5923, 5939, 5987, 6011, 6029, 6053, 6067, 6101, 6131, 6173, 6197, 6203, 6211, 6229, 6269, 6277, 6299, 6317, 6323, 6373, 6379, 6389, 6397, 6469, 6491, 6547, 6619, 6637, 6653, 6659, 6691, 6701, 6709, 6733, 6763, 6779, 6781, 6803, 6827, 6829, 6869, 6883, 6899, 6907, 6917, 6947, 6949, 6971, 7013, 7019, 7027, 7043, 7069, 7109, 7187, 7211, 7219, 7229, 7237, 7243, 7253, 7283, 7307, 7331, 7349, 7411, 7451, 7459, 7477, 7499, 7507, 7517, 7523, 7541, 7547, 7549, 7573, 7589, 7603, 7621, 7643, 7669, 7691, 7717, 7757, 7789, 7829, 7853, 7877, 7883, 7901, 7907, 7933, 7949, 8053, 8069, 8093, 8117, 8123, 8147, 8171, 8179
13	8219, 8221, 8237, 8243, 8269, 8291, 8293, 8363, 8387, 8429, 8443, 8467, 8539, 8563, 8573, 8597, 8627, 8669, 8677, 8693, 8699, 8731, 8741, 8747, 8803, 8819, 8821, 8837, 8861, 8867, 8923, 8933, 8963, 8971, 9011, 9029, 9059, 9173, 9181, 9203, 9221, 9227, 9283, 9293, 9323, 9341, 9349, 9371, 9397, 9419, 9421, 9437, 9467, 9491, 9533, 9539, 9547, 9587, 9613, 9619, 9629, 9643, 9661, 9677, 9733, 9749, 9803, 9851, 9859, 9883, 9901, 9907, 9923, 9941, 9949, 10037, 10067, 10069, 10091, 10093, 10099, 10133, 10139, 10141, 10163, 10181, 10253, 10259, 10267, 10301, 10331, 10357, 10427, 10459, 10477, 10499, 10501, 10589, 10613, 10667, 10691, 10709, 10723, 10733, 10789, 10837, 10853, 10859, 10861, 10867, 10883, 10891, 10909, 10949, 10973, 10979, 10987, 11003, 11027, 11069, 11083, 11093, 11131, 11171, 11197, 11213, 11261, 11317, 11437, 11443, 11483, 11549, 11579, 11587, 11621, 11677, 11699, 11717, 11779, 11789, 11813, 11821, 11827, 11867, 11909, 11933, 11939, 11981, 11987, 12011, 12043, 12107, 12149, 12157, 12197, 12203, 12211, 12227, 12251, 12253, 12269, 12277, 12301, 12323, 12347, 12373, 12379, 12413, 12437, 12491, 12539, 12547, 12589, 12611, 12613, 12619, 12637, 12653, 12659, 12739, 12757, 12763, 12781, 12821, 12829, 12899, 12907, 12917, 12923, 12941, 12979, 13037, 13043, 13109, 13147, 13163, 13187, 13229, 13291, 13331, 13339, 13397, 13411, 13451, 13469, 13477, 13523, 13613, 13619, 13627, 13691, 13709, 13723, 13757, 13763, 13829, 13859, 13877, 13883, 13901, 13907, 13931, 13933, 13997, 14011, 14051, 14107, 14173, 14221, 14243, 14341, 14387, 14389, 14411, 14419, 14461, 14533, 14549, 14557, 14621, 14627, 14629, 14653, 14669, 14699, 14717, 14723, 14741, 14747, 14771, 14797, 14813, 14821, 14827, 14843, 14851, 14867, 14869, 14891, 14923, 14939, 14947, 14957, 15013, 15053, 15061, 15077, 15083, 15091, 15101, 15107, 15131, 15139, 15149, 15173, 15187, 15227, 15259, 15269, 15299, 15331, 15349, 15373, 15413, 15427, 15443, 15461, 15581, 15629, 15661, 15667, 15683, 15731, 15739, 15749, 15773, 15787, 15797, 15803, 15859, 15907, 15923, 15971, 16067, 16069, 16139, 16187, 16189, 16229, 16253, 16301, 16333, 16339, 16349, 16363, 16381

Table 16.2: Values of q Giving Rise to Binary ℓ-sequences for Length 12 and 13.

Length	Values of q giving trinary ℓ-sequences
1	5, 7
2	17, 19
3	29, 31, 43, 53, 79
4	89, 101, 113, 127, 137, 139, 149, 163, 173, 197, 199, 211, 223, 233
5	257, 269, 281, 283, 293, 317, 331, 353, 379, 389, 401, 449, 461, 463, 487, 509, 521, 557 569, 571, 593, 607, 617, 631, 641, 653, 677, 691, 701
6	739, 751, 773, 797, 809, 811, 821, 823, 857, 859, 881, 907, 929, 941, 953, 977, 1013 1039, 1049, 1061, 1063, 1087, 1097, 1109, 1123, 1193, 1217, 1229, 1231, 1277, 1279, 1291, 1301, 1327, 1361, 1373, 1409, 1423, 1433, 1447, 1459, 1481, 1483, 1493, 1553, 1567, 1579, 1601, 1613, 1627, 1637, 1663, 1697, 1699, 1709, 1721, 1723, 1733, 1747, 1831, 1889, 1901, 1913, 1949, 1951, 1973, 1987, 1997, 1999, 2011, 2069, 2081, 2083, 2129, 2141, 2143, 2153
7	2213, 2237, 2239, 2273, 2309, 2311, 2333, 2347, 2357, 2371, 2381, 2393, 2417, 2467, 2477, 2503, 2539, 2549, 2609, 2633, 2647, 2657, 2659, 2683, 2693, 2707, 2719, 2729, 2731, 2741, 2753, 2767, 2777, 2789, 2801, 2837, 2861, 2897, 2909, 2957, 2969, 3041, 3089, 3137, 3163, 3209, 3257, 3259, 3271, 3307, 3329, 3331, 3389, 3391, 3413, 3449, 3461, 3463, 3533, 3547, 3557, 3559, 3571, 3581, 3583, 3593, 3617, 3643, 3677, 3701, 3727, 3761, 3797, 3821, 3823, 3833, 3917, 3919, 3929, 3931, 3943, 3989, 4001, 4003, 4013, 4027, 4049, 4073, 4133, 4157, 4159, 4217, 4219, 4229, 4231, 4241, 4243, 4253, 4289, 4327, 4337, 4349, 4363, 4373, 4397, 4409, 4421, 4423, 4447, 4457, 4481, 4493, 4507, 4517, 4519, 4567, 4603, 4637, 4639, 4649, 4651, 4663, 4673, 4723, 4759, 4793, 4817, 4831, 4877, 4889, 4903, 4937, 4973, 4987, 4999, 5009, 5021, 5023, 5081, 5119, 5189, 5237, 5261, 5273, 5297, 5309, 5333, 5347, 5381, 5393, 5407, 5417, 5419, 5431, 5441, 5443, 5477, 5479, 5503, 5563, 5573, 5647, 5657, 5669, 5683, 5693, 5717, 5741, 5779, 5801, 5813, 5827, 5849, 5861, 5897, 5923, 5981, 6007, 6029, 6053, 6089, 6113, 6151, 6163, 6173, 6197, 6199, 6221, 6257, 6269, 6317, 6329, 6343, 6353, 6367, 6379, 6389, 6427, 6449, 6451, 6473, 6547
8	6569, 6571, 6607, 6619, 6653, 6689, 6691, 6737, 6761, 6763, 6823, 6833, 6857, 6869, 6871, 6907, 6917, 6977, 7001, 7013, 7039, 7109, 7121, 7159, 7193, 7207, 7229, 7243, 7253, 7349, 7411, 7433, 7457, 7459, 7517, 7529, 7541, 7577, 7603, 7649, 7673, 7699, 7723, 7759, 7793, 7817, 7829, 7867, 7877, 7879, 7901, 7927, 7937, 7949, 8009, 8059, 8069, 8081, 8093, 8117, 8167, 8179, 8237, 8263, 8273, 8287, 8297, 8311, 8369, 8419, 8429, 8431, 8443, 8467, 8537, 8539, 8563, 8573, 8597, 8599, 8609, 8623, 8647, 8669, 8693, 8719, 8731, 8741, 8753, 8837, 8839, 8849, 8861, 8863, 8887, 8923, 8969, 8971, 9007, 9029, 9041, 9043, 9067, 9091, 9127, 9137, 9151, 9161, 9173, 9187, 9199, 9209, 9257, 9281, 9283, 9293, 9319, 9377, 9391, 9403, 9413, 9461, 9463, 9473, 9497, 9511, 9521, 9533, 9547, 9629, 9631, 9643, 9677, 9679, 9689, 9739, 9749, 9787, 9811, 9833, 9871, 9883, 9907, 9929, 9967

Table 16.3: Values of q Giving Rise to Trinary ℓ-sequences for $q \leq 10,000$.

Length	Values of q giving 5-ary ℓ-sequences
1	7, 17, 23
2	37, 43, 47, 53, 73, 83, 97, 103, 107, 113
3	137, 157, 167, 173, 193, 197, 223, 227, 233, 257, 263, 277, 283, 293, 307, 317, 347, 353, 373, 383, 397, 433, 443, 463, 467, 503, 523, 547, 557, 563, 577, 587, 593, 607, 613, 617
4	647, 653, 673, 677, 683, 727, 743, 757, 773, 787, 797, 857, 863, 877, 887, 907, 937, 947, 953, 967, 977, 983, 1013, 1033, 1093, 1097, 1103, 1153, 1163, 1187, 1193, 1213, 1217, 1223, 1237, 1277, 1283, 1307, 1327, 1367, 1373, 1427, 1433, 1483, 1487, 1493, 1523, 1543, 1553, 1567, 1583, 1607, 1613, 1637, 1663, 1667, 1693, 1697, 1733, 1747, 1777, 1787, 1823, 1847, 1877, 1907, 1913, 1933, 1987, 1993, 1997, 2003, 2017, 2027, 2053, 2063, 2083, 2087, 2113, 2143, 2153, 2203, 2207, 2213, 2237, 2243, 2267, 2273, 2293, 2297, 2333, 2347, 2357, 2377, 2383, 2393, 2417, 2423, 2437, 2447, 2467, 2473, 2477, 2503, 2543, 2557, 2617, 2633, 2647, 2657, 2663, 2677, 2683, 2687, 2693, 2713, 2753, 2767, 2777, 2797, 2833, 2837, 2843, 2887, 2897, 2903, 2917, 2927, 2957, 2963, 3023, 3037, 3067, 3083
5	3163, 3167, 3187, 3203, 3217, 3253, 3307, 3323, 3343, 3347, 3373, 3407, 3413, 3433, 3463, 3467, 3517, 3527, 3533, 3547, 3557, 3583, 3593, 3607, 3613, 3617, 3623, 3643, 3673, 3677, 3697, 3767, 3793, 3797, 3803, 3833, 3847, 3863, 3907, 3917, 3923, 3943, 3947, 4003, 4007, 4013, 4027, 4057, 4073, 4093, 4127, 4133, 4153, 4157, 4177, 4217, 4253, 4273, 4283, 4297, 4337, 4357, 4363, 4373, 4397, 4423, 4457, 4463, 4483, 4493, 4507, 4517, 4523, 4547, 4567, 4583, 4597, 4603, 4637, 4643, 4673, 4703, 4733, 4787, 4793, 4813, 4817, 4877, 4903, 4933, 4937, 4957, 4967, 4973, 4987, 4993, 5003, 5087, 5147, 5153, 5237, 5273, 5297, 5303, 5323, 5333, 5347, 5387, 5393, 5413, 5417, 5437, 5443, 5477, 5483, 5507, 5527, 5557, 5563, 5573, 5623, 5647, 5653, 5657, 5693, 5717, 5737, 5807, 5813, 5843, 5867, 5897, 5903, 5923, 5927, 5987, 6007, 6037, 6043, 6047, 6053, 6113, 6133, 6143, 6173, 6197, 6203, 6217, 6247, 6257, 6263, 6277, 6317, 6323, 6353, 6367, 6373, 6473, 6547, 6563, 6577, 6607, 6637, 6653, 6673, 6703, 6737, 6763, 6803, 6827, 6833, 6857, 6863, 6907, 6917, 6947, 6967, 6977, 6983, 6997, 7013, 7027, 7043, 7057, 7103, 7127, 7187, 7193, 7213, 7237, 7247, 7253, 7283, 7297, 7307, 7393, 7417, 7433, 7457, 7487, 7517, 7523, 7547, 7573, 7577, 7583, 7607, 7643, 7673, 7703, 7727, 7757, 7793, 7817, 7823, 7867, 7873, 7877, 7883, 7907, 7927, 7937, 7963, 7993, 8017, 8053, 8087, 8117, 8123, 8147, 8237, 8243, 8273, 8287, 8297, 8353, 8363, 8377, 8387, 8423, 8447, 8513, 8527, 8537, 8543, 8573, 8597, 8623, 8627, 8647, 8663, 8677, 8693, 8707, 8713, 8737, 8747, 8753, 8783, 8803, 8807, 8837, 8863, 8867, 8893, 8923, 8963, 9007, 9013, 9043, 9067, 9127, 9137, 9173, 9187, 9203, 9227, 9257, 9277, 9293, 9323, 9337, 9343, 9377, 9413, 9433, 9437, 9467, 9473, 9497, 9533, 9587, 9613, 9623, 9643, 9677, 9733, 9743, 9767, 9787, 9803, 9817, 9833, 9857, 9883, 9887, 9907, 9967

Table 16.4: Values of q Giving Rise to 5-ary ℓ-sequences for $q \leq 10,000$.

Length	Values of q giving 6-ary ℓ-sequences
1	11, 13, 17
2	41, 59, 61, 79, 83, 89, 103, 107, 109, 113, 127, 131, 137, 151, 157, 179, 199
3	223, 227, 229, 233, 251, 257, 271, 277, 347, 367, 373, 397, 401, 419, 443, 449, 467, 487, 491, 521, 563, 569, 587, 593, 613, 641, 659, 661, 683, 709, 733, 757, 761, 809, 823, 827, 829, 853, 857, 881, 929, 947, 953, 967, 971, 977, 991, 1019, 1039, 1049, 1063, 1069, 1091, 1097, 1117, 1163, 1187, 1193, 1213, 1217, 1237, 1259, 1279, 1283, 1289
4	1303, 1307, 1327, 1361, 1381, 1409, 1423, 1427, 1429, 1433, 1451, 1471, 1481, 1499, 1523, 1543, 1549, 1553, 1567, 1601, 1619, 1621, 1663, 1667, 1669, 1759, 1789, 1811, 1831, 1861, 1879, 1889, 1907, 1913, 1979, 1999, 2027, 2029, 2053, 2081, 2099, 2129, 2143, 2153, 2221, 2243, 2267, 2269, 2273, 2293, 2297, 2389, 2393, 2411, 2417, 2437, 2441, 2459, 2503, 2531, 2551, 2557, 2579, 2609, 2633, 2657, 2699, 2729, 2749, 2767, 2777, 2791, 2797, 2801, 2819, 2843, 2887, 2897, 2917, 2939, 2963, 2969, 3011, 3041, 3061, 3079, 3083, 3089, 3109, 3137, 3203, 3209, 3229, 3251, 3253, 3257, 3271, 3299, 3301, 3319, 3323, 3329, 3347, 3371, 3391, 3449, 3463, 3467, 3469, 3491, 3517, 3539, 3583, 3593, 3617, 3659, 3709, 3733, 3761, 3779, 3803, 3823, 3833, 3847, 3853, 3929, 3943, 3947, 3967, 4001, 4019, 4049, 4073, 4091, 4093, 4139, 4211, 4217, 4231, 4241, 4259, 4283, 4289, 4327, 4337, 4409, 4423, 4447, 4451, 4457, 4481, 4519, 4523, 4547, 4549, 4567, 4597, 4621, 4639, 4643, 4649, 4663, 4673, 4691, 4721, 4759, 4783, 4787, 4789, 4793, 4813, 4909, 4931, 4933, 4937, 4951, 4957, 5003, 5009, 5051, 5077, 5081, 5099, 5101, 5147, 5153, 5167, 5171, 5273, 5297, 5387, 5393, 5407, 5413, 5417, 5437, 5441, 5483, 5507, 5527, 5557, 5581, 5623, 5647, 5651, 5653, 5657, 5701, 5791, 5821, 5839, 5843, 5867, 5897, 5939, 5987, 6007, 6011, 6037, 6089, 6113, 6131, 6133, 6151, 6199, 6229, 6247, 6257, 6277, 6299, 6323, 6353, 6367, 6373, 6421, 6449, 6473, 6491, 6521, 6569, 6607, 6637, 6659, 6661, 6689, 6703, 6709, 6737, 6761, 6779, 6803, 6827, 6829, 6833, 6857, 6871, 6899, 6947, 6971, 6977, 6991, 7001, 7019, 7039, 7043, 7069, 7121, 7187, 7193, 7207, 7283, 7307, 7309, 7331, 7333, 7351, 7433, 7451, 7457, 7477, 7481, 7499, 7523, 7529, 7547, 7549, 7573, 7577, 7591, 7621, 7643, 7649, 7669, 7687, 7717
5	7793, 7817, 7907, 7927, 7933, 7937, 7951, 8009, 8081, 8101, 8123, 8147, 8167, 8219, 8243, 8263, 8273, 8287, 8291, 8293, 8297, 8317, 8369, 8387, 8389, 8461, 8513, 8527, 8537, 8581, 8599, 8609, 8623, 8627, 8629, 8647, 8677, 8699, 8719, 8747, 8753, 8819, 8839, 8863, 8941, 8963, 8969, 9013, 9041, 9059, 9103, 9133, 9137, 9157, 9161, 9181, 9203, 9209, 9227, 9257, 9281, 9323, 9343, 9349, 9371, 9377, 9421, 9463, 9467, 9473, 9491, 9497, 9511, 9521, 9539, 9587, 9613, 9631, 9661, 9689, 9781, 9803, 9833, 9851, 9857, 9901, 9923, 9929, 9949, 9967

Table 16.5: Values of q Giving Rise to 6-ary ℓ-sequences for $q \leq 10,000$.

Length	Values of q giving 7-ary ℓ-sequences
1	11, 13, 17, 23, 41
2	61, 67, 71, 79, 89, 97, 101, 107, 127, 151, 163, 173, 179, 211, 229, 239, 241, 257, 263, 269, 293
3	347, 349, 359, 379, 397, 431, 433, 443, 461, 491, 499, 509, 521, 547, 577, 593, 599, 601, 631, 659, 677, 683, 733, 739, 743, 761, 773, 797, 823, 827, 857, 863, 907, 919, 929, 937, 941, 967, 991, 997, 1013, 1019, 1049, 1051, 1069, 1097, 1103, 1109, 1163, 1181, 1187, 1193, 1217, 1237, 1249, 1277, 1283, 1301, 1303, 1361, 1367, 1433, 1439, 1451, 1471, 1523, 1553, 1601, 1607, 1609, 1613, 1619, 1637, 1667, 1669, 1693, 1697, 1721, 1747, 1753, 1759, 1787, 1831, 1889, 1949, 1973, 1993, 2003, 2011, 2027, 2039, 2083, 2087, 2089, 2111, 2141, 2143, 2179, 2207, 2251, 2273, 2281, 2309, 2339, 2341, 2357, 2393
4	2447, 2459, 2477, 2503, 2531, 2543, 2591, 2593, 2609, 2621, 2647, 2671, 2677, 2693, 2699, 2711, 2729, 2731, 2777, 2789, 2843, 2851, 2879, 2897, 2917, 2957, 2963, 3041, 3119, 3121, 3169, 3181, 3203, 3209, 3253, 3271, 3299, 3343, 3371, 3449, 3457, 3467, 3511, 3517, 3533, 3539, 3541, 3571, 3607, 3617, 3623, 3673, 3701, 3719, 3739, 3767, 3769, 3793, 3797, 3803, 3821, 3853, 3931, 3943, 3989, 4019, 4049, 4073, 4093, 4127, 4133, 4139, 4157, 4177, 4211, 4243, 4261, 4271, 4273, 4289, 4297, 4327, 4373, 4409, 4447, 4457, 4463, 4493, 4513, 4519, 4523, 4547, 4549, 4597, 4603, 4637, 4643, 4663, 4691, 4721, 4793, 4799, 4801, 4831, 4877, 4889, 4943, 4951, 4967, 4973, 4999, 5081, 5107, 5119, 5147, 5197, 5231, 5279, 5281, 5297, 5303, 5309, 5333, 5381, 5387, 5393, 5399, 5417, 5437, 5449, 5471, 5477, 5483, 5501, 5503, 5557, 5623, 5639, 5651, 5669, 5717, 5779, 5783, 5791, 5801, 5807, 5813, 5857, 5867, 5869, 5897, 5903, 5923, 5953, 5981, 5987, 6043, 6053, 6089, 6091, 6121, 6143, 6173, 6199, 6203, 6211, 6221, 6229, 6257, 6287, 6311, 6317, 6323, 6343, 6367, 6379, 6389, 6397, 6473, 6481, 6491, 6529, 6547, 6563, 6569, 6619, 6653, 6659, 6679, 6703, 6709, 6733, 6737, 6763, 6781, 6791, 6827, 6899, 6911, 6949, 6959, 6967, 6977, 6983, 7013, 7039, 7043, 7069, 7079, 7127, 7129, 7151, 7207, 7211, 7213, 7219, 7229, 7237, 7247, 7321, 7331, 7349, 7351, 7369, 7433, 7459, 7481, 7487, 7489, 7499, 7517, 7537, 7547, 7549, 7573, 7577, 7583, 7603, 7639, 7649, 7741, 7789, 7817, 7823, 7829, 7853, 7873, 7879, 7883, 7901, 7907, 7919, 7937, 7963, 8053, 8059, 8069, 8081, 8087, 8161, 8171, 8209, 8219, 8221, 8237, 8243, 8273, 8293, 8329, 8377, 8387, 8423, 8467, 8501, 8563, 8573, 8581, 8609, 8647, 8663, 8669, 8693, 8741, 8747, 8753, 8803, 8807, 8831, 8837, 8863, 8893, 8999, 9001, 9011, 9029, 9049, 9059, 9133, 9151, 9161, 9173, 9227, 9257, 9283, 9311, 9319, 9337, 9341, 9397, 9403, 9413, 9419, 9431, 9479, 9497, 9533, 9587, 9649, 9677, 9721, 9733, 9739, 9749, 9767, 9811, 9833, 9839, 9871, 9907, 9923, 9929

Table 16.6: Values of q Giving Rise to 7-ary ℓ-sequences for $q \leq 10,000$.

Length	Values of q giving 10-ary ℓ-sequences
1	17, 19, 23, 29, 47, 59, 61, 97
2	109, 113, 131, 149, 167, 179, 181, 193, 223, 229, 233, 257, 263, 269, 313, 337, 367, 379, 383, 389, 419, 433, 461, 487, 491, 499, 503, 509, 541, 571, 577, 593, 619, 647, 659, 701, 709, 727, 743, 811, 821, 823, 857, 863, 887, 937, 941, 953, 971, 977, 983
3	1019, 1021, 1033, 1051, 1063, 1069, 1087, 1091, 1097, 1103, 1109, 1153, 1171, 1181, 1193, 1217, 1223, 1229, 1259, 1291, 1297, 1301, 1303, 1327, 1367, 1381, 1429, 1433, 1447, 1487, 1531, 1543, 1549, 1553, 1567, 1571, 1579, 1583, 1607, 1619, 1621, 1663, 1697, 1709, 1741, 1777, 1783, 1789, 1811, 1823, 1847, 1861, 1873, 1913, 1949, 1979, 2017, 2029, 2063, 2069, 2099, 2113, 2137, 2141, 2143, 2153, 2179, 2207, 2221, 2251, 2269, 2273, 2297, 2309, 2339, 2341, 2371, 2383, 2389, 2411, 2417, 2423, 2447, 2459, 2473, 2539, 2543, 2549, 2579, 2593, 2617, 2621, 2633, 2657, 2663, 2687, 2699, 2713, 2731, 2741, 2753, 2767, 2777, 2789, 2819, 2833, 2851, 2861, 2887, 2897, 2903, 2909, 2927, 2939, 2971, 3011, 3019, 3023, 3137, 3167, 3221, 3251, 3257, 3259, 3299, 3301, 3313, 3331, 3343, 3371, 3389, 3407, 3433, 3461, 3463, 3469, 3527, 3539, 3571, 3581, 3593, 3607, 3617, 3623, 3659, 3673, 3701, 3709, 3727, 3767, 3779, 3821, 3833, 3847, 3863, 3943, 3967, 3989, 4007, 4019, 4051, 4057, 4073, 4091, 4099, 4127, 4139, 4153, 4177, 4211, 4217, 4219, 4229, 4259, 4261, 4327, 4337, 4339, 4349, 4421, 4423, 4447, 4451, 4457, 4463, 4567, 4583, 4651, 4673, 4691, 4703, 4783, 4793, 4817, 4931, 4937, 4943, 4967, 5021, 5059, 5087, 5099, 5153, 5167, 5179, 5189, 5233, 5273, 5297, 5303, 5309, 5381, 5393, 5417, 5419, 5501, 5503, 5527, 5531, 5581, 5623, 5651, 5657, 5659, 5669, 5701, 5737, 5741, 5743, 5749, 5779, 5783, 5807, 5821, 5857, 5861, 5869, 5897, 5903, 5927, 5939, 5981, 6011, 6029, 6047, 6073, 6113, 6131, 6143, 6211, 6217, 6221, 6247, 6257, 6263, 6269, 6287, 6301, 6337, 6343, 6353, 6367, 6389, 6473, 6553, 6571, 6619, 6659, 6661, 6673, 6691, 6701, 6703, 6709, 6737, 6779, 6793, 6823, 6829, 6833, 6857, 6863, 6869, 6899, 6949, 6967, 6971, 6977, 6983, 7019, 7057, 7069, 7103, 7109, 7177, 7193, 7207, 7219, 7229, 7247, 7309, 7349, 7393, 7411, 7433, 7451, 7457, 7459, 7487, 7499, 7541, 7577, 7583, 7607, 7673, 7687, 7691, 7699, 7703, 7727, 7753, 7793, 7817, 7823, 7829, 7873, 7901, 7927, 7937, 7949, 8017, 8059, 8069, 8087, 8171, 8179, 8219, 8233, 8263, 8269, 8273, 8287, 8291, 8297, 8353, 8377, 8389, 8423, 8429, 8447, 8501, 8513, 8537, 8543, 8623, 8647, 8663, 8669, 8699, 8713, 8731, 8741, 8753, 8783, 8807, 8819, 8821, 8861, 8863, 8887, 8971, 9011, 9029, 9059, 9103, 9109, 9137, 9221, 9257, 9341, 9343, 9371, 9377, 9421, 9461, 9473, 9491, 9497, 9539, 9623, 9629, 9697, 9739, 9743, 9749, 9767, 9781, 9811, 9817, 9829, 9833, 9851, 9857, 9887, 9931, 9949, 9967

Table 16.7: Values of q Giving Rise to 10-ary ℓ-sequences for $q \leq 10,000$.

Chapter 17 Maximal Period d-FCSRs

In this chapter we investigate the distribution properties (Section 17.2) and correlation properties (Section 17.3) of maximal period d-FCSR sequences. As in Chapter 9, we assume we are given integers $N \geq 2$ and $d \geq 1$ such that the polynomial $x^d - N$ is irreducible over the rational numbers $\mathbb{Q}$. Let $\pi \in \mathbb{C}$ be a root of this polynomial in an extension field of $\mathbb{Q}$. In this chapter we consider maximum period AFSRs based on the ring

$$R = \mathbb{Z}[\pi] = \left\{ \sum_{i=0}^{d-1} a_i \pi^i : a_i \in \mathbb{Z} \right\}.$$

The fraction field of R is

$$F = \mathbb{Q}[\pi] = \left\{ \sum_{i=0}^{d-1} a_i \pi^i : a_i \in \mathbb{Q} \right\} = \left\{ \frac{a}{b} : a \in R, b \in \mathbb{Z} \right\}.$$

The algebra in this ring is somewhat more complicated than that of $\mathbb{Z}$, and is not completely understood. For example, R may not be a UFD and it may not be integrally closed. The group of units in R may be infinite. A complete characterization of the cases where R is a Euclidean domain is not known, even when $d = 2$.

A d-FCSR is an AFSR based on R, π, and $S = \{0, 1, \cdots, N-1\}$.

17.1 Identifying maximal length sequences

By Lemma 3.4.8, if $q \in R$ is a non-unit, then the absolute value $|\mathsf{N}_{\mathbb{Q}}^F(q)|$ is equal to the number of elements in the quotient ring $R/(q)$.

Theorem 17.1.1. *The maximum period for a periodic sequence generated by a d-FCSR with connection element q is $|\mathsf{N}_{\mathbb{Q}}^F(q)| - 1$.*

Thus the largest possible period for an ℓ-sequence occurs when $R/(q)$ is a field. Table 17.1 gives a list of all such connection elements and the resulting periods for $N = 2$, $d = 2$, and length at most 11.

More generally, we consider sequences for which the period is the cardinality of the group of units in $R/(\pi)$.

Length	$(q, \mathsf{N}_{\mathbb{Q}}^{F}(q))$ giving binary 2-ℓ-sequences
4	$(5+2\pi, 17), (5+3\pi, 7)$
6	$(7+1\pi, 47), (9+1\pi, 79), (7+2\pi, 41), (11+3\pi, 103), (7+4\pi, 17), (13+4\pi, 137),$ $(11+7\pi, 23), (13+6\pi, 97), (13+7\pi, 71)$
7	$(13+8\pi, 41)$
8	$(15+4\pi, 193), (17+5\pi, 239), (21+4\pi, 409), (19+7\pi, 263), (23+2\pi, 521),$ $(29+3\pi, 823), (23+5\pi, 479), (29+4\pi, 809), (29+6\pi, 769), (29+7\pi, 743),$ $(15+8\pi, 97), (21+8\pi, 313), (21+11\pi, 199), (23+8\pi, 401), (23+9\pi, 367),$ $(25+9\pi, 463), (27+11\pi, 487), (29+11\pi, 599), (23+13\pi, 191), (23+14\pi, 137),$ $(29+14\pi, 449)$
9	$(23+16\pi, 17), (25+17\pi, 47), (27+19\pi, 7), (29+18\pi, 193), (29+20\pi, 41)$
10	$(33+1\pi, 1087), (37+2\pi, 1361), (31+4\pi, 929), (33+5\pi, 1039), (37+6\pi, 1297),$ $(39+4\pi, 1489), (45+4\pi, 1993), (31+10\pi, 761), (35+11\pi, 983), (33+13\pi, 751),$ $(31+14\pi, 569), (37+14\pi, 977), (43+11\pi, 1607), (45+11\pi, 1783), (39+14\pi, 1129),$ $(47+1\pi, 2207), (49+1\pi, 2399), (53+2\pi, 2801), (47+6\pi, 2137), (53+4\pi, 2777),$ $(51+7\pi, 2503), (53+7\pi, 2711), (55+1\pi, 3023), (61+7\pi, 3623), (47+8\pi, 2081),$ $(49+9\pi, 2239), (53+10\pi, 2609), (47+13\pi, 1871), (49+13\pi, 2063), (53+12\pi, 2521),$ $(53+14\pi, 2417), (55+8\pi, 2897), (61+8\pi, 3593), (55+14\pi, 2633), (61+12\pi, 3433),$ $(61+14\pi, 3329), (31+16\pi, 449), (31+17\pi, 383), (31+18\pi, 313), (37+16\pi, 857),$ $(35+19\pi, 503), (37+19\pi, 647), (31+21\pi, 79), (37+20\pi, 569), (35+23\pi, 167),$ $(37+22\pi, 401), (37+23\pi, 311), (39+16\pi, 1009), (45+19\pi, 1303), (45+23\pi, 967),$ $(39+25\pi, 271), (45+31\pi, 103), (47+16\pi, 1697), (49+17\pi, 1823), (53+16\pi, 2297),$ $(51+19\pi, 1879), (53+18\pi, 2161), (53+19\pi, 2087), (47+20\pi, 1409),$ $(51+23\pi, 1543), (55+17\pi, 2447), (55+18\pi, 2377), (61+16\pi, 3209),$ $(61+19\pi, 2999), (59+23\pi, 2423), (61+22\pi, 2753), (61+23\pi, 2663),$ $(49+25\pi, 1151), (47+26\pi, 857), (49+29\pi, 719), (47+30\pi, 409), (53+30\pi, 1009),$ $(53+31\pi, 887), (55+24\pi, 1873), (57+29\pi, 1567), (61+28\pi, 2153), (59+31\pi, 1559)$
11	$(53+32\pi, 761), (53+35\pi, 359), (55+32\pi, 977), (59+35\pi, 1031),$ $(61+34\pi, 1409), (55+38\pi, 137), (61+36\pi, 1129), (61+40\pi, 521), (61+42\pi, 193),$ $(61+43\pi, 23)$

Table 17.1: Values of q Giving Rise to binary 2-ℓ-sequences for Length ≤ 11.

Definition 17.1.2. *A d-ℓ-sequence (or simply an ℓ-sequence if the context is clear) is a periodic sequence* $\mathbf{a}$ *which is obtained from a d-FCSR with connection element q such that the period of* $\mathbf{a}$ *is the cardinality of the multiplicative group of $R/(q)$.*

By Theorem 9.4.2, there is an exponential representation

$$a_i = (u\pi^{-i} \pmod q) \pmod \pi, \tag{17.1}$$

for an ℓ-sequence $\mathbf{a}$, where the reductions modulo q are taken in the fundamental half-open parallelepiped P as in Section 9.4. Thus $\mathbf{a}$ is an ℓ-sequence if and only if π is primitive in $R/(q)$. Different choices of u with $u \in P$ give different left shifts of the same sequence.

It follows from Theorem 3.4.9 and the Chinese Remainder Theorem (Theorem 2.2.18) that, as with ℓ-sequences in the FCSR case, if q is the connection element of a d-ℓ-sequence, then $q = ur^t$ where u is a unit and r is any irreducible factor of q. This holds regardless of whether R is a UFD. Moreover, irreducibility of $r \in \mathbb{Z}[\pi]$ can be guaranteed by requiring that $n = \mathsf{N}_{\mathbb{Q}}^F(r)$ is a prime in $\mathbb{Z}$. It follows that $Q = \mathsf{N}_{\mathbb{Q}}^F(q) = n^t$, so by Proposition 9.3.2,

$$\mathbb{Z}[\pi]/(q) \cong \mathbb{Z}/(\mathsf{N}_{\mathbb{Q}}^F(q)) \cong \mathbb{Z}/(n^t).$$

The group of units $\mathbb{Z}[\pi]/(q)^\times$ in this ring is cyclic (see Section 2.2.d) and has order

$$\phi(n^t) = n^{t-1}(n-1).$$

Thus such a q is the connection element of a d-ℓ-sequence if and only the $n^{t-1}(n-1)$ different powers of π exactly account for the elements in $\mathbb{Z}[\pi]/(q)^\times$.

17.2 Distribution properties of d-ℓ-sequences

We have seen in Theorem 16.2.1 that if $\mathbf{a}$ is an N-ary ℓ-sequence based on a prime connection integer q, then the numbers of occurrences of any two n-element patterns differ at most by one. The difference is at most 2 if the connection integer q is a power of a prime. In this section we consider the same question for the d-ℓ-sequences of Definition 17.1.2. As in Section 9.2 suppose that $x^d - N$ is irreducible over $\mathbb{Q}$ and that $q \in \mathbb{Z}[\pi]$ is invertible modulo π. Assume that $q = r^t$ for some $r \in \mathbb{Z}[\pi]$ such that $\mathsf{N}_{\mathbb{Q}}^F(r)$ is prime, and that the image of π in $\mathbb{Z}[\pi]/(q)$ is primitive. Let $Q = \mathsf{N}_{\mathbb{Q}}^F(q) \in \mathbb{Z}$.

17.2.a Reduction to counting lattice points

In this section we show that the number of occurrences of an s-element pattern in an ℓ-sequence is equal to the number of points in a certain integer lattice in a certain hypercube.

First observe that determining the distribution of occurrences of s-element patterns in a single period of $\mathbf{a}$ is equivalent to determining the distribution of occurrences of length s element patterns as the first s elements of left shifts of $\mathbf{a}$. Every left shift of $\mathbf{a}$ is the coefficient sequence of the

π-adic expansion of an element of the form u/q for some u. Moreover, the shift corresponding to u/q begins with the s-element pattern $\mathbf{b}$ if and only if

$$b_0 + b_1\pi + b_2\pi^2 + \cdots b_{s-1}\pi^{s-1} \equiv \frac{u}{q} \pmod{\pi^s}.$$

In Section 9.4 we analyzed the periodicity of d-FCSR sequences. Recall that

$$P_0 = \left\{ \sum_{i=0}^{d-1} v_i q \pi^i |\ v_i \in \mathbb{Q} \text{ and } -1 < v_i < 0 \right\} \subset \mathbb{Q}[\pi]$$

is the open parallelepiped for q, and

$$\Delta_0(q) = P_0 \cap \mathbb{Z}[\pi].$$

The following theorem follows from Corollary 9.4.3.

Theorem 17.2.1. *Suppose that $q \in \mathbb{Z}[\pi]$ is invertible modulo π, $\mathsf{N}_{\mathbb{Q}}^F(q)$ is prime, and π is primitive modulo q. Let $\mathbf{a}$ be an ℓ-sequence with connection element q. Then the π-adic expansion of v/q is a left shift of $\mathbf{a}$ if and only if $v \in \Delta_0(q)$.*

Suppose that $q \in \mathbb{Z}[\pi]$ is invertible modulo π, $q = ur^t$, $\mathsf{N}_{\mathbb{Q}}^F(r)$ is prime, and π is primitive modulo q. Let $\mathbf{a}$ be an ℓ-sequence with connection element q. Then the π-adic expansion of v/q is a left shift of $\mathbf{a}$ if and only if $v \in \Delta_0(q) - r\Delta_0(q/r)$.

Thus by Theorem 17.2.1, determining the distribution of occurrences of s-element patterns in a single period of $\mathbf{a}$ is equivalent to determining the distribution of elements

$$b = \sum_{i=0}^{s-1} b_i \pi^i, b_i \in \{0, 1, \cdots, N-1\},$$

where each such element is counted once for every $v \in \Delta_0(q)$ with

$$\frac{v}{q} \equiv b \pmod{\pi^s}.$$

But q is invertible modulo π^s, so multiplying by q modulo π^s is a permutation. Since our goal is just to determine the set of these multiplicities, we can simply determine for each b the number of $v \in \Delta_0(q)$ with $v \equiv b \pmod{\pi^s}$.

For $v \in \Delta_0(q)$, we have

$$v = \sum_{i=0}^{d-1} v_i \pi^i q \in \mathbb{Z}[\pi], \tag{17.2}$$

with $-1 < v_i < 0$. Recall that by Lemma 9.3.1 $\delta = Q/q$ is in R. Multiplying equation (17.2) by δ, we find that

$$\delta v = \sum_{i=0}^{d-1} v_i \pi^i Q \in \mathbb{Z}[\pi],$$

so that $v_i = z_i/Q$ for some integers z_i, $i = 0, \cdots, d-1$ with $-N < z_i < 0$. Thus for fixed b, we want to find the number of solutions to the equation

$$v = \sum_{i=0}^{d-1} \frac{z_i}{Q} \pi^i q \equiv b \pmod{\pi^s},$$

with $z_i \in \mathbb{Z}$ and $-Q < z_i < 0$. Equivalently, we want the number of $w \in \mathbb{Z}[\pi]$ such that

$$\sum_{i=0}^{d-1} \frac{z_i}{Q} \pi^i q = b + \pi^s w,$$

for some $z_i \in \mathbb{Z}$ with $-Q < z_i < 0$. Multiplying by δ again and dividing by π^s, we see that this is the same as the number of $w \in \mathbb{Z}[\pi]$ such that

$$\sum_{i=0}^{d-1} \frac{z_i}{\pi^s} \pi^i q = \frac{b\delta}{\pi^s} + w\delta,$$

for some $z_i \in \mathbb{Z}$ with $-Q < z_i < 0$.

Lemma 17.2.2. *Suppose that $(q) = (r^t)$ and the norm of r is prime. For every $w \in \mathbb{Z}[\pi]$ there are elements $k \in \mathbb{Z}$ and $y \in \mathbb{Z}[\pi]$ so that $w\delta = k\delta + yQ$.*

Proof. This is equivalent to showing that there exist elements $y \in \mathbb{Z}[\pi]$ and $k \in \mathbb{Z}$ so that $w = k + yq$. This follows from the fact that $\mathbb{Z}[\pi]/(q)$ is isomorphic to $\mathbb{Z}/(Q)$ when $(q) = (r^t)$ and the norm of r is prime. □

Note that it is this step in our analysis that depends on the special form of q. It follows that it suffices for us to count the number of pairs $k \in \mathbb{Z}, y \in \mathbb{Z}[\pi]$ such that

$$\sum_{i=0}^{d-1} \frac{z_i}{\pi^s} \pi^i q = \frac{b\delta}{\pi^s} + k\delta + yQ, \tag{17.3}$$

for some $z_i \in \mathbb{Z}$ with $-Q < z_i < 0$. If s is a multiple of d, then this is the number of points $k\delta + yQ$ in a real hypercube with faces parallel to the coordinate planes each of whose edges have length $Q/N^{s/d}$ and whose vertex with minimal coordinates is $b\delta/\pi^s$.

17.2.b When $d = 2$

In this section we apply the results of Section 17.2.a to the case when $d = 2$. We show that the number of occurrences of each s-tuple, s even, differs from the average number of occurrences by at most a small constant times the square root of the average number of occurrences.

Since $d = 2$ we have $q = q_0 + q_1\pi$, $q_0, q_1 \in \mathbb{Z}$, and $\delta = \pm(q_0 - q_1\pi)$. Let s be even and let

$$a = \frac{Q}{N^{s/2}} \in \mathbb{R}^{>0}.$$

Let B be an a by a square with sides parallel to the axes. We want to bound the cardinality of the set W of points in B of the form $k(q_0, -q_1) + (Qw_0, Qw_1)$, with $k, w_0, w_1 \in \mathbb{Z}$. Let us assume that $q_1 < 0$. The case when $q_1 > 0$ is similar. For fixed $w = (w_0, w_1) \in \mathbb{Z}^2$, let L_w be the real line

$$L_w = \{k(q_0, -q_1) + Qw : k \in \mathbb{R}\}.$$

Then W is the union over all w of the set of points

$$k(q_0, -q_1) + (Qw_0, Qw_1) \in L_w \cap B.$$

That is, W is a union of line segments. The number of lattice points on each such segment is approximately the length of the segment divided by the (constant) distance between consecutive lattice points. Alternatively, it is approximately the variation in the second coordinate along the segment divided by the variation in the second coordinate between two consecutive lattice points. The error in this estimate is at most one per line segment. Thus we can bound the size of W by the following steps:

1. Find numbers m_0 and m_1 so that the sum of the variations in the second coordinates along the segments is between m_0 and m_1.
2. The variation in the second coordinate between consecutive lattice points is $|q_1|$.
3. Bound the error: the actual number of lattice points on a segment whose second coordinate varies by c is greater than $(c/|q_1|) - 1$ and at most $(c/|q_1|) + 1$.
4. Let ℓ be the number of segments. Then the total number of lattice points is at least $(m_0/|q_1|) - \ell$ and at most $(m_1/|q_1|) + \ell$.

The slope of the segments is positive, so there are three types of lines: (a) those that start on the left hand vertical side of B and end on the upper horizontal side; (b) those that start on the lower horizontal side and end on the upper horizontal side; and (c) those that start on the lower horizontal side and end on the right hand vertical side. We can count the number ℓ of segments by counting the x-intercepts and y-intercepts. The x-intercept of a line L_w, $w = (w_0, w_1)$, is a point $k(q_0, -q_1) + Q(w_0, w_1)$ such that $Qw_1 - kq_1 = 0$. That is, $k = Qw_1/q_1$. The intercept is the x

coordinate, $Q(q_0w_1 + q_1w_0)/q_1$. But q_0 and q_1 are relatively prime, so as we let w vary all possible numbers of the form Qe/q_1 in the range of the x-coordinates along the length a side of B occur as intercepts. Thus the number of x-intercepts is between $aq_1/Q - 1$ and $aq_1/Q + 1$. Similarly, the number of y-intercepts is between $aq_0/Q - 1$ and $aq_0/Q + 1$. That is

$$\frac{a(q_0 + q_1)}{Q} - 2 \leq \ell \leq \frac{a(q_0 + q_1)}{Q} + 2.$$

Next we want to bound the sum of the variations in the second coordinates along the segments. For a lower bound, let

$$a' = \left\lfloor \frac{a|q_1|}{Q} \right\rfloor \frac{Q}{|q_1|}$$

be the largest integral multiple of $Q/|q_1|$ that is less than or equal to a. We can shrink B slightly to obtain an a' by a rectangle B' and just measure the parts of the segments in B'. Every segment of type (a) in B' matches up with a segment of type (c) in B' so that the sum of the differences in the second coordinate along the two segments is exactly a. If we call these combined segments and the segments of type (b) *super segments*, then the number of super segments in B' is the number of x-intercepts in B' and each super segment varies in its second coordinate by a'. Thus the number of super segments is at least

$$\left\lfloor \frac{a'}{Q/|q_1|} \right\rfloor = \left\lfloor \frac{a|q_1|}{Q} \right\rfloor \geq \frac{a|q_1|}{Q} - 1 = \frac{|q_1|}{N^{s/2}} - 1,$$

and the sum of the variation in the second coordinates is at least

$$m_0 = \frac{a|q_1|}{N^{s/2}} - a = \frac{Q|q_1|}{N^s} - \frac{Q}{N^{s/2}}.$$

It follows that the number of lattice points in B is at least

$$\begin{aligned} \frac{m_0}{|q_1|} - \ell &\geq \frac{Q}{N^s} - \frac{Q}{N^{s/2}|q_1|} - \frac{a(|q_0| + |q_1|)}{Q} - 2 \\ &= \frac{Q}{N^s} - \frac{Q}{N^{s/2}|q_1|} - \frac{|q_0| + |q_1|}{N^{s/2}} - 2 \end{aligned}$$

Next we show that we can choose the connection element q so the error term is bounded. For any v we have $vu/vq = u/q$, so we can multiply u and q by a unit and leave the number of lattice points unchanged. For general N and q, the best we can do is guarantee that $|q_0|, |q_1| \in O(Q)$, which results in a useless bound. We can do better when $N = 2$.

Proposition 17.2.3. *Suppose that $N = 2$. Let $q = q_0 + q_1\pi \in R$ with q_0 and q_1 relatively prime. Let $Q = |q_0^2 - 2q_1^2|$ be the absolute norm of q. Then there is a unit v in R such that $vq = q_0' + q_1'\pi$ with $|q_0'| < |q_1'| < Q^{1/2}$.*

Proof. The product of the elements $1 + \pi$ and $1 - \pi$ is -1, hence they are units. We have

$$(1 - \pi)(q_0 + q_1\pi) = q_0 - 2q_1 + (q_1 - q_0)\pi$$

and

$$(1 + \pi)(q_0 + q_1\pi) = q_0 + 2q_1 + (q_1 + q_0)\pi.$$

Suppose that $|q_0| > |q_1|$. Then by checking the four possible cases, we find that multiplying q by either $1 + \pi$ or $1 - \pi$ gives an element $r = r_0 + r_1\pi$ with $|r_0| < |q_1|$. After finitely many such multiplications we must obtain an element $q'q_0' + q_1'\pi$ with $|q_0'| < |q_1'|$. But then $Q = |q_0'^2 - 2q_1'^2| = |q_1'^2 + (q_1'^2 - q_0'^2)| > q_1'^2$, as desired. □

Thus we may assume that q has the form in the proposition. Also, we have $|q_1| \geq (Q/2)^{1/2}$.

Lemma 17.2.4. *If $(Q/N)^{1/2} \leq |q_1| \leq Q^{1/2}$ and $Q = |q_0^2 - Nq_1^2|$, then $(Q/|q_1|) + |q_0| + |q_1| \leq (N^{1/2} + (N-1)^{1/2} + 1)Q^{1/2}$. If $N = 2$, then this can be improved to $(Q/|q_1|) + |q_0| + |q_1| \leq 3Q^{1/2}$.*

Proof. We have $|q_1| \leq Q^{1/2}$,

$$\frac{Q}{|q_1|} \leq \frac{Q}{(Q/N)^{1/2}} = N^{1/2}Q^{1/2},$$

and

$$|q_0| = |(Nq_1^2 - Q)^{1/2}| \leq |(NQ - Q)^{1/2}| = (N-1)^{1/2}q^{1/2},$$

from which the first statement follows. The second statement is left as an exercise. □

It follows from Lemma 17.2.4 that if $(Q/N)^{1/2} \leq |q_1| \leq Q^{1/2}$, then

$$m_0 \geq \frac{Q}{2^s} - \frac{(N^{1/2} + (N-1)^{1/2} + 1)Q^{1/2}}{2^{s/2}} - 2.$$

If $N = 2$, then it follows from Proposition 17.2.3 and Lemma 17.2.4 that

$$m_0 \geq \frac{Q}{2^s} - \frac{3Q^{1/2}}{2^{s/2}} - 2.$$

A similar derivation of an upper bound gives the following theorems.

Theorem 17.2.5. *Suppose that $q \in \mathbb{Z}[\pi]$ is invertible modulo π, $\mathsf{N}_{\mathbb{Q}}^{F}(q)$ is prime, and π is primitive modulo q. Let* $\mathbf{a}$ *be an ℓ-sequence defined over $\mathbb{Z}[\pi]$, $\pi^2 = N$, whose connection element $q = q_0 + q_1\pi$ has absolute norm Q. If s is even, then the number K of occurrences of any s-tuple in one period of* $\mathbf{a}$ *satisfies*

$$\left| K - \frac{Q}{N^s} \right| \leq \frac{Q}{N^{s/2}|q_1|} + \frac{|q_0| + |q_1|}{N^{s/2}} + 2.$$

If $(Q/N)^{1/2} \leq |q_1| \leq Q^{1/2}$, then

$$\left| K - \frac{Q}{N^s} \right| \leq \frac{(N^{1/2} + (N-1)^{1/2} + 1)Q^{1/2}}{N^{s/2}} + 2.$$

If $N = 2$, then

$$\left| K - \frac{Q}{2^s} \right| \leq 3\left(\frac{Q}{2^s}\right)^{1/2} + 2.$$

Theorem 17.2.6. *Suppose that $q \in \mathbb{Z}[\pi]$ is invertible modulo π, for some unit u in R and $q_0, q_1 \in \mathbb{Z}$ we have $q = ur^t = q_0 + q_1\pi$, $T = \mathsf{N}_{\mathbb{Q}}^{F}(r)$ is prime, and π is primitive modulo q. Suppose also that for some unit u' in R and $q_0', q_1' \in \mathbb{Z}$ we have $u'r^{t-1} = q_0' + q_1'\pi$. Let* $\mathbf{a}$ *be an ℓ-sequence defined over $\mathbb{Z}[\pi]$, $\pi^2 = N$, whose connection element q has absolute norm Q. If s is even, then the number K of occurrences of any s-tuple in one period of* $\mathbf{a}$ *satisfies*

$$\left| K - \frac{Q(1-1/T)}{N^s} \right| \leq \frac{Q}{N^{s/2}|q_1|} + \frac{|q_0| + |q_1|}{N^{s/2}} + \frac{Q/T}{N^{s/2}|q_1'|} + \frac{|q_0'| + |q_1'|}{N^{s/2}} + 4.$$

If $(Q/N)^{1/2} \leq |q_1| \leq Q^{1/2}$ and $(Q/(TN))^{1/2} \leq |q_1'| \leq (Q/T)^{1/2}$, then

$$\left| K - \frac{Q(1-1/T)}{N^s} \right| \leq \frac{(N^{1/2} + (N-1)^{1/2} + 1)Q^{1/2}(1 + 1/T^{1/2})}{N^{s/2}} + 4.$$

If $N = 2$, then

$$\left| K - \frac{Q(1-1/T)}{2^s} \right| \leq 3\frac{Q^{1/2}(1 + 1/T^{1/2})}{2^{s/2}} + 4.$$

Proof. In this case $K = \Delta_0(q) - r\Delta_0(q')$. Here we have used the fact that multiplying a connection element by a unit has no effect on the set of sequences generated by AFSRs with the connection element. □

In fact, it is likely that in general the error is much smaller than this since this analysis assumes the maximum possible error along every line segment.

For $N > 2$, this bound is not in general very good since $|q_1|$ may be close to T. Thus Theorem 17.2.5 tells us that to obtain nearly uniform distributions, we should choose a connection element q with $(Q/N)^{1/2} \leq |q_1| \leq Q^{1/2}$.

When $d > 2$, a similar analysis can be tried. However, the results are not as good for two reasons. First, the errors are larger — it is possible to give a heuristic argument that the error estimates from this approach will be at best $O((Q/N^s)^{(d-1)/d})$. Second, the difference between the degree of F over $\mathbb{Q}$ and the rank of the unit group is larger, so it is harder to find a connection element equivalent to q whose components are small.

17.3 Arithmetic correlations

Let $\mathbf{a} = a_0, a_1, \cdots$ be an N-ary sequence and let a be the associated π-adic number. Let τ be a nonnegative integer. As before, if $\mathbf{b} = b_0, b_1, \ldots$ is another N-ary sequence, denote by $\mathbf{b}^\tau$ the shift of $\mathbf{b}$ by τ steps, that is, $b_i^\tau = b_{i+\tau}$. If b is the π-adic integer corresponding to $\mathbf{b}$, let $b^{(\tau)}$ denote the π-adic integer corresponding to $\mathbf{b}^\tau$.

Because of Lemma 5.5.6 we can generalize the notion of arithmetic cross-correlation from Section 11.3.c to the setting of d-FCSRs as follows. As before, we let ζ be a primitive Nth root of unity in the complex numbers. We let χ be the character of $\mathbb{Z}/(N)$ defined by $\chi(u) = \zeta^u$. If $\mathbf{a}$ is eventually periodic and for each $u = 0, 1, \cdots, N-1$, μ_u is the number of occurrences of u in one complete period of $\mathbf{a}$, then the imbalance is

$$Z(a) = Z(\mathbf{a}) = \sum_{u=0}^{N-1} \mu_u \chi(u).$$

Definition 17.3.1. *Let* $\mathbf{a}$ *and* $\mathbf{b}$ *be two eventually periodic* N*-ary sequences with period* T*. Let* a *and* $b^{(\tau)}$ *be the* π*-adic numbers whose coefficients are given by* $\mathbf{a}$ *and* $\mathbf{b}^\tau$*, respectively. Then the sequence of coefficients associated with* $a - b^{(\tau)}$ *is eventually periodic and its period divides* T*. The* shifted d-arithmetic cross-correlation *(or just the* arithmetic cross-correlation *when* d *is clear from the context) of* $\mathbf{a}$ *and* $\mathbf{b}$ *is*

$$\mathcal{C}^{(d)}_{\mathbf{a},\mathbf{b}}(\tau) = Z(a - b^{(\tau)}). \tag{17.4}$$

When $\mathbf{a} = \mathbf{b}$*, the arithmetic cross-correlation is called the* d-arithmetic autocorrelation *(or just the* arithmetic cross-correlation*) of* $\mathbf{a}$ *and is denoted* $\mathcal{A}^{(d)}_{\mathbf{a}}(\tau)$*.*

The sequences $\mathbf{a}, \mathbf{b}$ (of period T) has *ideal d-arithmetic correlations* if,

$$C^{(d)}_{\mathbf{a},\mathbf{b}}(\tau) = \begin{cases} T & \text{if } \mathbf{a} = \mathbf{b}^\tau \\ 0 & \text{otherwise} \end{cases}$$

for each τ with $0 \leq \tau < T$. A family of sequences has ideal d-arithmetic correlations if every pair of sequences in the family has ideal d-arithmetic correlations. Now we focus on the binary case, $N = 2$.

If $\mathbf{a} = a_0, a_1, \cdots$ is a sequence and $k \neq 0$ is an integer, then recall that the sequence $\mathbf{b} = b_0, b_1, \ldots$ is the k-fold decimation of $\mathbf{a}$ if $b_i = a_{di}$ for every i. Let $\mathbf{a}$ be a d-ℓ-sequence. Let $\mathcal{F}_{\mathbf{a}}$ be the family of all k-fold decimations of $\mathbf{a}$, where $k \geq 1$ is allowed to vary over all integers which are relatively prime to the period of $\mathbf{a}$.

Theorem 17.3.2. *Let $\pi^d = 2$. Let $\mathbf{a}$ be a d-ℓ-sequence with connection integer $q \in \mathbb{Z}[\pi]$. Suppose that $(q) = (r^t)$ for some $r \in \mathbb{Z}[\pi]$ such that $n = \mathsf{N}_{\mathbb{Q}}^F(r)$ is prime. Suppose also that*

$$\gcd(d, \phi(\mathsf{N}_{\mathbb{Q}}^F(q))) = 1.$$

Then the family $\mathcal{F}_{\mathbf{a}}$ has ideal d-arithmetic correlations.

The remainder of this section consists of a proof of Theorem 17.3.2. Let $Q = \mathsf{N}_{\mathbb{Q}}^F(q) = \pm n^t$. We first need a constraint on the sequences that can occur as d-ℓ-sequences.

Proposition 17.3.3. *Under the hypotheses of Theorem 17.3.2, the second half of one period of $\mathbf{a}$ is the bitwise complement of the first half.*

Proof. We have

$$\mathbb{Z}[\pi]/(q) \cong \mathbb{Z}/(Q),$$

via an isomorphism $\psi : \mathbb{Z}/(Q) \to \mathbb{Z}[\pi]/(q)$. Therefore

$$\pi^{\phi(|Q|)} \equiv 1 \pmod{q}.$$

That is,

$$q | \pi^{n^{t-1}(n-1)} - 1 = (\pi^{n^{t-1}(n-1)/2} - 1)(\pi^{n^{t-1}(n-1)/2} + 1).$$

These two factors differ by 2, which is a power of π, so r can only divide one of them. Thus q divides exactly one of them. Since π is primitive modulo q, q cannot divide the first factor, hence it divides the second factor. It follows that

$$\pi^{\phi(Q)/2} \equiv -1 \pmod{q}.$$

Let $m = \psi^{-1}(\pi)$. Note that $Q \pmod 2 = 1$. By Corollary 9.5.2 we have

$$\begin{aligned} a_{i+\phi(Q)/2} &= (zm^{-i-\phi(Q)/2} \pmod{Q}) \pmod 2 \\ &= (-zm^{-i} \pmod{Q}) \pmod 2 \\ &= (Q - (zm^{-i} \pmod{Q})) \pmod 2 \\ &= 1 - ((zm^{-i} \pmod{Q}) \pmod 2) \\ &= 1 - a_i \end{aligned}$$

for some integer z. □

The above property extends to decimations of d-ℓ-sequences.

Lemma 17.3.4. *Let* $\mathbf{a}$ *be a binary sequence of even period* T *whose second half is the complement of its first half. Let* $k > 0$ *be relatively prime to* T*, and let* $\mathbf{b}$ *be a* k*-fold decimation of* $\mathbf{a}$*. Then the second half of one period of* $\mathbf{b}$ *is the complement of the first half.*

Proof. Note that k must be odd and $a_j = 1 - a_{j+T/2}$. Thus we have

$$\begin{aligned} b_i &= a_{ik} = 1 - a_{ik+T/2} \\ &= 1 - a_{(i+T/2)k} = 1 - b_{i+T/2}. \end{aligned}$$ □

Lemma 17.3.5. *Let* T *be an even integer. Let* $\mathbf{b} = b_0, b_1, \ldots$ *be a strictly periodic sequence with period* T *such that the second half of each period is the bitwise complement of the first half and let*

$$b = \sum_{i=0}^{\infty} b_i \pi^i.$$

Then

$$b = \frac{v}{\pi^{T/2} + 1}$$

for some $v \in \mathbb{Z}[\pi]$.

Conversely, if $\mathbf{b} = b_0, b_1, \ldots$ *is the* π*-adic expansion of a* π*-adic integer* $w/(\pi^{T/2} + 1)$ *with eventual period dividing* T*,* $\gcd(T, d) = 1$ *and* $w \notin \{0, (\pi^{T/2} + 1)/(\pi - 1)\}$*, then the second half of each* T *bit period of* $\mathbf{b}$ *is the bitwise complement of the first half.*

Proof. If a is a π-adic number, then we let $\bar{a}$ be the complementary π-adic number. That is, we replace each 1 by 0 and each 0 by 1 in the π-adic expansion of a. For the first part, we have

$$\begin{aligned} b + \bar{b} &= 1 + \pi + \pi^2 + \cdots \\ &= \frac{-1}{\pi - 1} \\ &= -(1 + \pi + \pi^2 + \cdots + \pi^{d-1}) \end{aligned}$$

since $\pi^d = 2$. Therefore

$$\begin{aligned} -(1 + \pi + \pi^2 + \cdots + \pi^{d-1}) - b &= \bar{b} \\ &= c + \pi^{T/2} b \end{aligned}$$

for some $c \in \mathbb{Z}[\pi]$. Solving for b we have

$$b = \frac{-c - 1 - \pi - \cdots - \pi^{d-1}}{\pi^{T/2} + 1},$$

as claimed. Now consider the converse. We first reduce to the case when $\mathbf{b}$ is strictly periodic. We can write

$$\frac{w}{\pi^{T/2}+1} = c + \pi^k \gamma$$

with $c \in \mathbb{Z}[\pi]$, k a positive integer, and γ a π-adic integer with strictly periodic π-adic expansion. Since the π-adic expansion of γ is periodic, we can write $\gamma = u/(\pi^T - 1)$ for some $u \in \mathbb{Z}[\pi]$. Thus

$$(w - c(\pi^{T/2}+1))(\pi^{T/2}-1) = \pi^k u. \tag{17.5}$$

In particular, π^k divides the left hand side of equation (17.5). It follows that π divides $w - c(\pi^{T/2}+1)$, since $\pi^{T/2}-1$ is congruent to -1 modulo π. Note that this requires a careful argument since $\mathbb{Z}[\pi]$ may not be a unique factorization domain. By induction, π^k divides $w - c(\pi^{T/2}+1)$. Thus we can write $w - c(\pi^{T/2}+1) = \pi^k z$ for some $z \in \mathbb{Z}[\pi]$. Therefore

$$\gamma = \frac{z}{\pi^{T/2}+1}$$

and we may assume that $\mathbf{b}$ is strictly periodic.

Let

$$\mu = \pi^{(d-1)T/2} - \pi^{(d-2)T/2} + - \cdots - \pi^{T/2} + 1,$$

so $\mu(\pi^{T/2}+1) = \pi^{dT/2} + 1 = 2^{T/2} + 1$. Then

$$\begin{aligned} \frac{w}{\pi^{T/2}+1} &= \frac{w\mu}{2^{T/2}+1} \\ &= \frac{v_0}{2^{T/2}+1} + \frac{v_1}{2^{T/2}+1}\pi + \cdots + \frac{v_{d-1}}{2^{T/2}+1}\pi^{d-1} \end{aligned}$$

for some integers $v_0, v_1, \ldots, v_{d-1}$. Since $\mathbf{b}$ is strictly periodic, we have

$$-(2^{T/2}+1) \le v_j \le 0$$

for all j.

We claim that if any of the v_j is zero, then they all are, and that if any of the v_j is $-(2^{T/2}+1)$, then they all are. For any j we have

$$\frac{v_j}{2^{T/2}+1} = \sum_{i=0}^{\infty} b_{di+j} 2^i.$$

Fix a particular j. Since T and d are relatively prime, we can choose $n, m \in \mathbb{Z}$ so that $nT + j = md$, and we can assume that n and m are nonnegative. Then by the fact that $\mathbf{b}$ is periodic with period dividing T, for every i

$$\begin{aligned} b_{di+j} &= b_{di+nT+j} \\ &= b_{d(i+m)}. \end{aligned}$$

It follows that

$$\begin{aligned}\frac{v_j}{2^{T/2}+1} &= \sum_{i=0}^{\infty} b_{d(i+m)}2^i \\ &= 2^{-dm}\sum_{i=m}^{\infty} b_{di}2^i \\ &= 2^{-dm}\left(\frac{v_0}{2^{T/2}+1} - c\right),\end{aligned}$$

where c is an ordinary integer between 0 and $2^m - 1$. That is, $2^{dm}v_j = v_0 - (2^{T/2}+1)c$. It follows that if $2^{T/2}+1$ divides one of v_j and v_0, then it divides the other. If $v_0 = 0$ and $v_j = -(2^{T/2}+1)$, then $c = 2^{dm}$, which is impossible. If $v_0 = -(2^{T/2}+1)$ and $v_j = 0$, then $c = -1$, which is also impossible. The claim follows.

Now, by hypothesis the v_j are not all zero and are not all $-(2^{T/2}+1)$. Hence by the above claim, none is zero and none is $-(2^{T/2}+1)$. In particular there is an exponential representation for $v_0/(2^{T/2}+1)$, $b_{di} = (-v_0 2^{-i} \pmod{2^{T/2}+1}) \pmod 2$. Since d is odd,

$$\begin{aligned}b_{di+T/2} &= b_{d(i+T/2)} \\ &= (-v_0 2^{-i-T/2} \pmod{2^{T/2}+1}) \pmod 2 \\ &= (v_0 2^{-i} \pmod{2^{T/2}+1}) \pmod 2 \\ &= 1 - b_{di}.\end{aligned}$$

Now let m be arbitrary and pick i and k so $m = id + kT$ and $i \geq 0$. Then

$$\begin{aligned}b_{m+T/2} &= b_{id+kT+T/2} \\ &= b_{id+T/2} \\ &= 1 - b_{id} \\ &= 1 - b_{m-kT} \\ &= 1 - b_m,\end{aligned}$$

which proves the lemma. □

Theorem 17.3.2 is an immediate consequence of the following computation.

Theorem 17.3.6. *Let the hypotheses be as in Theorem 17.3.2. Let d and d' be relatively prime to $\phi(Q)$ with $1 \leq d, d' < \phi(Q)$. Let $\mathbf{b}$ and $\mathbf{c}$ be d and d'-fold decimations of $\mathbf{a}$, respectively. Then the d-arithmetic cross-correlation of $\mathbf{b}$ and $\mathbf{c}$ with shift τ is*

$$C^{(d)}_{\mathbf{b},\mathbf{c}}(\tau) = \begin{cases} \phi(Q) & \text{if } \mathbf{b} = \mathbf{c}^\tau \\ 0 & \text{otherwise} \end{cases}$$

Proof. Let $T = \phi(Q)$. Let $\mathbf{b}$ and $\mathbf{c}$ have associated π-adic numbers $b = u/(\pi^{T/2}+1)$ and $c = v/(\pi^{T/2}+1)$, respectively, with $u, v \in R$. These elements have this form by Lemma 17.3.5. The shift of $\mathbf{c}$ by τ corresponds to a π-adic integer $\pi^{T-\tau}c + e$ for some $e \in \mathbb{Z}[\pi]$. The d-arithmetic cross-correlation of $\mathbf{b}$ and $\mathbf{c}$ with shift τ is the number of zeros minus the number of ones in one length T period of

$$\begin{aligned} b - (\pi^{T-\tau}c + e) &= \frac{u - \pi^{T-\tau}v - e(\pi^{T/2}+1)}{\pi^{T/2}+1} \\ &\stackrel{\text{def}}{=} \frac{w}{\pi^{T/2}+1} \\ &\stackrel{\text{def}}{=} f. \end{aligned}$$

If $\mathbf{c}$ is a shift of $\mathbf{b}$ with shift τ, then $w = 0$ and the result follows.

Suppose $\mathbf{c}$ is not a shift of $\mathbf{b}$ with shift τ. Let $\mathbf{f}$ be the sequence associated to f. It suffices to show that any period of $\mathbf{f}$ is balanced. The element w is nonzero and by Lemma 5.5.6 the coefficient sequence of $w/(\pi^{T/2}+1) \in \mathbb{Z}_\pi$ has eventual period dividing T. The theorem then follows from the second part of Lemma 17.3.5. □

Remark. It is not in general true that the coefficient sequence of a π-adic integer of the form $w/(\pi^{T/2}+1)$ has eventual period dividing T. For example, if we take $w = -\pi^{T/2} - 1$, then the period is d. However, the π-adic integer $w/(\pi^{T/2}+1)$ that arises in the preceding proof is the difference of two π-adic integers whose coefficient sequences have period dividing T.

17.4 Exercises

1. Prove the second statement in Lemma 17.2.4.

2. Find the distribution of 2-tuples and 3-tuples in the sequences studied in exercises 9.7.1 and 9.7.2. Is this consistent with what you would expect?

3. Let $R = \mathbb{Z}[\pi]$ with $\pi^2 = 2$ and $S = \{0, 1\}$. Let $q = 5 + 2\pi$. Show that any nonzero periodic sequence with minimal connection element q is an ℓ-sequence of period 16. Describe the distribution of 4-tuples in such a sequence.

4. Let $R = \mathbb{Z}[\pi]$ with $\pi^2 = 3$. Find a connection element for a 3-ary 2-ℓ-sequence with period 22 in this setting. Note that 3 is not primitive modulo 23 in $\mathbb{Z}$, so there is no period 22 3-ary ordinary ℓ-sequence.

Part IV

Register Synthesis and Security Measures

Chapter 18 Register Synthesis and LFSR Synthesis

18.1 Sequence generators and the register synthesis problem

If $\mathcal{F}$ is any class of sequence generators, then, loosely speaking, the *register synthesis problem for* $\mathcal{F}$ is the problem of finding the smallest generator in $\mathcal{F}$ that outputs a sequence $\mathbf{a}$ given only a prefix of $\mathbf{a}$. In order to analyze the complexity of synthesis algorithms it is necessary to have precise understandings of the concepts of a class of generators and the size of a generator and of what it means for a register synthesis algorithm to be successful.

Recall from Section 5.1.c that a *sequence generator* $F = (U, \Sigma, f, g)$ is a discrete state machine with output, consisting of a set U of *states*, an alphabet Σ of output values, a state change function $f : U \to U$ and an output function $g : U \to \Sigma$. For a given initial state $u_0 \in U$, such a sequence generator determines an infinite sequence $F(u_0) = (a_0, a_1, \cdots)$ of elements of Σ, defined recursively by $a_n = g(u_n)$ and $u_{n+1} = f(u_n)$ for $n = 0, 1, \cdots$. If the set of states U is finite then the output sequence is eventually periodic.

Let us say that a *class* $\mathcal{F}$ of sequence generators is a set of generators that share a common alphabet Σ. We usually ask that the class $\mathcal{F}$ have a notion of the *size* of each generator $F \in \mathcal{F}$, a real number which reflects the number of symbols used to represent a state of F. If $\mathbf{a}$ is a (finite or infinite) sequence over an alphabet Σ and if $\mathcal{F}$ is a class of sequence generators with alphabet Σ then the *$\mathcal{F}$-span of* $\mathbf{a}$ is

$$\lambda^{\mathcal{F}}(\mathbf{a}) = \inf_{F \in \mathcal{F}} \text{size}(F)$$

where the infimum is taken over all sequence generators F that output the sequence $\mathbf{a}$.

Definition 18.1.1. *A* register synthesis algorithm *for a class $\mathcal{F}$ of sequence generators with output alphabet Σ is an algorithm A which, given a finite string $\mathbf{b}$ with symbols in Σ, outputs a sequence generator $F \in \mathcal{F}$ and a state s_0 of F such that $\mathbf{b}$ is a prefix of $F(s_0)$.*

The Berlekamp-Massey algorithm is a register synthesis algorithm for the class of linear feedback shift registers (LFSRs). For this class, the size function is simply the number of cells in the shift register. Several register synthesis algorithms are also known for the class of FCSRs. These ar described in the next few chapters.

18.2 LFSRs and the Berlekamp-Massey algorithm

18.2.a Linear span

In this section we consider sequences and shift registers over a field F. Let $\mathbf{a} = (a_0, a_1, \cdots)$ be a sequence (with $a_i \in F$) and suppose that it can be generated by a LFSR. The *linear span* or *linear complexity*, $\lambda(\mathbf{a})$, is the number of cells in the shortest LFSR that generates $\mathbf{a}$. (So $\lambda(\mathbf{a}) = \lambda^{\mathcal{F}}(\mathbf{a})$ is the $\mathcal{F}$-span of $\mathbf{a}$ where $\mathcal{F}$ is the class of linear feedback shift registers over the field F.) If no LFSR can generate $\mathbf{a}$ (i.e., if $\mathbf{a}$ is not eventually periodic), then $\lambda(\mathbf{a}) = \infty$.

Proposition 18.2.1. *Let* $\mathbf{a}$ *be a sequence with* $m = \lambda(\mathbf{a}) < \infty$. *Then the LFSR of length* m *that generates* $\mathbf{a}$ *is uniquely determined by any string of* $2m$ *consecutive symbols* $a_k, a_{k+1}, \cdots, a_{k+2m-1}$.

Proof. For convenience we may take $k = 0$. We search for feedback coefficients $q_1, q_2, \cdots, q_m$ so that

$$\begin{aligned} a_m &= q_1 a_{m-1} + q_2 a_{m-2} + \cdots + q_m a_0 \\ a_{m+1} &= q_1 a_m \quad + q_2 a_{m-1} + \cdots + q_m a_1 \\ a_{m+2} &= q_1 a_{m+1} + q_2 a_m \quad + \cdots + q_m a_2 \end{aligned}$$

and so on. This is an infinite collection of linear equations in m unknowns, so we will need at least m of these in order to determine $q_1, q_2, \cdots, q_m$. Therefore we will need to use the $2m$ coefficients $a_0, a_1, \cdots, a_{2m-1}$. These m equations will have a unique solution if and only if the determinant of the coefficient matrix

$$\begin{pmatrix} a_{m-1} & a_{m-2} & \cdots & a_0 \\ a_m & a_{m-1} & \cdots & a_1 \\ \cdots & & & \\ a_{2m-1} & a_{2m-2} & \cdots & a_{m-1} \end{pmatrix}$$

is nonzero, or equivalently, if the rows $R_1, R_2, \cdots, R_m$ of this matrix are linearly independent. But suppose a relation of linear dependence exists, say $c_1 R_1 + c_2 R_2 + \cdots + c_m R_m = 0$. If $c_m \neq 0$ we may solve for R_m as a linear combination of the other rows, $R_m = q_1' R_{m-1} + q_2' R_{m-2} + \cdots + q_{m-1}' R_1$ whence

$$\begin{aligned} a_{m-1} &= q_1' a_{m-2} + q_2' a_{m-3} + \cdots + q_{m-1}' a_0 \\ a_m &= q_1' a_{m-1} + q_2' a_{m-2} + \cdots + q_{m-1}' a_1 \\ a_{m+1} &= q_1' a_m \quad + q_2' a_{m-1} + \cdots + q_{m-1}' a_2 \end{aligned}$$

(and so on). This is a linear recurrence of length $m-1$ for the sequence $\mathbf{a}$, contradicting the minimality of m. If $c_m = 0$ then the same argument applies to the last nonzero coefficient c_k, giving a linear recurrence of length $k - 1 < m$ for the sequence $\mathbf{a}$. □

18.2.b The Berlekamp-Massey algorithm

In this section we study the register synthesis problem for linear feedback shift registers over a field F. Given a sequence $\mathbf{a} = (a_0, a_1, \cdots)$ with (finite) linear span m, the algorithm determines the list of coefficients $q_1, \cdots, q_m$ for the smallest linear feedback shift register that generates $\mathbf{a}$ (or, equivalently, the smallest linear recurrence satisfied by $\mathbf{a}$.) It does so using only $2m$ terms in the sequence which (according to Proposition 18.2.1) is optimal.

The algorithm was first discovered in 1968 by Berlekamp [6] as a decoding algorithm for BCH codes. Massey [133] then reformulated Berlekamp's algorithm as a tool for cryptanalyzing stream ciphers. It was later discovered ([147], [191], [26], [35]) that this algorithm essentially computes the continued fraction (see Section 5.7.b) expansion of the power series associated to $\mathbf{a}$. The algorithm runs in quadratic time in the length of its input because it is *adaptive*: when a new symbol of the sequence $\mathbf{a}$ is discovered, the algorithm updates the current guess (for the smallest LFSR) at low cost, rather than starting over from the beginning. Such a strategy is a great advantage from the cryptanalytic point of view, and it is a common one for register synthesis algorithms, see Section 19.3 and Chapter 20.

The Berlekamp-Massey algorithm can most easily be described in terms of the power series model (see Section 6.4) for LFSR sequences. Let $a(x) = \sum_{i=0}^{\infty} a_i x^i$ be the generating function associated with $\mathbf{a}$. Suppose $\mathbf{a}$ can be generated by an LFSR with connection polynomial $q(x)$. According to Theorem 6.4.1, there is a polynomial $f(x)$ in $R[x]$ so that $a(x) = f(x)/q(x)$. Equivalently,

$$q(x)a(x) = f(x). \tag{18.1}$$

Equation (18.1) is sometimes called the *key equation*. The size of the LFSR is

$$\Phi(f, q) = \max(\deg(f) + 1, \deg(q)).$$

Thus $\lambda(\mathbf{a})$ is the minimum over all f, q with $a(x) = f/q$ of $\Phi(f, q)$. We remark that

$$\Phi(f, q) + \Phi(f', q') \geq \Phi(fq' + qf', qq') \tag{18.2}$$

for any $f, f', q, q' \in R[x]$. The LFSR synthesis problem can be rephrased as follows:

Given a prefix $a_0, a_1, \cdots, a_{k-1}$ of $\mathbf{a}$.
Find a pair (f, q) that minimizes $\Phi(f, q)$ among all polynomials f, q that satisfy the key equation.

Let us say that a pair $(f(x), q(x))$ of polynomials form a *degree i approximation* to $a(x)$ if

$$q(x)a(x) \equiv f(x) \pmod{x^i}. \tag{18.3}$$

At the ith stage the Berlekamp-Massey algorithm, having read the symbols $a_0, a_1, \cdots, a_{i-1}$, finds a degree i approximation (f_i, q_i) for which the function $\Phi(f, q)$ is minimal. To process the next symbol, a_i, there are two possibilities. If

$$q_i(x)a(x) \equiv f_i(x) \pmod{x^{i+1}},$$

then the best approximation modulo x^i is also the best approximation modulo x^{i+1}. Thus we simply set $f_{i+1} = f_i$ and $q_{i+1} = q_i$. Otherwise we say a *discrepancy* occurs at stage i. In this case we need to compute a new approximation. There exists $b \in F$, $b \neq 0$, so that

$$q_i(x)a(x) \equiv f_i(x) + bx^n \pmod{x^{i+1}}.$$

Suppose a discrepancy also occurred at some earlier stage, say, the mth stage. Then

$$q_m(x)a(x) \equiv f_m(x) + cx^m \pmod{x^{m+1}},$$

for some $c \in F$, $c \neq 0$. Let

$$f_{i+1} = f_i - (b/c)x^{i-m}f_m \quad \text{and} \quad q_{i+1} = q_i - (b/c)x^{i-m}q_m.$$

Then we obtain a degree $i+1$ approximation because

$$q_{i+1}(x)a(x) \equiv f_{i+1}(x) \pmod{x^{i+1}}.$$

It might not, however, correspond to the shortest LFSR that generates the segment $a_0, a_1, \cdots, a_i$ of $\mathbf{a}$. To guarantee that it does, we must carefully choose m. In the Berlekamp-Massey algorithm, m is taken to be the most recent stage at which $\Phi(f_i, q_i)$ changed. The Berlekamp-Massey algorithm is given in Figure 18.1.

Let us say that a natural number i is a *turning point* if it occurs as a value of m in the algorithm. That is, i is a turning point[1] if

$$\Phi(f_{i+1}, q_{i+1}) > \Phi(f_i, q_i). \tag{18.4}$$

Theorem 18.2.2 says that at each stage the Berlekamp-Massey algorithm generates a Φ-minimizing approximation. If i is a turning point then there is a unique such approximation. If $i \geq 2\lambda(\mathbf{a})$ (λ = linear span) then this approximation is exact: it generates the whole sequence $\mathbf{a}$.

Theorem 18.2.2. *Let $\mathbf{a} = a_0, a_1, \cdots$ and let $a(x) \in K[[x]]$ be its generating function. Let (f_i, q_i) be the output of the Berlekamp-Massey algorithm at stage $i \geq 1$. Then (f_i, q_i) is a degree i approximation to $a(x)$. Suppose (f, q) is another degree i approximation to $a(x)$. Then*

$$\Phi(f_i, q_i) \leq \Phi(f, q). \tag{18.5}$$

If $\Phi(f_i, q_i) = \Phi(f, q)$ and if i is a turning point then $f_i/q_i = f/q$. If $i \geq 2\lambda(\mathbf{a})$, then $\Phi(f_i, q_i) = \lambda(\mathbf{a})$ and $f_i(x)/q_i(x) = a(x)$.

[1] The normalized span $\Phi(f_i, q_i)/i$ is decreasing as long as $\Phi(f_i, q_i)$ is constant. When $\Phi(f_i, q_i)$ changes, the normalized span goes from small to large.

BERLEKAMP_MASSEY$(a_0, \cdots, a_{n-1})$

begin
if all $a_i = 0$ **then**
    return(0,1)
fi
$a(x) = \sum_{i=0}^{n-1} a_i x^i$
Let m be minimal with $a_m \neq 0$
$f_m(x) = 0$
$q_m(x) = 1$
$f_{m+1}(x) = x^m$
$q_{m+1}(x) = \begin{cases} 1 + x^m & \text{if } m > 0 \\ 1 & \text{else} \end{cases}$
$c = a_m$
for $i = m + 1$ to $n - 1$ **do**
    Let $a(x)q_i(x) - f_i(x) \equiv bx^i \pmod{x^{i+1}}$
    if $b = 0$ **then**
        $f_{i+1}(x) = f_i(x)$
        $q_{i+1}(x) = q_i(x)$
    else
        $f_{i+1}(x) = f_i(x) - (b/c)x^{i-m} f_m(x)$
        $q_{i+1}(x) = q_i(x) - (b/c)x^{i-m} q_m(x)$
        if $\Phi(f_{i+1}, q_{i+1}) > \Phi(f_i, q_i)$ **then**
            $m = i$
            $c = b$
    fi fi
    $i = i + 1$
od
return (f_n, q_n)
end

Figure 18.1: The Berlekamp-Massey Algorithm.

The first statement follows from the preceding discussion. We need several lemmas before we prove the other statements in Theorem 18.2.2.

Lemma 18.2.3. *Let $f(x)$ and $q(x)$ be polynomials. Suppose that x^{i+1} divides $a(x)q(x) - f(x)$ but does not divide $a(x)q_i(x) - f_i(x)$ (so a discrepancy occurs at stage i). Then $\Phi(f, q) \geq i+1-\Phi(f_i, q_i)$.*

Proof. We have

$$\frac{f}{q} - \frac{f_i}{q_i} = \frac{fq_i - f_iq}{qq_i} = \frac{bx^i}{qq_i}$$

for some $b \neq 0$. Thus, using equation (18.2),

$$i + 1 \leq \Phi(fq_i - f_iq, qq_i) \leq \Phi(f, q) + \Phi(f_i, q_i). \qquad \square$$

Lemma 18.2.4. *If m is a turning point then $\Phi(f_{m+1}, q_{m+1}) = m + 1 - \Phi(f_m, q_m)$.*

Proof. It is straightforward to check that the lemma is true for $m = 0$, so we proceed by induction. Assume the lemma is true for some turning point m', and let m be the next turning point. Since m is a turning point there exists $u \in F$ so that:

$$\begin{aligned} f_{m+1}(x) &= f_m(x) - ux^{m-m'} f_{m'}(x) \\ q_{m+1}(x) &= q_m(x) - ux^{m-m'} q_{m'}(x). \end{aligned}$$

But m' and m are consecutive turning points so

$$\begin{aligned} \Phi(f_m, q_m) &= \Phi(f_{m'+1}, q_{m'+1}) \\ &= m' + 1 - \Phi(f_{m'}, q_{m'}), \end{aligned}$$

using the induction hypothesis. Then

$$\begin{aligned} \Phi(f_{m+1}, q_{m+1}) &= \Phi(x^{m-m'} f_{m'}(x), x^{m-m'} q_{m'}(x)) \\ &= m - m' + \Phi(f_{m'}, q_{m'}) \\ &= m - m' + m' + 1 - \Phi(f_{m'+1}, q_{m'+1}) \\ &= m + 1 - \Phi(f_m, q_m). \qquad \square \end{aligned}$$

Proof of Theorem 18.2.2. Suppose that $f(x), q(x)$ are polynomials such that

$$\Phi(f, q) < \Phi(f_i(x), q_i(x)) \quad \text{and} \quad a(x)q(x) \equiv f(x) \pmod{x^i}.$$

First assume that $i - 1$ is a turning point. We have

$$a(x)q_{i-1}(x) \equiv f_{i-1}(x) \pmod{x^{i-1}},$$

but not modulo x^i. Also,

$$\Phi(f,q) < \Phi(f_i,q_i) = i - \Phi(f_{i-1},q_{i-1})$$

by Lemma 18.2.4, and

$$\Phi(f,q) \geq i - \Phi(f_{i-1},q_{i-1}),$$

by Lemma 18.2.3, which is a contradiction. Now suppose $i-1$ is not a turning point. By induction we may assume that $\Phi(f_{i-1},q_{i-1})$ is minimal for stage $i-1$. Then

$$\Phi(f,q) < \Phi(f_i,q_i) = \Phi(f_{i-1},q_{i-1}) \leq \Phi(f,q),$$

which again is a contradiction, hence equation (18.5) holds.

For the uniqueness statement, suppose that i is a turning point and that $f,q \in F[x]$ is a degree i approximation with $\Phi(f,q) = \Phi(f_i,q_i)$. Then there exists $b,c \in F$ $(b \neq 0)$ so that

$$q_i(x)a(x) \equiv f_i(x) + bx^i \pmod{x^{i+1}} \quad \text{and} \quad q(x)a(x) \equiv f(x) + cx^i \pmod{x^{i+1}}.$$

Let

$$f' = (c/b)f_i - f \quad \text{and} \quad q' = (c/b)q_i - q.$$

If (f,q) is not a multiple of (f_i,q_i) then (f',q') is a degree $i+1$ approximation to $a(x)$ with $\Phi(f',q') \leq \Phi(f_i,q_i)$. But equations (18.4) and (18.5) give

$$\Phi(f',q') \geq \Phi(f_{i+1},q_{i+1}) > \Phi(f_i,q_i),$$

which is a contradiction.

Finally, suppose that $a(x)q(x) = f(x)$, so that $\lambda(\mathbf{a}) = \Phi(f(x),q(x))$. If $i \geq 2\lambda(\mathbf{a})$, then

$$\frac{f(x)}{q(x)} - \frac{f_i(x)}{q_i(x)} = \frac{f(x)q_i(x) - f_i(x)q(x)}{q(x)q_i(x)} = \frac{b(x)x^i}{q(x)q_i(x)},$$

for some $b(x) \in F[x]$. Suppose $b(x) \neq 0$. By equation (18.5) and (18.2),

$$2\lambda(\mathbf{a}) < i+1 \leq \Phi(b(x)x^i, q(x)q_i(x)) \leq \Phi(f,q) + \Phi(f_i,q_i) \leq 2\Phi(f,q) = 2\lambda(\mathbf{a})$$

which is a contradiction. Hence $b(x) = 0$ and $f_i(x)/q_i(x) = a(x)$. □

18.2.c Complexity of the Berlekamp-Massey algorithm

At each iteration of the loop in the Berlekamp-Massey algorithm it is necessary to compute

$$\delta_{i+1}(x) = a(x)q_{i+1}(x) - f_{i+1}(x).$$

However, it is not necessary to carry out a full polynomial multiplication since we either have

$$\delta_{i+1}(x) = \delta_i(x) \quad \text{or} \quad \delta_{i+1}(x) = \delta_i(x) - (b/c)x^{i-m}\delta_m(x).$$

In either case, $\delta_{i+1}(x)$ can be computed in linear time in n (since $a(x)$, $q_i(x)$, and $f_i(x)$ all have degree at most n). Hence the overall time complexity is $O(n^2)$.

18.2.d Continued fractions and the Berlekamp-Massey algorithm

"The Berlekamp-Massey algorithm is equivalent to the continued fraction expansion for Laurent series" (folklore). In this section we will explain in more precise terms the meaning of this often-heard sentence. We refer to Section 5.7 for background on continued fractions. The continued fractions (CF) algorithm and the Berlekamp-Massey (BM) algorithm both provide a sequence of approximations, which are rational functions $f_n(x)/q_n(x)$, to a Laurent series. One might guess that each BM approximation $f_n(x)/q_n(x)$ is equal to the corresponding CF approximation, but this is not true. In the first place, as explained in Section 5.7.b, the CF expression approximates a *reciprocal* Laurent series, that is, a Laurent series in x^{-1}, whereas the BM algorithm provides (rational function) approximations for Laurent series in x. Consequently, the polynomials in the CF approximations must be replaced by their reciprocal polynomials (cf. Section 5.3) to obtain the BM approximations. Next, it turns out that each stage of the CF algorithm corresponds to many stages of the BM algorithm, and finally there is a shift of degree by one. With these caveats, the two approximations become identical, as described in the following theorem. Our proof of Theorem 18.2.5 differs from other proofs in the literature in that it is not computational, but relies instead on the uniqueness properties of the CF and BM approximations as described in Theorems 5.7.6 and 18.2.2.

Let $\mathbf{a} = a_0, a_1, a_2, \cdots$ be a sequence of elements of a field F. Let

$$\begin{aligned} \widehat{a}(x) &= a_0 + a_1 x + a_2 x^2 + \cdots \\ a(x) &= a_0 x^{-1} + a_1 x^{-2} + a_2 x^{-3} + \cdots \end{aligned}$$

be "the" Laurent series and reciprocal Laurent series associated to this sequence.

Theorem 18.2.5. *Let $(u_i(x), v_i(x))$ be the output of the BM algorithm at stage i for the input $\widehat{a}(x)$. Let $f_n(x)/q_n(x)$ be the nth convergent associated with the CF expansion of $a(x)$. Let $e_n = \deg(q_n)$ and define*

$$\widehat{f}_n(x) = x^{e_n - 1} f_n(1/x) \quad \text{and} \quad \widehat{q}_n(x) = x^{e_n} q_n(1/x).$$

Then there exists $0 \neq c \in F$ so that $\widehat{f}_n(x) = c u_{2e_n}(x)$ and $\widehat{g}_n(x) = c v_{2e_n}(x)$, and in particular,

$$\frac{\widehat{f}_n(x)}{\widehat{q}_n(x)} = \frac{u_{2e_n}(x)}{v_{2e_n}(x)} \in F[[x]].$$

Proof. First we check that $\widehat{f}_n$ is in fact a polynomial. Since $a(x)$ has no constant term, it follows from Proposition 5.7.5 that $\deg(f_1) < \deg(q_1)$ and by induction that $\deg(f_n) < \deg(q_n) = e_n$. Consequently $\widehat{f}_n = x^{e_n - 1} f_n(1/x)$ is a polynomial (no terms of negative degree) of degree at most $e_n - 1$ while $\widehat{q}_n(x)$ is a polynomial of degree at most e_n. In particular, $\Phi(\widehat{f}_n, \widehat{q}_n) \leq e_n$.

If the remainder r_n (at the nth stage of the CF expansion for $a(x)$) vanishes then $f_n(x)/q_n(x) = a(x)$ so $\widehat{f}_n(x)/\widehat{q}_n(x) = \widehat{a}(x)$. Consequently $\lambda(a_0, a_1, \cdots) \le e_n$ and Theorem 18.2.2 implies that

$$\frac{u_{2e_n}}{v_{2e_n}} = \frac{\widehat{f}_n}{\widehat{q}_n} = \widehat{a}(x).$$

So we may assume that $r_n \neq 0$. According to Theorem 5.7.6, the nth convergent satisfies

$$f_n(1/x) \equiv q_n(1/x)a(1/x) \pmod{x^{e_n+e_{n+1}}}$$

in the power series ring $F[[x]]$. Since $r_n \neq 0$, equation (5.24) says that $e_{n+1} > e_n$, so

$$\frac{\widehat{f}_n(x)}{\widehat{q}_n(x)} \equiv \widehat{a}(x) = x^{-1}a(1/x) = a_0 + a_1x + a_2x^2 + \cdots. \pmod{x^{2e_n+1}}.$$

In other words, $(\widehat{f}_n, \widehat{q}_n)$ is a degree $2e_n + 1$ approximation to $\widehat{a}(x)$. Now let (u_{2e_n}, v_{2e_n}) be the polynomials constructed by BM at stage $2e_n$, so that

$$\frac{u_{2e_n}(x)}{v_{2e_n}(x)} \equiv \widehat{a}(x) \pmod{x^{2e_n}}.$$

By Theorem 18.2.2,

$$\Phi(u_{2e_n}, v_{2e_n}) \le \Phi(\widehat{f}_n, \widehat{q}_n) \le e_n$$

so

$$\deg(v_{2e_n}) \le e_n \quad \text{and} \quad \deg(u_{2e_n}) \le e_n - 1.$$

Going backwards, let

$$\widehat{u}_{2e_n}(x) = x^{e_n-1}u_{2e_n}(1/x) \quad \text{and} \quad \widehat{v}_{2e_n}(x) = x^{e_n}v_{2e_n}(1/x).$$

Then $\widehat{u}_{2e_n}$ and $\widehat{v}_{2e_n}$ are polynomials, with $\deg(\widehat{v}_{2e_n}) \le e_n$, and

$$\frac{\widehat{u}_{2e_n}(x)}{\widehat{v}_{2e_n}(x)} \equiv a(x) = x^{-1}\widehat{a}(1/x) \pmod{x^{-2e_n-1}}$$

in $F[[x^{-1}]]$. But according to Theorem 5.7.6, the convergent f_n/q_n is the unique such approximation, so

$$\frac{\widehat{u}_{2e_n}}{\widehat{v}_{2en}} = \frac{f_n}{q_n}.$$

Therefore

$$\frac{u_{2e_n}}{v_{2e_n}} = \frac{\widehat{f}_n}{\widehat{q}_n}.$$

□

Algorithm	Ring	series	convergent	approximation
CF	$F[[x^{-1}]]$	$a(x)$	f_n/g_n	$\widehat{u}_{2e_n}/\widehat{v}_{2e_n}$
BM	$F[[x]]$	$\widehat{a}(x)$	$\widehat{f}_n/\widehat{g}_n$	u_{2e_n}/v_{2e_n}

Table 18.1: Translation between continued fraction convergents and Berlekamp-Massey approximations.

18.3 Blahut's theorem

In this section we describe a remarkable connection between the linear span of a sequence and the discrete Fourier transform. The result is credited to R. Blahut by J. Massey in [134] because it appears implicitly in [10]. See also [136] for a self-contained exposition. We use standard facts from Sections 2.3 and 3.2.h concerning the Fourier transform.

Theorem 18.3.1. *Let F be a field and let $\mathbf{a} = (a_0, a_1, \cdots)$ be a periodic sequence, with period T, and symbols $a_i \in F$. Define $f : \mathbb{Z}/(T) \to F$ by $f(k) = a_k$. Assume either (a) the field F has characteristic zero or (b) $\mathrm{char}(F)$ does not divide T. Then the linear span of $\mathbf{a}$ is equal to the number of nonzero discrete Fourier coefficients of the function f.*

Proof. Let $a(x) = \sum_{i=0}^{\infty} a_i x^i$ be the generating function of the sequence $\mathbf{a}$. Then

$$a(x) = a'(x)(1 + x^T + x^{2T} + \cdots) = \frac{a'(x)}{1 - x^T} \in F[[x]]$$

where $a'(x) = a_0 + a_1 x + \cdots + a_{T-1}x^{T-1}$ corresponds to a single period of $\mathbf{a}$. We claim the Fourier coefficients of f which vanish are in one to one correspondence with the Tth roots of unity that are also roots of $a'(x)$ as follows. The assumptions on $\mathrm{char}(F)$ imply the existence of an extension field K of F that contains T distinct Tth roots of unity, see Section 3.2.d. Let $\zeta \in K$ be a primitive Tth root of unity and let $\chi_1 : \mathbb{Z}/(T) \to K$ be the resulting character, $\chi_1(k) = \zeta^k$. All the other characters are of the form $\chi_m = \chi_1^m$ so the mth Fourier coefficient of f is

$$\widehat{f}(\chi_m) = \sum_{k \in \mathbb{Z}/(T)} f(k)\chi_1^m(k) = \sum_{k=0}^{T-1} a_k(\zeta^m)^k = a'(\zeta^m) \tag{18.6}$$

which proves the claim (and gives a way to compute $\widehat{f}(\chi_m)$).

Now let $q(x)$ be the connection polynomial of the smallest LFSR that outputs the sequence $\mathbf{a}$. By Theorem 6.4.1 there exists a polynomial $g(x)$ with $\deg(g) < \deg(q)$ such that

$$\frac{a'(x)}{1 - x^T} = \frac{g(x)}{q(x)} \quad \text{or} \quad a'(x)q(x) = (1 - x^T)g(x). \tag{18.7}$$

The polynomial $q(x)$ divides $1 - x^T$ and $\gcd(q(x), g(x)) = 1$ since $q(x)$ is minimal. It follows that $q(x)$ and $g(x)$ have no common roots in the extension field K (otherwise, by Theorem 3.2.9, they would share an irreducible factor). Consequently the roots of $q(x)$ and the roots of $a'(x)$ are distinct: if $c \in K$ were a common root then it would also be a double root of the right side of (18.7), so it would be a root of $g(x)$ which is a contradiction. Therefore the linear span of $\mathbf{a}$ is $\lambda(\mathbf{a}) = \deg(q) = T - E$ where E is the above number of Tth roots of unity that are also roots of $a'(x)$. That is, E is the number of Fourier coefficients of f which vanish. Therefore $T - E$ is the number of non-vanishing Fourier coefficients. □

18.4 The Günther-Blahut theorem

Blahut's Theorem only applies to sequences whose period T is relatively prime to the characteristic of the field containing the elements of the sequence. If this is not the case, then the roots of $x^T - 1$ are not distinct, so we cannot define the appropriate discrete Fourier coefficients. However, Massey and Serconek have developed a generalization (the GDFT) of the discrete Fourier transform that can be used to analyze the linear span of an arbitrary periodic sequence over a finite field. [136, 137]. They refer to the result as the Günther-Blahut theorem because it is equivalent to a result by Günther [70]. It makes use of the Hasse derivative [76] (called the *hyperderivative* in [123]). Let F be a field.

18.4.a The Hasse derivative

Let $f(x) = \sum_i f_i x^i \in F[x]$ be a polynomial. The jth (formal) derivative of f is the polynomial

$$f^{(j)}(x) = \sum_i i(i-1)\cdots(i-j+1) f_i x^{i-j} = j! \sum_i \binom{i}{j} f_i x^{i-j}.$$

If $\mathrm{char}(F) = 0$ then the multiplicity of a root $b \in F$ of f is the least i such that $f^{(i)}(b) \neq 0$. This is false when $\mathrm{char}(F) = p > 0$ and in fact $f^{(j)} = 0$ for all $j \geq p$. The jth *Hasse derivative* ([76], [68], [182]) of f is defined by

$$f^{[j]}(x) = \sum_i \binom{i}{j} f_i x^{i-j}.$$

Lemma 18.4.1. *The following properties of the Hasse derivative hold.*

1. *If $f, g \in F[x]$ and $i \geq 0$, then $(f+g)^{[i]} = f^{[i]} + g^{[i]}$.*
2. *If $f_1, \cdots, f_t \in F[x]$, then*
$$(f_1 \cdots f_t)^{[i]} = \sum f_1^{[i_1]} \cdots f_t^{[i_t]},$$
where the sum is extended over all t-tuples $(i_1, \cdots, i_t)$ with $i_j \geq 0$ and $i_1 + \cdots + i_t = i$.

3. *For any $c \in F$ and $t \in \mathbb{Z}^+$, if $f(x) = (x-c)^t$ then*

$$f^{[i]}(x) = \binom{t}{i}(x-c)^{t-i}.$$

4. *For $0 \leq n \leq t$ and $g, h \in F[x]$, we have*

$$(gh^t)^{[n]} = uh^{t-n}$$

for some $u \in F[x]$ with $\deg(u) \leq \deg(g) + n(\deg(h)-1)$.

5. *Let $f, h \in F[x]$. Suppose that h is irreducible and $h^{[1]} \neq 0$. Let $m \geq 1$. Then h^m divides f if and only if h divides each of $f^{[0]} = f, f^{[1]}, \cdots, f^{[m-1]}$.*

Proof. Parts (1) through (4) are left to the exercises. For part (5), first suppose that $f = gh^m$. Then by part (4), $f^{[n]} = uh^{m-n}$ for some $u \in F[x]$ with $\deg(u) \leq \deg(g) + n(\deg(h)-1)$. Thus if $n < m$, then h divides $f^{[n]}$.

Conversely, suppose that $f = gh^k$ with $k < m$ and $\gcd(g,h) = 1$. Then by part (2) we have

$$f^{[k]} = \sum g^{[j]} h^{[i_1]} \cdots h^{[i_k]},$$

where the sum is extended over all $k+1$-tuples $(j, i_1, \cdots, i_k)$ with $j, i_\ell \geq 0$ and $j + i_1 + \cdots + i_k = k$. Thus in each term either some $i_j = 0$, or $j = 0$ and $i_1 = \cdots = i_k = 1$. The former type of term is divisible by h. The remaining term is $g(h^{[1]})^k \neq 0$. We have $\deg(h^{[1]}) < \deg(h)$ and h is irreducible, so h does not divide $h^{[1]}$. Nor does it divide g. Thus h does not divide $f^{[k]}$, which proves part (5). □

18.4.b The GDFT and Günther weight

There are various notions of "generalized discrete Fourier transform", used in various circumstances [11], [138]. In this section we review the construction of Günther, Massey and Serconek [70], [136], [137]. Let $\mathbf{a} = a_0, a_1, \cdots$ be a sequence over F of period $T = np^v$ for positive integers n and v, with $p = \mathrm{char}(F)$ relatively prime to n. Associate to this sequence the polynomial $f(x) = a_0 + a_1 x + \cdots + a_{N-1}x^{N-1}$. By passing to an extension field if necessary, we may assume that F contains n distinct nth roots of unity. Let $b \in F$ be a primitive nth root of unity.

Definition 18.4.2. *The* generalized discrete Fourier transform *of the sequence* $\mathbf{a}$ *is the matrix*

$$\widehat{\mathbf{a}} = GDFT_b(a_0, \cdots, a_{N-1}) = \begin{bmatrix} f(1) & f(b) & \cdots & f(b^{n-1}) \\ f^{[1]}(1) & f^{[1]}(b) & & f^{[1]}(b^{n-1}) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ f^{[p^v-1]}(1) & f^{[p^v-1]}(b) & & f^{[p^v-1]}(b^{n-1}) \end{bmatrix}.$$

If $v = 0$ then the GDFT of $\mathbf{a}$ coincides with its discrete Fourier transform. The following theorem says that the GDFT is indeed a "transform", meaning that it has an inverse [137]. The proof is delayed until Section 18.4.c.

Theorem 18.4.3. *If the field* F *contains* n *distinct* n*th roots of unity then the GDFT determines a one to one correspondence between* T*-periodic sequences* $\mathbf{a}$ *and matrices of size* $n \times p^v$*.*

The *Günther weight* of a matrix is the number of entries that are nonzero or lie below a nonzero entry. The following theorem is an appropriate generalization of Blahut's theorem to the case when $\text{char}(F) = p > 0$ divides the period, T, of the sequence.

Theorem 18.4.4. *(Günther-Blahut Theorem) Let* $\mathbf{a}$ *be a periodic sequence of period* $T = p^v n$ *over a finite field with characteristic* p*, where* p *does not divide* n*. Then the linear span of* $\mathbf{a}$ *is equal to the Günther weight of the GDFT of* $\mathbf{a}$*.*

Proof. As above, the GDFT of $\mathbf{a}$ is fabricated from the polynomial $f(x) = 1 + a_1 x + \cdots + a_{T-1} x^{T-1}$. Let $m_i \geq 0$ denote the order of b^i as a root of f. Consider the ith column of the GDFT. By Lemma 18.4.1 part (5), this column begins with $\min\{m_i, p^v\}$ zeroes so it contributes $p^v - \min\{p^v, m_i\}$ to the Günther weight w. Hence the Günther weight of the GDFT of f is

$$w = T - \sum_{i=0}^{n-1} \min\{p^v, m_i\}.$$

As in the proof of Blahut's theorem (Section 18.3), the generating function of $\mathbf{a}$ may be written:

$$a_0 + a_1 x + a_2 x^2 + \cdots = \frac{f(x)}{1 - x^T} = \frac{g(x)}{q(x)} \tag{18.8}$$

or $f(x)q(x) = (1 - x^T)g(x)$, where g, q are relatively prime and $\deg(q)$ is the linear span of $\mathbf{a}$. Recall from Theorem 3.2.13 that every root of $1 - x^T$ is an nth root of unity and is repeated with

multiplicity p^v. If such a root b^i is also a root of q then it is not a root of g so its multiplicity as a root of q is $p^v - m_i$ (where as above, m_i denotes its multiplicity as a root of f). If b^i is not a root of q then its multiplicity as a root of f may be greater than p^v. Therefore

$$\deg(q) = \sum_{i=0}^{n-1} (p^v - \min\{p^v, m_i\}) = w. \qquad \square$$

18.4.c Proof of Theorem 18.4.3

We follow [137]. The *Hasse matrix* $H_k(x)$ over the field F is the $k \times k$ matrix whose i,jth entry is the $(i-1)$st Hasse derivative of the monomial x^{j-1}. That is, the i,jth entry is $\binom{j-1}{i-1} x^{j-i}$ if $j \geq i$ and is 0 if $j < i$.

Lemma 18.4.5. *The Hasse matrix $H_k(x)$ is invertible and its inverse is $H_k(-x)$.*

Proof. The (i,j)th entry in $H_k(x)H_k(-x)$ is

$$\sum_{m=1}^{k} \binom{m-1}{i-1} x^{m-i} \binom{j-1}{m-1} (-x)^{j-m} = x^{j-i} \sum_{m=1}^{k} \binom{m-1}{i-1} \binom{j-1}{m-1} (-1)^{j-m}.$$

When $i > j$, one of the two binomial coefficients in each summand is zero. When $i = j$, the only summand that is nonzero is the one for which $m = i$, and the result is then 1. When $i < j$, if we let $u = m - i$, then this becomes

$$(-x)^{j-i} \sum_{u=0}^{j-i} \binom{j-1}{i-1} \binom{j-i}{u} (-1)^u = 0$$

after recombining terms of the binomial coefficients. Thus $H_k(x)H_k(-x)$ is the identity matrix. $\square$

To prove Theorem 18.4.3 we need to recover the original sequence $\mathbf{a}$ (or equivalently, the polynomial $f(x)$) from its GDFT. For $0 \leq j \leq p^v - 1$, associate to the sequence $\mathbf{a}$ the function

$$A_{(j)}(x) = a_j + a_{j+p^v} x + \cdots + a_{j+(n-1)p^v} x^{n-1},$$

so that

$$f(x) = A_{(0)}(x^{p^v}) + A_{(1)}(x^{p^v}) x + \cdots + A_{(p^v-1)}(x^{p^v}) x^{p^v-1}. \tag{18.9}$$

The coefficients of the polynomial $A_{(j)}(x)$ are the first n symbols in the jth phase of the decimation of $\mathbf{a}$ by p^v.

Lemma 18.4.6. *Combining these polynomials into a column vector, we have:*

$$\begin{bmatrix} f(x) \\ f^{[1]}(x) \\ . \\ . \\ . \\ f^{[p^v-1]}(x) \end{bmatrix} = H_{p^v}(x) \begin{bmatrix} A_{(0)}(x^{p^v}) \\ A_{(1)}(x^{p^v}) \\ . \\ . \\ . \\ A_{(p^v-1)}(x^{p^v}) \end{bmatrix}. \tag{18.10}$$

Proof. Let $A_{(j,p^v)}(x) = A_{(j)}(x^{p^v})$. Take the ith Hasse derivative of equation (18.9) and use part 2 of Lemma 18.4.1 to obtain

$$f^{[i]}(x) = \sum_{j=0}^{p^v-1} \sum_{\ell=0}^{i} \binom{j}{\ell} x^{j-\ell} A^{[i-\ell]}_{(j,p^v)}(x).$$

A typical term has a factor of

$$\binom{tp^v}{s}$$

with $0 \le t \le n$ and $0 \le s = i - \ell \le i < p^v$. We claim that this binomial coefficient is zero if $\ell < i$. That is, it is zero if $1 \le s < p^v$. Indeed, it can be written in the form

$$\binom{tp^v}{s} = \frac{tp^v}{s} \cdot \frac{tp^v - 1}{1} \cdot \frac{tp^v}{2} \cdots \frac{tp^v - (s-1)}{s-1}.$$

Now count the powers of p dividing the binomial coefficient. In the first term, the numerator is divisible by a higher power of p than the denominator. In the subsequent terms, since $s < p^v$ each numerator and corresponding denominator are divisible by the same power of p. Thus the binomial coefficient is a multiple of p, hence is zero as claimed. It follows that

$$f^{[i]}(x) = \sum_{j=0}^{p^v-1} \binom{j}{i} x^{j-i} A_{(j,p^v)}(x),$$

which is what the matrix equation says. □

The left hand side of equation (18.10) is a column vector which, when evaluated at b^{i-1}, coincides with the ith column of the GDFT of $\mathbf{a}$. So applying the inverse of the Hasse matrix to this equation gives

$$H_{p^v}(-b^i) \begin{bmatrix} f(b^i) \\ f^{[1]}(b^i) \\ . \\ . \\ . \\ f^{[p^v-1]}(b^i) \end{bmatrix} = \begin{bmatrix} A_{(0)}(c^i) \\ A_{(1)}(c^i) \\ . \\ . \\ . \\ A_{(p^v-1)}(c^i) \end{bmatrix} \tag{18.11}$$

(for $1 \le i \le n$) where $c = b^{p^v}$. So it suffices to show that the n columns on the right side of equation (18.11) determine the polynomials $A_{(0)}, A_{(1)}, \cdots, A_{(p^v-1)}$. Placing these columns side by side so as to make a matrix, and observing that c is a primitive nth root of unity in F (since $\gcd(p^v, n) = 1$), we see that the jth row of the resulting matrix is precisely the DFT (with respect to c) of the n-periodic sequence $a_j, a_{j+p^v}, \cdots, a_{j+(n-1)p^v}$ of coefficients of $A_{(j)}$. But the DFT is invertible, hence $A_{(j)}$ is uniquely determined by this data. This completes the proof of Theorem 18.4.3. □

18.5 Generating sequences with large linear span

In considering the security of keystream generators, it is customary to make one of several assumptions concerning the knowledge of the attacker. A stream cipher is considered very weak if it can be broken simply with knowledge of the plaintext. In other cases one assumes that the attacker knows part of the keystream and needs to predict the rest. The attacker may even know the architecture of the generator, and need to find the particular parameters (e.g. connection polynomial) and key. Correlation attacks are analyzed under the assumption that the attacker has full knowledge of the specific generator and only needs to find the initial state. In each of these cases, specific attacks have been developed, and specific pseudo-random sequences have been analyzed for their vulnerabilities.

The existence of the Berlekamp-Massey algorithm implies that sequences with low linear span are useless from a cryptologic point of view. High linear span is a necessary condition for security, but it is by no means sufficient. A sequence each of whose periods consists of $T - 1$ zeros followed by a single one has linear span T. The termwise sum of this sequence with an m-sequence of the same length will give a sequence with excellent statistical properties and large linear span, but it is still insecure because any message of length less than T has been encoded with a sequence with low linear span. The *summation combiner* (see sect. 18.5.d) is a keystream generator with high linear span and good statistical properties, but it can be attacked by other means, in this case using an FCSR synthesis algorithm. Many keystream generators with large linear span are vulnerable to other types of attack such as correlation attacks, a topic not addressed in this book.

Nevertheless, it is a major goal of stream cipher research to find efficient generators of sequences with statistical properties that are similar to those of m-sequences, but with high linear span. In this section we briefly describe, mostly without proof, some the approaches that have been tried. General references for this section include [42], [55], and [169].

18.5.a Linear registers

One might hope to generate sequences with high linear span using a *linear register* (see Section 10.6), that is, a sequence generator

$$A \circlearrowright V \xrightarrow{\text{f}} F.$$

where V is a k-dimensional vector space over a field F, $f : V \to F$ is a linear mapping and $A : V \to V$ is a linear (but not necessarily a shift) mapping. For any initial state $v \in V$ the output sequence is $f(v), f(A(v)), f(A(A(v)), \cdots$ as in Section 5.1.c. However, according to Theorem 10.6.2, there exists a linear *feedback* shift register of the same or smaller size, k, that outputs all the same sequences, and in particular the linear span of each such sequence is at most k. This means that linear modifications to LFSRs are of no help cryptologically. If we wish to construct sequences with large linear spans then we must use nonlinearity.

18.5.b Nonlinear filters

A common method for generating new pseudorandom sequences from an m-sequence is to apply a *feed forward function* or *nonlinear filter* f to the state of the LFSR as in Figure 18.2. In this way a sequence may be obtained with the same period as the original m-sequence but with greater linear span. General references for this section include [169] and [21].

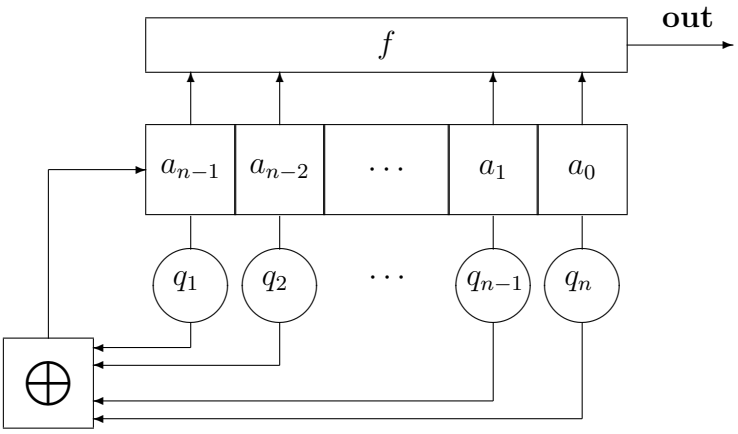


Figure 18.2: LFSR with feedforward function.

A well known result of Key [89] gives an upper bound on the linear span of the resulting sequence. As pointed out in [136] this result is most easily proven using Blahut's theorem. The *Hadamard product* $\mathbf{c} = \mathbf{a} \cdot \mathbf{b}$ of sequences $\mathbf{a}, \mathbf{b}$ is the termwise product, $\mathbf{c} = (a_0 b_0, a_1 b_1, \cdots)$.

Theorem 18.5.1. *([89], [136]) Let* $\mathbf{a} = (a_0, a_1, \cdots)$ *be a binary m-sequence of order* n*. Let* $\mathbf{c} = \mathbf{a} \cdot \mathbf{a}[\tau]$ *be the Hadamard product of* $\mathbf{a}$ *with the* τ*-shift of* $\mathbf{a}$*. Then its linear span satisfies:* $\lambda(\mathbf{c}) \leq n + \binom{n}{2}$*. More generally, if* $\mathbf{c} = \mathbf{c}^{(1)} + \mathbf{c}^{(2)} + \cdots + \mathbf{c}^{(s)}$ *is the termwise sum of* s *sequences, each of which is a Hadamard product of degree* $\leq d$ *of shifts of the sequence* $\mathbf{a}$*, then*

$$\lambda(\mathbf{c}) \leq \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{d}. \tag{18.12}$$

Returning to the case of the Hadamard product $\mathbf{c} = \mathbf{a} \cdot \mathbf{a}[\tau]$*, if* n *is prime then* $\lambda(\mathbf{c}) = n + \binom{n}{2}$*.*

Proof. Let us first consider the case $\mathbf{c} = \mathbf{a} \cdot \mathbf{a}[\tau]$. We can assume that $\mathbf{a} = (\mathsf{Tr}(\alpha^0), \mathsf{Tr}(\alpha^1), \cdots)$ where $\alpha \in F$ is a primitive elment, $F = \mathbb{F}_{2^n}$, and $\mathsf{Tr} : F \to \mathbb{F}_2$ is the trace mapping. By Blahut's theorem (Section 18.3), it suffices to count the number of nonzero Fourier coefficients of the sequence $\mathbf{c}$. This sequence is T-periodic, where $T = 2^n - 1$. Moreover, $\alpha \in F$ is a primitive Tth root of unity and in fact, the choice of α determines an isomorphism $F^\times \cong \mathbb{Z}/(T)$. Therefore the linear span of $\mathbf{c}$ is the number of values for m (with $0 \leq m \leq T-1$) such that $f(\alpha^m) \neq 0$ where, as in equation (18.6),

$$f(x) = c_0 + c_1 x + \cdots + c_{T-1} x^{T-1}$$

with $c_i = a_i a_{i+\tau}$. Let $A = \alpha^\tau$ and compute,

$$\begin{aligned} f(\alpha^m) &= \sum_{i=0}^{T-1} \mathsf{Tr}(\alpha^i)\mathsf{Tr}(\alpha^{i+\tau})\alpha^{mi} \\ &= \sum_{y \in F^\times} \mathsf{Tr}(y)\mathsf{Tr}(Ay)y^m \\ &= \sum_{a=0}^{n-1}\sum_{b=0}^{n-1}\sum_{y \in F^\times} y^{2^a}(Ay)^{2^b} y^m \\ &= \sum_{b=0}^{n-1} A^{2^b} \sum_{a=0}^{n-1}\sum_{y \in F^\times} y^{m+2^a+2^b}. \end{aligned}$$

The inner sum vanishes unless $m + 2^a + 2^b \equiv 0 \pmod{T}$, so $T - m = 2^a + 2^b$. If it does not vanish then a, b are uniquely determined (up to reordering) by m. Therefore the number of nonzero Fourier coefficients is no more than the number of possible values of $2^a + 2^b$ with $0 \leq a, b \leq n-1$, which is $n + \binom{n}{2}$.

Similarly, for a Hadamard product of degree $\leq d$ the mth Fourier coefficient is

$$f(\alpha^m) = \sum_{a_1=0}^{n-1}\sum_{a_2=0}^{n-1}\cdots\sum_{a_d=0}^{n-1} C(a_1, a_2, \cdots, a_d) \sum_{y \in F^\times} y^{m+2^{a_1}+2^{a_2}+\cdots+2^{a_d}}$$

(where C is a certain function that is straightforwardly determined). Again, the inner sum vanishes unless $T - m = 2^{a_1} + 2^{a_2} + \cdots + 2^{a_d}$ has dyadic weight $\leq d$. This gives the estimate (18.12) since the right hand side is simply the number of positive integers with dyadic weight $\leq d$. If $\mathbf{c} = \mathbf{c}^{(1)} + \mathbf{c}^{(2)} + \cdots + \mathbf{c}^{(s)}$ is the termwise sum of several such sequences, then for each of its component sequences $\mathbf{c}^{(j)}$ it is always the same Fourier coefficients $f(\alpha^m)$ that may fail to vanish: those for which the dyadic weight of $T - m$ is $\leq d$. Therefore the same estimate (18.12) continues to hold.

Returning to the case of a single product, $\mathbf{c} = \mathbf{a} \cdot \mathbf{a}[\tau]$, consider the above formula for $f(\alpha^m)$. Even if we suppose that $T - m$ can be written in the form $T - m = 2^u + 2^v$, the Fourier coefficient $f(\alpha^m)$ may still vanish. Assuming $T - m$ has this form (say, with $u \leq v$) we obtain

$$f(\alpha^m) = (A^{2^u} + A^{2^v})T = TA^{2^u}(1 + A^{2^v - 2^u})$$

Thus, $f(\alpha^m)$ vanishes if and only if

$$0 = 1 + A^{2^v - 2^u} = 1 + \alpha^{\tau(2^v - 2^u)}$$

which can only hold if $2^v - 2^u$ divides T. Since T is odd we must have $u = 0$ and $2^v - 1$ divides $T = 2^n - 1$. If n is prime then no such v exists (see Table 3.6). □

According to Key's theorem, a necessary condition for high linear span is that the degree of the feedforward function f should be high. But high degree alone does not guarantee high linear span. It is not known how to construct filter generators with maximal linear span where the feedforward function f has few terms. In fact, lower bounds on the linear span are difficult to come by. We give two examples.

Theorem 18.5.2. *([8], [167], [169]) Let* $\mathbf{a}$ *be a binary m-sequence of order* n*. Let* $\mathbf{c}$ *be the Hadamard product of* s *consecutive* τ*-shifts of* $\mathbf{a}$*. Assume* $\gcd(\tau, 2^n - 1) = 1$*. Then* $\lambda(\mathbf{c}) \geq \binom{n}{s}$.

Let $F = \mathbb{F}_{2^n}$. Recall from Section 13.4 that a function $f : F \to \mathbb{F}_2$ is *bent* if all of its (complex) Fourier coefficients $\widehat{f}(\chi)$ have the same magnitude $|\widehat{f}(\chi)|$. For such a function and a primitive element $\alpha \in F$ the resulting sequence $\mathbf{a} = (f(\alpha^0), f(\alpha^1), \cdots)$ is a *bent sequence* of order n.

Theorem 18.5.3. *([114]) Let* $\mathbf{a}$ *be a Bent sequence of order* n*. Assume* 4 *divides* n*. Then*

$$\lambda(\mathbf{a}) \geq 2^{n/4} \binom{n/2}{n/4} \sim 2\frac{2^{n/2}}{\sqrt{\pi n}}.$$

Another approach to nonlinear filtering is to take an LFSR over one field $\mathbb{F}_q$ and apply a feed forward function $f : \mathbb{F}_q \to \mathbb{F}_r$ with values in a different field $\mathbb{F}_r$. The function f might even be defined just on the rightmost cell, producing a sequence $b_i = f(a_i)$, where $\mathbf{a}$ is an m-sequence over

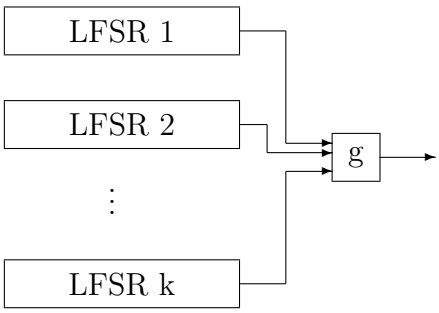


Figure 18.3: Nonlinear Combiner

$\mathbb{F}_q$. If q is a power of r, then f can be represented as a polynomial. In the case when $r = 2$ and f is even, results of Herlestam [77] and Brynielsson show that if $f(x) = \sum_{i=0}^{q-1} c_i x^i$, then

$$\lambda(\mathbf{b}) \;=\; \sum_{c_i \neq 0} \lambda(\mathbf{a})^{||i||} \leq q^{\log_2(\lambda(\mathbf{a})+1)}.$$

To be cryptographically strong, a register of this size would need a linear span close to $q^{\lambda(\mathbf{a})}$, so this method is not very good. Chan and Games showed that if q is odd and $r = 2$, then $\lambda(\mathbf{b})$ can be made as large as $q^{\lambda(\mathbf{a})-1}$. In this case the sequences also have optimal autocorrelations (and in some cases low cross-correlations). However, that if one considers these sequences as sequences over $GF(q)$ (that simply happen to have only two values), then the linear span is low and, by exploiting the imbalance of the sequences, the parameter q can even be found with a probabilistic attack [91].

Filter generators are also vulnerable to correlation attacks, although we do not address this topic here.

18.5.c Nonlinear combiners

A nonlinear combiner takes the outputs from a set of k LFSRs and combines them with a nonlinear function $g : \mathbb{F}_q^k \to \mathbb{F}_q$. We denote the output from the jth LFSR by

$$\mathbf{a}^j = a_0^j, a_1^j, \cdots.$$

Let n_i be the period of $\mathbf{a}^i$. The output $\mathbf{b}$ from the nonlinear combiner is the sequence whose ith element is $b_i = g(a_i^1, \cdots, a_i^k)$. A diagram of a nonlinear combiner is given in Figure 18.3.

The *Geffe generator* is a special case using three binary LFSRs (so $q = 2$). The output function is $g(a^1, a^2, a^3) = a^1a^3 + a^2(1 - a^3)$. The period of the Geffe generator is $n_1n_2n_3$ and the linear span is

$$\lambda(\mathbf{a}^1)\lambda(\mathbf{a}^3) + \lambda(\mathbf{a}^2)(1 + \lambda(\mathbf{a}^3)).$$

Another special case is the binary *threshold generator*. This generator combines k LFSR sequences by outputting a 1 if and only if a majority of the outputs are 1. Thus

$$g(a^1, \cdots, a^k) = \begin{cases} 1 & \text{if } \sum_{i=1}^{k} a^i > k/2 \\ 0 & \text{otherwise.} \end{cases}$$

In case $k = 3$ the period is $n_1n_2n_3$ and the linear span is

$$\lambda(\mathbf{a}^1)\lambda(\mathbf{a}^2) + \lambda(\mathbf{a}^2)\lambda(\mathbf{a}^3) + \lambda(\mathbf{a}^1)\lambda(\mathbf{a}^3).$$

In the general case (even over nonbinary fields) it is known that if $\rho(\mathbf{a})$ denotes the period of a periodic sequence $\mathbf{a}$, then

$$\rho(\mathbf{b}) \leq \text{lcm}(\rho(\mathbf{a}^1), \cdots, \rho(\mathbf{a}^k)).$$

The linear span is bounded by

$$\lambda(\mathbf{b}) \leq g^{\mathbb{Z}}(\lambda(\mathbf{a}^1), \cdots, \lambda(\mathbf{a}^k)),$$

where $g^{\mathbb{Z}}$ is g thought of as a polynomial over the integers [170] with coefficients in $\{0, 1\} \subset \mathbb{Z}$. Key [89] showed further that these inequalities become equalities when the $\mathbf{a}^j$ are m-sequences with relatively prime periods.

18.5.d Summation combiner

Combiner generators are vulnerable to correlation attacks. This was first observed by Siegenthaler. Rueppel proposed the summation combiner (or addition with carry generator) as a solution to this problem [168, 169]. In its simplest form, the summation combiner accepts as input two information bits a, b and one "memory" bit m. It outputs the information bit $c = a + b + m \pmod 2$ while modifying the memory to $m' = a + b + c \text{ (div } 2)$. For use in a stream cipher, the inputs are m-sequences $(a_0, a_1, \cdots)$ and $(b_0, b_1, \cdots)$. If m_i denotes the state of the memory at the ith clock cycle then the output stream is $c_i = a_i + b_i + c_i \pmod 2$ and $m_{i+1} = a_i + b_i + m_i \text{ (div } 2)$. In the language of Section 5.4 the bit streams $\mathbf{a}, \mathbf{b}$ are added as 2-adic integers.

Rueppel showed that if the periods of the m-sequences are relatively prime, then the period of the combined sequence is the product of the periods of the component sequences. He further gave a heuristic argument that the linear span of the the combined sequence is close to its period. However, this keystream generator is vulnerable to an FCSR synthesis attack (cf. Chapter 19). There are also successful correlation attacks against it.

18.5.e Clock-controlled generators

Nonlinearity can be introduced by irregularly clocking the generator. A survey of these clock-controlled generators as of 1989 was given by Gollman and Chambers [67]. In a simple case, two LFSRs over a finite field F are used: L_1 and L_2 of sizes n_1 and n_2, respectively. We are also given a function

$$f : F \to \mathbb{Z}.$$

At each step we change the state of L_1 once and extract $f(t)$ where t is the current output of L_1. Register L_2 then changes state $f(t)$ times and the output from the last state change is taken as the next output of the clock controlled generator. If $\mathbf{b}$ is the output sequence from L_2 and t_j is the jth output from L_1, then this is

$$c_i = b_{\sigma(i)}, \qquad \text{where} \qquad \sigma(i) = \sum_{j=0}^{i} f(t_j). \tag{18.13}$$

The case when $F = \mathbb{F}_2$ and $f(t) = t$ is called the *stop-and-go generator*. It is weak since each change in the output reveals that a 1 has been generated by L_1. Also, there is a large correlation between consecutive output bits. The strength is improved by taking $f(t) = 1 + t$, giving rise to the *step-once-twice generator*.

In fact there is no need for L_1 and L_2 to be LFSRs. If $L_1 = (U_1, \Sigma_1, h_1, g_1)$ and $L_2 = (U_2, \Sigma_2, h_2, g_2)$ are arbitrary sequence generators (see Section 5.1.c), and $f : \Sigma_1 \to \mathbb{Z}$, then equation (18.13) defines the output from a clock-controlled sequence generator based on L_1 and L_2.

For general clock-controlled generators, if ρ_i is the period of L_1 and T is the sum of the f values of the states of L_1 in one period, then the period of the output sequence $\mathbf{c}$ is

$$\rho(\mathbf{c}) \left| \frac{\rho_1 \rho_2}{\gcd(T, \rho_2)}. \right.$$

The period is not maximal unless $\gcd(T, \rho_2) = 1$. We make this assumption for the remainder of this discussion. The linear span of $\mathbf{c}$ is upper bounded by $\lambda(\mathbf{c}) \leq n_2 \rho_1$, hence is large but not maximal. If L_1 generates an m-sequence, then the stop-and-go generator achieves the upper bound. If L_1 and L_2 generate the same m-sequence, then both the stop-and-go and step-once-twice generators achieve the maximum linear span [9]. Several authors have described correlation attacks on clock controlled shift registers.

Clock-controlled shift registers can be extended by using a cascade of registers, each output sequence clocking the next register. The structure may be modified so that the output from stage $k-1$ both clocks stage i and is added to the output of stage k modulo 2. A diagram of a cascaded clock-controlled shift register of height 3 is given in Figure 18.4.

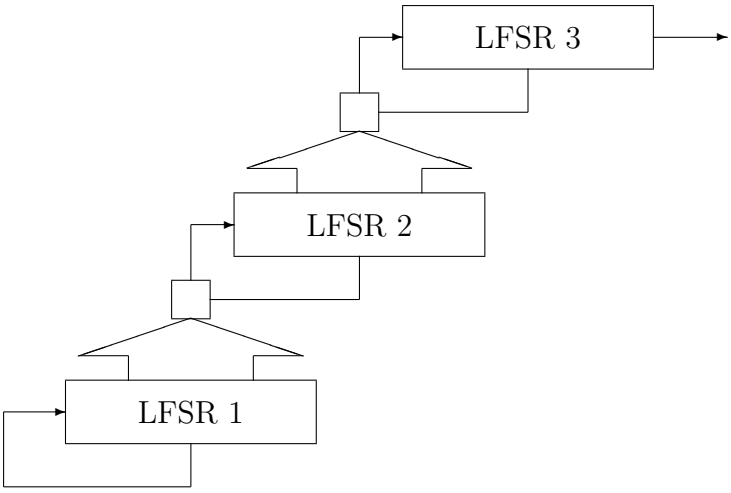


Figure 18.4: Cascaded Clock Controlled Shift Register of Height = 3

If 0,1 clocking is used (that is, a cascaded stop-and-go generator), each LFSR has maximal period and size n, and all LFSRs have distinct primitive connection polynomials, then the output period is $(2^n - 1)^k$ and the linear span is $n(2^n - 1)^{k-1}$. For arbitrary s, t clocking suppose the LFSRs have irreducible degree d connection polynomials and period p with $p^2 \not| (2^{p-1} - 1)$. Then the output period is p^n and the output linear span is at least $d(p^n - 1)/(p - 1)$. The distribution of subsequences of any fixed length approaches the uniform distribution as k approaches infinity.

With a *self-clocking generator* a single binary LFSR is used to clock itself. The case when the clocking function satisfies $f(s) = d$ if $s = 0$ and $f(s) = k$ if $s = 1$ is called the $[d, k]$-*self decimation generator*. Let us focus on this generator. Assume that the underlying LFSR has maximal period $2^n - 1$. The state graph of a $[d, k]$-self decimation generator may not be purely cyclic. If $d \neq k$, then the (eventual) period is at most $(3/4)(2^n - 1)$. In case $\gcd(d, 2^n - 1) = 1$ and $2d \equiv k \pmod 2^n - 1$ or $2^{n-1}d \equiv k \pmod 2^n - 1$ the period is exactly $(2/3)(2^n - 1)$. The distribution of short subsequences is close to uniform. There is evidence that the linear span is at least 2^{n-1} but this has not been proved. The drawback to the $[d, k]$-self decimation generator is that each output bit reveals a state bit and reveals which state bit is revealed by the next output bit. One way of avoiding this is to use different state bits for output and clocking control.

18.5.f The Shrinking generator

The *shrinking generator* represents another approach to irregular clocking. Again, we start with a pair of binary LFSRs, L_1 and L_2, with sizes n_1 and n_2 and output sequences $\mathbf{a}$ and $\mathbf{b}$. At each stage, if $a_i = 1$, then b_i is output. Otherwise no bit is output. Thus the output $\mathbf{c}$ is a shrunken version of $\mathbf{b}$: $c_j = b_{i_j}$ if i_j is the position of the jth 1 in $\mathbf{a}$. If $\mathbf{a}$ and $\mathbf{b}$ are m-sequences and $\rho(\mathbf{a})$ and $\rho(\mathbf{b})$ are relatively prime, then $\mathbf{c}$ has period

$$\rho(\mathbf{c}) = \rho(\mathbf{b})2^{n_1-1} = (2^{n_2} - 1)2^{n_1-1}.$$

The linear span of $\mathbf{c}$ satisfies

$$n_1 2^{n_2-2} < \lambda(\mathbf{c}) \leq n_1 2^{n_2-1}.$$

It can also be shown that the distribution of fixed length subsequences in $\mathbf{c}$ is close to uniform. One drawback is the irregularity of the output. A string of zeros in $\mathbf{a}$ leads to a delay in generating the next bit. Buffering can be used to alleviate this problem. A variation called the *self shrinking generator*, where a single register clocks itself, has also been considered.

18.5.g Linear span of ℓ-sequences

It is natural to investigate the linear span of ℓ-sequences. On might think that the function field and number field universes are essentially independent and that ℓ-sequences are essentially random with respect to linear span. This is not true. First, we can obtain a linear recurrence from Proposition 16.2.2.

Theorem 18.5.4. *If* $\mathbf{a}$ *is an* N*-ary* ℓ*-sequence based on a connection integer* $q = p^t$*, with* p *an odd prime, or if* $\mathbf{a}$ *is the* d*-fold decimation of such an* ℓ*-sequence, where* d *is odd, then* $\mathbf{a}$ *satisfies the linear recurrence*

$$a_{i+(q+1)/2} = a_{i+(q-1)/2} - a_{i+1} + a_i$$

Thus the linear span of $\mathbf{a}$ *is at most* $(q+1)/2$*, and the minimal polynomial of* $\mathbf{a}$ *is a divisor of*

$$x^{(q+1)/2} - x^{(q-1)/2} + x - 1 = (x-1)(x^{(q-1)/2} + 1).$$

Proof. By Proposition 16.2.2, for every i we have

$$a_{i+(q-1)/2} = q - 1 - a_i.$$

It follows that we also have

$$a_{i+1+(q-1)/2} = q - 1 - a_{i+1}.$$

The recurrence follows by subtracting these equations. □

Seo, Lee, Sung, Han, and Kim used the last fact in this theorem to obtain a lower bound on the linear span of a binary ℓ-sequence [178]. Their main theorem can be generalized to N-ary ℓ-sequences.

Theorem 18.5.5. *Suppose that N and $q = 2p+1$ are prime, $q \neq N$, p is prime, and $p \neq N$. Let m be the multiplicative order of N modulo p. If $\gcd(d, 2p) = 1$, then the linear span of a d-fold decimation of an ℓ-sequence with connection integer q is at least $m+2$.*

This is proved by observing that $(q-1)/2 = p$ and $x^p + 1$ can be factored using the theory of cyclotomic polynomials. See pages 64-66 of Lidl and Niederreiter's book for details on cyclotomic polynomials [123]. A similar bound can be obtained when q is a nontrivial power of a prime, but we omit the details.

If N is primitive modulo p (so that q is a so-called *strong N-prime*), then $m = p - 1$, so the lower and upper bounds for the linear span coincide. In this case the linear span λ of an ℓ-sequence is $p + 1 = m + 2$. More generally there may be a rather large gap between the bounds. For example, the values of $q \leq 227$ for which $N = 2$ is primitive modulo q and p is prime are $5, 11, 59, 83, 107, 179$, and 227. Of these, $5, 11, 59$, and 107 are 2-primes. If $q = 83$, then $m = 20$ and we have the bound $22 \leq \lambda \leq 41$. If $q = 179$, then $m = 11$ and we have the bound $13 \leq \lambda \leq 90$. If $q = 227$, then $m = 28$ and then we have the bound $30 \leq \lambda \leq 114$.

18.6 Exercises

1. Consider a sequence each of whose periods consists of $N - 1$ zeros followed by a single one. Show that this sequence has linear span N and can be generated by a finite state device whose state has $\lceil \log_2(N) \rceil$ bits.

2. Prove Parts (1) – (4) of Lemma 18.4.1.

3. Let $\mathbf{a}$ be an m-sequence with linear span k over a finite field F with q elements. Let $\mathbf{b}$ be the sequence obtained by changing one symbol (in each period) of $\mathbf{a}$. That is, we fix a j and for each i we replace $a_{i(q^k-1)+j}$ by $c \neq a_{i(q^k-1)+j}$. Determine the linear span of $\mathbf{b}$.

4. Let $\mathbf{a}$ and $\mathbf{b}$ be sequences over a finite field F. Let $\mathbf{c}$ be the term by term sum of $\mathbf{a}$ and $\mathbf{b}$ (that is, $c_i = a_i + b_i$ for every i). Prove that the linear span of $\mathbf{c}$ is at most the sum of the linear spans of $\mathbf{a}$ and $\mathbf{b}$. When is it equal to the sum of these linear spans? What is the smallest possible linear span of $\mathbf{c}$?

5. (GDFT) Consider the sequence $1, 1, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, \cdots$ with period $12 = 2^2 \cdot 3$ of elements in $F = \mathbb{F}_2$. Let $b \in \mathbb{F}_4$ be a primitive cube root of unity, so $b^2 + b + 1 = 0$. Using $f(x) =$

$1 + x + x^2 + x^4 + x^7$, show that the resulting GDFT of this sequence is

$$\widehat{\mathbf{a}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & b^2 & b \\ 1 & b & b^2 \end{bmatrix}.$$

6. (GDFT continued) Show that the Günther weight of the above matrix is 8 and that this agrees with equation (18.8):

$$\frac{f(x)}{1 - x^{12}} = \frac{1 + x + x^3}{(1 + x)^4(1 + x + x^2)^2}.$$

7. Let $L_1 = (U_1, \Sigma_1, h_1, g_1)$ and $L_2 = (U_2, \Sigma_2, h_2, g_2)$ be sequence generators generating sequences of period ρ_1 and ρ_2, respectively. Let $f : \Sigma_1 \to \mathbb{Z}$. Suppose that a clock controlled generator is constructed from these ingredients as in Section 18.5.e. Let be ρ the period of the resulting sequence. Also let T denote the sum of the f values over one period of L_1.

a. Show that ρ divides $\rho_1\rho_2/\gcd(T, \rho_2)$.
b. Find an example where $\rho \neq \rho_1\rho_2/\gcd(T, \rho_2)$.

8. Prove Theorem 18.5.5.

Chapter 19 FCSR Synthesis

19.1 N-adic span and complexity

As in the case of linear span, the N-adic span of a sequence is intended to measure how large an FCSR is required to output the sequence. In the LFSR case, this is given by the number of cells in an LFSR that outputs the sequence, and coincides with the degree of the connection polynomial, i.e., the denominator of the rational function giving the power series whose coefficients are the elements of the sequence.

In the N-ary FCSR case, things are more complicated. The number of N-ary coefficients in the connection integer equals the size of the basic register, but additional space is required for the memory. For purely periodic sequences, this extra memory is small (at most the $\log_N$ of the number of cells in the basic register), and if such sequences were our only concern we could ignore it. However, an *eventually* periodic sequence may require a considerable amount of extra memory. We would like to define the N-adic span of an eventually periodic sequence $\mathbf{a}$ to be the number of cells in the register plus the number of elements needed for the memory of an FCSR which outputs the sequence a. However even this definition must be approached with care because the memory value may grow as the FCSR runs (see the discussion at the end of this section). In the following paragraph we define two natural notions: the *span* (an integer which counts the number of N-ary cells in the register plus memory) and the *complexity* (a real number) of a sequence $\mathbf{a}$. The N-adic span is an engineering definition, while the N-adic complexity is a closely related mathematical definition. If $\mathbf{a}$ is strictly periodic, then the number of cells in the basic register (not counting the memory) is $\lceil \text{complexity}(\mathbf{a}) \rceil$. We show that these two measures differ at most by $\log_N(\text{complexity}(\mathbf{a}))$.

Let $\mathbf{a} = (a_0, a_1, \ldots)$ be an eventually period sequence of bits. Suppose an FCSR with connection integer

$$q = -1 + q_1 N^1 + \ldots + q_m N^m$$

and initial memory z ouputs this sequence, and that $q_m \neq 0$ (i.e. that $m = \lfloor \log_N(q+1) \rfloor$). We associate to this FCSR the number

$$\lambda = \begin{cases} m + \max(\lfloor \log_N(\text{wt}(q+1)) \rfloor, \lfloor \log_N |z| \rfloor + 1) & \text{if } z \neq 0 \\ m + \lfloor \log_N(\text{wt}(q+1)) \rfloor & \text{otherwise} \end{cases}$$

of N-ary cells in the FCSR, where $\text{wt}(q+1)$ denotes the sum of the q_i for $1 \leq i \leq m$. (See Section 7.6: memory values within the range

$$0 \leq m < \text{wt}(q+1)$$

may grow and shrink within this range; memory values outside this range will move monotonically toward this range. The "+1" in the second argument of the maximum is a "sign bit" which allows for possible negative memory values.)

Definition 19.1.1. *The N-adic span $\lambda_N(\mathbf{a})$ of an N-ary, eventually period sequence $\mathbf{a}$ is the smallest value of λ which occurs among all FCSR's whose output is the sequence $\mathbf{a}$.*

Now suppose $\mathbf{a} = \mathbf{seq}_N(f/q)$, with coprime integers f, q, so that $f/q = a_0 + a_1 N + \cdots \in \mathbb{Z}_N$ is the fraction in lowest terms whose N-adic expansion agrees with the sequence $\mathbf{a}$.

Definition 19.1.2. *The N-adic complexity of the sequence $\mathbf{a}$ is the real number*

$$\varphi_N(\mathbf{a}) = \log_N(\Phi(f, q)),$$

where $\Phi(f, q) = \max(|f|, |q|)$ is the Weil height *of f/q.*

Proposition 19.1.3. *If $\mathbf{a} = \mathbf{seq}_N(f/q)$ with f, q coprime, then its N-adic span and complexity are related by*

$$|\lambda_N(\mathbf{a}) - \varphi_N(\mathbf{a})| \leq \log_N(\varphi_N(\mathbf{a})) + 2. \tag{19.1}$$

Remarks. In Corollary 19.3.6 we show that the sequence $\mathbf{a}$ is determined by its first $\varphi_N(\mathbf{a}) + 1$ symbols; compare Proposition 18.2.1.

It is also possible to estimate the above difference between the complexity and span in terms of the span. If $N > 2$ or $\varphi_N(\mathbf{a}) \geq 2/\ln(N)$ then

$$\log_N(\varphi_N(\mathbf{a})) \leq \frac{\varphi_N(\mathbf{a})}{2}$$

so equation (19.1) gives $\varphi_N(\mathbf{a}) \leq 2(\lambda_N(\mathbf{a}) - 2)$, hence

$$\log_N(\varphi_N(\mathbf{a})) \leq \log_N(\lambda_N(\mathbf{a}) + 2) + \log_N(2).$$

This gives the inequality

$$|\lambda_N(\mathbf{a}) - \varphi_N(\mathbf{a})| \leq \log_N(\lambda_N(\mathbf{a}) + 2)) + 2. \tag{19.2}$$

If $N = 2$ and $\varphi_N(\mathbf{a}) < 4$ then the inequality (19.2) may be checked directly – there are 240 cases with $0 \leq |f| \leq 15$ and $1 \leq |q| \leq 15$.

Proof of Proposition 19.1.3. By possibly multiplying both f and q by -1, we may assume that $q > 0$. For notational simplicity, let us write $\lambda = \lambda_N(\mathbf{a})$, $\varphi = \varphi_N(\mathbf{a})$ and $\Phi = \Phi(p, q)$. Let z be the initial memory, so that equation (7.11) holds. Let

$$w = \mathrm{wt}(q+1) = \sum_{i=1}^{m} q_i.$$

If $z = 0$, then $f \geq 0$, and the proof is a simplification of the proofs in Case 1 and Case 4 below. So assume from now on that $|z| \geq 1$. Expanding the absolute value and exponentiating, equation (19.1) is equivalent to

$$\frac{\Phi}{\log_N(\Phi)} \leq N^{m+2} \max(N^{\lfloor \log(w) \rfloor}, N^{\lfloor \log |z| \rfloor + 1}) \tag{19.3}$$

and

$$N^m \max(N^{\lfloor \log(w) \rfloor}, N^{\lfloor \log |z| \rfloor + 1}) \leq N^2 \Phi \log_N(\Phi). \tag{19.4}$$

We do not know a uniform proof for this statement, and there are many cases to consider.

Case 1: $\mathbf{f < -q}$. In this case $|f| \geq 3$. We have $\Phi(\mathbf{a}) = -f = |f|$ so that

$$\varphi_N(\mathbf{a}) = \log_N(-f) = \log_N |f|.$$

Thus the left hand side of equation (19.3) is $|f| / \log_N |f|$. Let $g(x) = x / \log_N(x)$. We have

$$g'(x) = (\log_N(x) - \log_N(e)) / \log_N(x)^2,$$

where e is the base of natural logarithms. Thus $g'(x) > 0$ for $x \geq 3$. By equations (7.11) and (7.13), we have $z \geq 1$ and $|f| \leq (z+1)N^m$. Thus

$$\frac{|f|}{\log_N |f|} \leq \frac{(z+1)N^m}{\log_N((z+1)N^m)} \leq \frac{(z+1)N^m}{m} \leq N^{\lfloor \log |z| \rfloor + 1} N^m.$$

This gives inequality (19.3).

Now consider inequality (19.4) under the same condition, $f < -q$. We have $q \geq N^m - 1$, so $|f| \geq N^m$. If $z \leq w$ then

$$N^m N^{\max(\lfloor \log(w) \rfloor, \lfloor \log |z| \rfloor + 1)} \leq N^{m+1} w \leq N^{m+2} m \leq N^2 |f| \log_N |f|$$

as desired. So suppose $z \geq w+1$. Then $(z-m+1)m \geq z$. Since $f < 0$ we have $|f| \geq (z-w+1)N^m$ (cf eq. (7.13)). So

$$\begin{aligned} N^2 |f| \log_N |f| &\geq (z-w+1) N^{m+2} (m + \log_N(z-w+1)) \\ &> (z-w+1) N^{m+2} m \\ &\geq z N^{m+1} \\ &\geq N^{\lfloor \log |z| \rfloor} N^{m+1}. \end{aligned}$$

This concludes the proof of Proposition 19.1.3 in Case 1.

Case 2: $\mathbf{f > q}$. In this case $\Phi(\mathbf{a}) = f$ and $\varphi_N(\mathbf{a}) = \log_N(f)$. We have

$$\frac{f}{\log_N(f)} \leq \frac{w - q_m - z}{m} N^m < \frac{w}{m} N^m \leq N^{m+1}.$$

This gives inequality (19.3).

We also have

$$N^m \max(N^{\lfloor \log(w) \rfloor}, N^{\lfloor \log |z| \rfloor + 1}) \leq N^m \max(w, N|z|).$$

If $z > 0$ then $z \leq w - q_m - 1 < w < Nm$, so $N^m \max(w, N|z|) \leq N^{m+2} m \leq N^2 f \log_N(f)$. If $z \leq 0$, then $(|z| - 1)N^m < f$ so $N^m \max(w, N|z|) \leq N^2 f \log_N(f)$. This concludes the proof in Case 2.

Case 3: $\mathbf{-q < f < 0}$. In this case $\Phi(\mathbf{a}) = q$ and $\varphi_N(\mathbf{a}) = \log_N(q)$. Since $q \leq N^{m+1}$,

$$\frac{\Phi}{\log_N(\Phi)} = \frac{q}{\log_N(q)} \leq \frac{N^{m+1}}{m} \leq N^{m+1},$$

which proves inequality (19.3). By Corollary 7.2.2 the sequence $\mathbf{a}$ is strictly periodic, and the memory z lies in the range $0 \leq z \leq w - 1$. In particular, $\max(w, N|z|) \leq Nw < N^2 m$. Then

$$N^m \max(N^{\lfloor \log(w) \rfloor}, N^{\lfloor \log |z| \rfloor + 1}) \leq N^{m+2} m \leq N^2 q \log_N(q)$$

unless $q = N^m - 1$. If $q = N^m - 1$ then $w = 1$ so $z = 0$, which we have assumed is not the case. This concludes the proof in Case 3.

Case 4: $\mathbf{0 < f < q}$. Inequality (19.3) is proven just as in Case 3 above. For inequality (19.4), first let $z \geq 0$. Then by Lemma 7.2.4 part (1), $z \leq w - 2$ so

$$N^m \max(N^{\lfloor \log(w) \rfloor}, N^{\lfloor \log |z| \rfloor + 1}) \leq N^{m+1} w \leq N^{m+2} m.$$

If $q \neq N^m - 1$, then $N^{m+2} m \leq N^2 q \log_N(q)$. If $q = N^m - 1$, then $w = 1$ and we still have $N^{m+1} w \leq N^2 q \log_N(q)$.

Next suppose $z < 0$. By equation (7.12), $N^{m+1} > q > f > (|z| - 1)N^m$. Thus, $|z| < N$ and

$$N^m \max(N^{\lfloor \log(w) \rfloor}, N^{\lfloor \log |z| \rfloor + 1}) \leq N^m \max(w, N) \leq N^2 q \log_N(q)$$

(A special argument must be made if $N = 2$ and $q = N - 1 = 1$.) This concludes the proof. □

Proposition 19.1.4. *Let* $\mathbf{a}$ *be an eventually periodic* N*-ary sequence. Then* $\varphi_N(\mathbf{d}) \leq \lambda_N(\mathbf{d}) + \log_N(\lambda_N(\mathbf{d}))$.

The proof is left as an exercise. It follows that $\log_N(\varphi_N(\mathbf{a})) \leq 2\log_N(\lambda_N(\mathbf{a})) + 2\log_N(3)$ (the constants here are a bit crude).

Proposition 19.1.3 allows us to relate the N-adic spans of two sequences to the 2-adic span of their with-carry sum, see equation (5.5).

Theorem 19.1.5. *Suppose* $\mathbf{a}$ *and* $\mathbf{b}$ *are periodic* N*-ary sequences. Let* $\mathbf{c}$ *denote the* N*-ary sequence obtained by adding the sequences* $\mathbf{a}$ *and* $\mathbf{b}$ *with carry, see equation* (5.5)*. Then the* N*-adic complexity of* $\mathbf{c}$ *is bounded as follows,*

$$\varphi_N(\mathbf{c}) \leq \varphi_N(\mathbf{a}) + \varphi_N(\mathbf{b}) + \log_N(2).$$

The N*-adic span is bounded as follows,*

$$\lambda_N(\mathbf{c}) \leq \lambda_N(\mathbf{a}) + \lambda_N(\mathbf{b}) + 2\log_N(\lambda_2(\mathbf{a})) + 2\log_N(\lambda_N(\mathbf{b})) + 3\log_N(2) + \log_N(3) + 2.$$

Proof. Write $\mathbf{a} = \mathbf{seq}_N(f_1/q_1)$ and $\mathbf{b} = \mathbf{seq}_N(f_2/q_2)$ as reduced fractions. The sum-with-carry sequence is

$$\mathbf{c} = \mathbf{seq}_N\left(\frac{f_1}{q_1} + \frac{f_2}{q_2}\right) = \mathbf{seq}_N\left(\frac{f_1q_2 + f_2q_1}{q_1q_2}\right), \tag{19.5}$$

so

$$\begin{aligned}
\varphi_N(\mathbf{c}) &= \log_N(\Phi(p_1q_2 + p_2q_1, q_1q_2)) \\
&\leq \log_N(2\Phi(p_1, q_1)\Phi(p_2, q_2)) \\
&= \varphi_N(\mathbf{a}) + \varphi_N(\mathbf{b}) + \log_N(2).
\end{aligned}$$

For the second claim, first note that for any $x_1, \cdots, x_k > 0$, $\log_N(\sum_k x_i) \leq \log_N(k\max(x, y, z)) = \log_N(k) + \max(\{\log_N(x_i)\}) \leq \log_N(3) + \sum_i \log_N(x_i)$. By Proposition 19.1.3

$$\begin{aligned}
\lambda_N(\mathbf{c}) &\leq \varphi_N(\mathbf{c}) + \log_N(\varphi_N(\mathbf{c})) + 2 \\
&\leq \varphi_N(\mathbf{a}) + \varphi_N(\mathbf{b}) + \log_N(2) + \log_N(\varphi_N(\mathbf{a}) + \varphi_N(\mathbf{b}) + \log_N(2)) + 2 \\
&\leq \varphi_N(\mathbf{a}) + \varphi_N(\mathbf{b}) + \log_N(\varphi_N(\mathbf{a})) + \log_N(\varphi_N(\mathbf{b})) + \log_N(2) + \log_N(3) + 2 \\
&\leq \lambda_N(\mathbf{a}) + \lambda_N(\mathbf{b}) + \log_N(\lambda_N(\mathbf{a})) + \log_N(\lambda_N(\mathbf{b})) + \log_N(\lambda_N(\mathbf{a})) + \log_N(\lambda_N(\mathbf{b})) \\
&\qquad +3\log_N(2) + \log_N(3) + 2 \\
&\leq \lambda_N(\mathbf{a}) + \lambda_N(\mathbf{b}) + 2\log_N(\lambda_2(\mathbf{a})) + 2\log_N(\lambda_N(\mathbf{b})) + 3\log_N(2) + \log_N(3) + 2
\end{aligned}$$

by Proposition 19.1.4. □

The span may be much less than this if the fraction on the right hand side of equation (19.5) is not in lowest terms.

What is the N-adic span of an m-sequence? Although we do not know, it is easy to prove that there exist m-sequences of maximal 2-adic span.

Theorem 19.1.6. *Suppose* $\mathbf{a}$ *is a periodic binary sequence with period* $T = 2^k - 1$. *Suppose that* $2^T - 1$ *is prime. Then the* 2*-adic span of* $\mathbf{a}$ *is* $T + 2$.

Proof. Consider an FCSR that generates the sequence $\mathbf{a}$, and let q denote its connection integer. Then $\mathrm{ord}_q(N) = T = 2^k - 1$. Therefore $2^T \equiv 1 \pmod{q}$. This says that $2^T - 1$ is divisible by q. However, by assumption, $2^T - 1$ is prime, hence $q = 2^T - 1$. The 2-adic span is then at least $\log_2(q+1) + 1 = T + 1$. However, any sequence of period T can be generated by an FCSR with T bits in the basic register, one bit of carry (which is always zero), and one sign bit (which is always zero). □

This argument doesn't quite work in the N-ary case with N odd since $N^T - 1$ is even and thus never prime. One can still obtain good estimates in some cases. For example, if $N^T - 1$ is twice a prime, then the N-adic span equals either the period or half the period.

More generally, the same proof shows that the N-adic span of any periodic sequence with period T is greater than or equal to $\log_N(r+1)+1$, where r is the smallest prime divisor of $N^T - 1$ that exceeds the period.

We remark that if $N = 2$, $T = 2^k - 1$, and $2^T - 1$ is prime, then both T and k are prime as well. However, the hypotheses of this theorem may be difficult to verify in practice. It is possible that there are only finitely many primes of the form $2^T - 1$, and in any case the largest prime known to date is $q = 2^T - 1$ where $T = 859,433$. An m-sequence generated by a LFSR with only 20 cells already has period $T = 1,048,575$. So, a verification of the hypothesis of the theorem for any larger m-sequence would mean discovering a new prime number.

Complexity Profile. A periodic sequence $\mathbf{a}$ with low complexity or span can be artificially made to have high complexity by changing only a few symbols. For this reason the linear span, N-adic span, etc. are at best crude upper bounds on the cryptographic security of the sequence. Rueppel [169] (p. 33) has suggested that the *complexity profile*, i.e., the full record of the span or complexity of the first k symbols of $\mathbf{a}$, as k varies between 1 and the period of $\mathbf{a}$, might be a more meaningful indicator. We are therefore led to consider the N-adic span and complexity of a finite segment of the sequence $\mathbf{a}$. It is not clear what the N-adic *span* of such a finite segment should mean, because the memory of the corresponding FCSR might grow as the FCSR runs. However it is possible to make sense of the N-adic complexity of a finite sequence.

Let $\mathbf{a} = a_0, a_1, a_2, a_{k-1}$ be a finite N-ary sequence. Define

$$\psi(\mathbf{a}) = \log_N(\min_{(f,q)} \Phi(f,q)) = \log_N(\min_{(f,q)} \max(|f|,|q|))$$

where the minimum is taken over all pairs $(f, q) \in \mathbb{Z} \times \mathbb{Z}$ of integers, with $q \equiv -1 \pmod{N}$, such that the first k coefficients in the N-adic expansion of the fraction f/q is precisely $a_0, a_1, \ldots, a_{k-1}$. (In the language of Section 19.3, $\psi(\mathbf{a})$ is the minimum value of $\log_N(\Phi(h))$ as h is allowed to vary in the k-th approximation lattice L_k of $a = \sum_{i=0}^{k-1} a_i N^i$.)

Now let $\mathbf{a} = a_0, a_1, \ldots$ be a possibly infinite N-ary sequence. The *N-adic complexity profile* is the function $\psi_{\mathbf{a}} : \mathbb{Z}^{>0} \to \mathbb{R}$ defined by

$$\psi_{\mathbf{a}}(k) = \psi(a_0, a_1, \ldots, a_{k-1}),$$

whose values are the N-adic complexity of the first k terms in the sequence $\mathbf{a}$. The algorithm presented in Section 19.3 may be used to compute the complexity profile. In fact, (using the notation of Figure 19.2), at the k-th step in the algorithm we have $\psi_{\mathbf{a}}(k) = \log_N(\max(|f|, |g|))$. A highly random sequence $\mathbf{a}$ will exhibit an N-adic complexity profile $\psi_{\mathbf{a}}(k)$ which grows approximately as $k/2$.

Maximum order complexity. The maximum order complexity of a sequence is the size of the smallest (possibly nonlinear) feedback shift register (without memory) which may be used to generate the sequence (cf. [86, 87, 88, 14]). If a sequence $\mathbf{a}$ is generated by a FCSR with nonnegative memory, then its N-adic span is no greater than its maximum order complexity. In fact an FCSR may be interpreted as a (nonlinear) feedback shift register without memory, and the N-adic span counts the *total* number of cells.

19.2 Symmetric N-adic span

Let $\mathbf{a} = (a_0, a_1, a_2, \cdots)$ be a periodic sequence with entries in a ring U and with period n. In general in cryptanalysis one may ask whether, given a block of elements of $\mathbf{a}$, it is easier to infer the elements that precede the block than it is to infer those that follow the block. Let $\mathbf{a}^{\text{rev}}$ be the sequence formed by reversing each period of $\mathbf{a}$, the *reversal of* $\mathbf{a}$. We are interested in whether the reversal of a sequence has smaller complexity (with respect to some class of sequence generators) than the original sequence.

First let us consider linear span. Let $a(x) = \sum_{i=0}^{\infty} a_i x^i$ be the power series associated with $\mathbf{a}$. Then the linear span of $\mathbf{a}$ is the degree of the smallest degree monic polynomial $q(x)$ such that for some polynomial $f(x)$ we have $a(x) = f(x)/q(x)$. The polynomial $f(x)$ necessarily has degree less than the degree of $q(x)$. If $g(x)$ is a polynomial of degree d, let $g^{\text{rev}}(x)$ denote the polynomial of degree d with same coefficients as $g(x)$ but in reverse order. This polynomial is called the *reversal of* $g(x)$. That is, $g^{\text{rev}}(x) = x^d g(1/x)$. Suppose that $g(x)$ and $h(x)$ are two polynomials whose product has degree equal to the sum of their degrees. (This is equivalent to saying that the product of their leading coefficients is nonzero, and always holds if U is an integral domain.) Then $g^{\text{rev}}(x)h^{\text{rev}}(x) = (gh)^{\text{rev}}(x)$.

Lemma 19.2.1. *The linear span of* $\mathbf{a}^{rev}$ *equals the linear span of* $\mathbf{a}$.

Proof. Let $h(x) = \sum_{i=0}^{n-1} a_i x^i$. Then $a(x) = h(x)/(1-x^n)$. The power series associated with $\mathbf{a}^{\rm rev}$ is just $a^{\rm rev}(x) = -h^{\rm rev}(x)/(1-x^n)$. We have $f(x)(1-x^n) = h(x)q(x)$, so that $f(x)^{\rm rev}(1-x^n)^{\rm rev} = h(x)^{\rm rev}q(x)^{\rm rev}$. Thus

$$\frac{-h(x)^{\rm rev}}{1-x^n} = \frac{-f(x)^{\rm rev}}{q(x)^{\rm rev}}.$$

It follows that the linear span of $\mathbf{a}^{\rm rev}$ is no greater than the linear span of $\mathbf{a}$. Since the reversal of the reversal of $\mathbf{a}$ is $\mathbf{a}$, the opposite inequality holds as well, and the two sequences have the same linear span. □

Thus reversal makes no difference for linear span. The picture is quite different for N-adic complexity. One can define the reversal of an integer with respect to its N-ary representation. However, this operation no longer commutes with multiplication. The problem is that the carries in the multiplications go in opposite directions.

Example 19.2.2. *Consider the binary* ℓ*-sequence* $\mathbf{a}$ *with period 18, one of whose periods is*

$$00010100111101011.$$

The associated 2-adic integer is the rational number

$$\frac{-110376}{2^{18}-1} = \frac{-8}{19}.$$

The 2-adic reversal of this sequence has a period

$$1101011110010100 00,$$

whose associated 2-adic integer is the rational number

$$\frac{-10731}{2^{18}-1} = \frac{-7}{171}.$$

Thus the reversal of $\mathbf{a}$ *has greater 2-adic span than that of* $\mathbf{a}$.

It follows that a cryptanalyst who is given a finite set of consecutive symbols of a keystream, wants to find a generator of the keystream might simultaneously apply an FCSR synthesis algorithm to the sequence and its reversal. If the number of symbols available is sufficient for either application of the algorithm to succeed, then the keystream must be considered vulnerable. The attacker might not know which has succeeded, but the context of the message suffice to tell.

Definition 19.2.3. *Let $N \geq 2$ be a natural number and let* $\mathbf{a}$ *be a periodic sequence over the alphabet* $\{0, 1, \cdots, N-1\}$. *Then the* symmetric N-adic span *of* $\mathbf{a}$ *is the minimum of the N-adic spans of* $\mathbf{a}$ *and* $\mathbf{a}^{rev}$.

Any algorithm for finding the N-adic span of a sequence $\mathbf{a}$ given a block $\mathbf{b}$ of symbols of $\mathbf{a}$ can be converted to one for finding the symmetric N-adic span by running the original algorithm in parallel on $\mathbf{b}$ and $\mathbf{b}^{\text{rev}}$. The minimum of the outputs of the parallel runs is taken as the output of the new algorithm.

In example 19.2.2 the sequence $\mathbf{a}$ is an ℓ-sequence. In general it is the case that the reversal of a binary ℓ-sequence has larger N-adic span than the ℓ-sequence. To see this we use our analysis of the distribution of patterns of consecutive elements in an ℓ-sequence.

Let q be prime, and suppose that 2 is primitive modulo q. Let s be any nonnegative integer, and let $\mathbf{b}$ and $\mathbf{c}$ be s-bit sequences. Let $\mathbf{a}$ be a binary ℓ-sequence with connection integer q. By Proposition 16.2.2 the number of occurrences of any bit pattern in $\mathbf{a}$ equals the number of occurrences of its bitwise complement. Also, recall from Corollary 16.2.1 that the numbers of occurences of $\mathbf{b}$ and $\mathbf{c}$ in $\mathbf{a}$ differ by at most 1. If $q = 2^t + e$ with $1 \leq e < 2^t$, then this says that every t-bit pattern occurs either once or twice, and every $(t+1)$-bit pattern occurs either once or not at all. In fact we can say exactly which t-bit and $(t+1)$-bit patterns occur once and which occur twice. For any integer y, we denote by $[y]$ the reduced residue of y modulo 2^t. That is $y \equiv [y] \bmod 2^t$ and $0 \leq [y] < 2^t$.

Lemma 19.2.4. *Let* $\mathbf{a}$ *be a binary ℓ-sequence with prime connection integer $q = 2^t + e$ and $1 \leq e < 2^t$. If $\mathbf{c} = (c_0, c_1, \cdots, c_{t-1})$ is a t-bit pattern, and $c = \sum_{i=0}^{t-1} c_i 2^i$, then* $\mathbf{c}$ *occurs twice in* $\mathbf{a}$ *if and only if $[-ec] < e$.*

Proof. The left shifts of $\mathbf{a}$ are the 2-adic expansions of the rational numbers $-x/q$ with $0 < x < q$. Thus the occurences of the t-bit pattern $\mathbf{c}$ correspond to the integers x such that $c \equiv -x/q \bmod 2^t$ and $0 < x < q$. The first condition is equivalent to $qc \equiv -x \bmod 2^t$, which is equivalent to $ec \equiv -x \bmod 2^t$, which is equivalent to $-ec \equiv x \bmod 2^t$. That is $[-ec] = x$.

If $[-ec] < e$, then $2^t + [-ec] < 2^t + e = q$, so we have two occurences of $\mathbf{c}$. Conversely, if $[-ec] \geq e$, then any integer congruent to $[-ec]$ modulo 2^t must either be greater than or equal to q or less than 0. Thus there is only one occurence of $\mathbf{c}$. □

For example, if $\mathbf{c}$ is the all one pattern of length t, then $c = 2^t - 1$, so $[-ec] = e$, so this $\mathbf{c}$ occurs once. Also, the all zero pattern of length t, occurs once.

Theorem 19.2.5. *Let* $\mathbf{a}$ *be a binary ℓ-sequence with prime connection integer q. Then the reversal* $\mathbf{a}^{rev}$ *is shift distinct from* $\mathbf{a}$*, so it is not an ℓ-sequence.*

Proof. Suppose to the contrary that $\mathbf{a}^{\text{rev}}$ is a left shift of $\mathbf{a}$. Then for some k we have $a_i = a_{k-i}$ for all i. Set $q = 2^t + e$ with $1 \leq e < 2^t$. We proceed by finding a series of constraints on e that arise from occurrences of t bit patterns in $\mathbf{a}$.

Consider the pattern of t ones. It occurs once and is its own reversal. If it occurs at index i, then its reversal occurs at index $k - i - t + 1$, so we must have $i = k - i - t + 1$ or $t = k - 2i + 1$. In particular, k and t have opposite parity. Suppose that $\mathbf{c}$ is a $(t+1)$-bit pattern that is equal to its own reversal. We claim that $\mathbf{c}$ cannot occur in $\mathbf{a}$. We know that $\mathbf{c}$ can occur at most once. Thus if it occurs at position j, then its reversal occurs at position $k - j - t$, so $t = k - 2j$. This would imply that k and t have the same parity, which is false. It follows that any $t+1$-bit pattern that equals its own reversal cannot occur in $\mathbf{a}$.

Suppose that $\mathbf{c}$ is a t-bit pattern that equals its own reversal and occurs just once in $\mathbf{a}$, say at index i. Then its reversal occurs at index $k - i - t + 1$, so we must have $i = k - i - t + 1$. That is, $i \in \{(k - t + 1)/2, (k - t + q)/2\}$. Thus there are at most two such bit patterns. By the comments preceeding the theorem, there are precisely two, the all zero and all one t-bit patterns.

Now suppose that $\mathbf{c}$ is a t-bit pattern and b is a bit such that the $(t+1)$-bit pattern $\mathbf{c}, b$ equals its own reversal. Then $\mathbf{c}, b$ cannot occur in $\mathbf{a}$. Let b' denote the complement of b. Then $\mathbf{c}, b'$ can occur at most once in $\mathbf{a}$, so $\mathbf{c}$ can also occur at most once. In particular, by Lemma 19.2.4 if $c = c_0 + 2c_1 + \cdots + 2^{t-1}c_{t-1}$, then $[-ec] \geq e$. This gives us a crank to turn. We let

$$e = \sum_{i=0}^{t-1} e_i 2^i,$$

where $e_0 = 1$ and $e_i \in \{0, 1\}$ for $i = 1, \cdots, t-1$.

1. Let $\mathbf{c} = 1, 0, 0, \cdots, 0$. Then $\mathbf{c}$ occurs once in $\mathbf{a}$ and $c = 1$, so $e \leq [-e] = 2^t - e$, so $e \leq 2^{t-1}$. In fact, since q is odd,

$$e < 2^{t-1}. \tag{19.6}$$

2. Let $\mathbf{c} = 0, 0, 1, 1, \cdots, 1, 0$. Thus $\mathbf{c}$ occurs once in $\mathbf{a}$ and $c = 2^{t-1} - 4$, so $e \leq [-(2^{t-1} - 4)e] = [2^{t-1} + 4e]$ since e is odd. We have $2^{t-1} + 4e < 2^{t-1} + 4 \cdot 2^{t-1} < 3 \cdot 2^t$ by equation (19.6). Suppose that $2^{t-1} + 4e \geq 2^{t+1}$. Then $[2^{t-1} + 4e] = 2^{t-1} + 4e - 2 \cdot 2^t \geq e$. It follows that $e \geq 2^{t-1}$, a contradiction. Therefore $2^{t-1} + 4e < 2^{t+1}$ so

$$e < \frac{3}{4} 2^{t-1}. \tag{19.7}$$

3. Let $\mathbf{c} = 0, 1, \cdots, 1, 0$. Then $\mathbf{c}$ occurs twice in $\mathbf{a}$ and $c = 2^{t-1} - 2$. Thus $e > [-(2^{t-1} - 2)e] = [2^{t-1} + 2e]$ since e is odd. We have $2^{t-1} + 2e < 5 \cdot 2^{t-2} < 2 \cdot 2^t$ by equation (19.7). Thus $[2^{t-1} + 2e] \in \{2^{t-1} + 2e, 2^{t-1} + 2e - 2^t = 2e - 2^{t-1}\}$. If $[2^{t-1} + 2e] = 2^{t-1} + 2e$ then $e > 2^{t-1} + 2e$,

which is impossible. Thus $[2^{t-1}+2e] = 2e - 2^{t-1}$ and $2^{t-1}+2e \geq 2^t$. That is, $e \geq 2^{t-2}$. Since e is odd, we have

$$e \geq 2^{t-2} + 1. \tag{19.8}$$

4. Let $\mathbf{c} = 1, 0, 1, 1, \cdots, 1, 0, 1$. Then $\mathbf{c}$ occurs twice in $\mathbf{a}$ and $c = 2^t - 2^{t-2} - 3$. Thus $e > [-(2^t - 2^{t-2} - 3)e] = [3e + 2^{t-2}(1 + 2e_1)]$. Suppose that $e_1 = 1$. Then $3e + 2^{t-2}(1 + 2e_1) = 3e + 3 \cdot 2^{t-2} \geq 3(2^{t-2}+1) + 3 \cdot 2^{t-2} > 2^t$ by equation (19.8). Also, $3e + 3 \cdot 2^{t-2} < (9/2)2^{t-2} + 3 \cdot 2^{t-2} = (15/2)2^{t-2} < 2 \cdot 2^t$ by equation (19.7). Thus $e > [3e + 3 \cdot 2^{t-2}] = 3e + 3 \cdot 2^{t-2} - 2^t = 3e - 2^{t-2}$, so $2^{t-2} > 2e$, which contradicts equation (19.8). Therefore $e_1 = 0$.

5. Let $\mathbf{c} = 0, 0, 1, 1, \cdots, 1, 0, 0$. Then $\mathbf{c}$ occurs twice in $\mathbf{a}$ and $c = 2^{t-2} - 4$. Thus $e > [-(2^{t-2} - 4)e] = [4e - 2^{t-2}]$. We have $4e - 2^{t-2} \geq 4(2^{t-2} + 1) - 2^{t-2} \geq 0$ by equation (19.8). If $4e - 2^{t-2} < 2^t$, then $e > [4e - 2^{t-2}] = 4e - 2^{t-2}$. This would imply that $2^{t-2} > 3e$, which contradicts equation (19.8). Thus $2^t \leq 4e - 2^{t-2}$ and so

$$e \geq \frac{5}{4} 2^{t-2}. \tag{19.9}$$

6. Let $\mathbf{c} = 0, 0, 0, 1, 1, \cdots, 1, 0, 0$. Then $\mathbf{c}$ occurs once in $\mathbf{a}$ and $c = 2^{t-2} - 8$. Thus $e \leq [-(2^{t-2} - 8)e] = [8e - 2^{t-2}]$. We have $8e - 2^{t-2} > 5 \cdot 2^{t-1} - 2^{t-2} \geq 2 \cdot 2^t$ by equation (19.9). Moreover, $8e - 2^{t-2} < 3 \cdot 2^t$ by equation (19.7). Thus $e \leq [8e - 2^{t-2}] = 8e - 2^{t-2} - 2 \cdot 2^t = 8e - 9 \cdot 2^{t-2}$ and so

$$e > \frac{9}{7} 2^{t-2}. \tag{19.10}$$

7. Let $\mathbf{c} = 1, 1, 0, 1, 1, \cdots, 1, 0, 1$. Then $\mathbf{c}$ occurs once in $\mathbf{a}$ and $c = 2^t - 2^{t-2} - 5$. Thus $e \leq [-(2^t - 2^{t-2} - 5)e] = [5e + 2^{t-2}]$. We have $5e + 2^{t-2} > (45/7)2^{t-2} + 2^{t-2} = (52/7)2^{t-2} > 2^t$ by equation (19.10). Suppose $5e + 2^{t-2} \geq 2 \cdot 2^t$. Then $e < [5e + 2^{t-2}] \leq 5e + 2^{t-2} - 2 \cdot 2^t = 5e - 7 \cdot 2^{t-2}$, so $7 \cdot 2^{t-2} < 4e$, which contradicts equation (19.7). Thus $5e + 2^{t-2} < 2 \cdot 2^t$, so

$$e < \frac{7}{5} 2^{t-2}. \tag{19.11}$$

8. Let $\mathbf{c} = 0, 1, 0, 1, 1, \cdots, 1, 0, 1$. Then $\mathbf{c}$ occurs once in $\mathbf{a}$ and $c = 2^t - 2^{t-2} - 6$. Thus $e \leq [-(2^t - 2^{t-2} - 6)e] = [6e + 2^{t-2}]$. We have $6e + 2^{t-2} > (54/7)2^{t-2} + 2^{t-2} = (61/7)2^{t-2} > 2 \cdot 2^t$ by equation (19.10). Similarly, $6e + 2^{t-2} < 3 \cdot 2^t$. Thus $e \leq [6e + 2^{t-2}] = 6e + 2^{t-2} - 2 \cdot 2^t = 6e - 7 \cdot 2^{t-2}$. It follows that

$$e \geq \frac{7}{5} 2^{t-2}. \tag{19.12}$$

Equation (19.12) contradicts equation (19.11), proving the theorem[1] if t is large enough that all the choices of $\mathbf{c}$ used are valid strings of length t. This holds as long as $t \geq 6$, so $q \geq 64$. For smaller values of q, the theorem can be verified by exhaustive search. □

19.3 Rational approximation

Suppose $\mathbf{a} = (a_0, a_1, a_2, \cdots)$ is an eventually periodic binary sequence. We consider the problem of finding an FCSR whose output sequence coincides with $\mathbf{a}$. In the language of Section 18.1, we want to solve the register synthesis problem for FCSRs. The analogous problem for LFSRs over a field is completely solved by the Berlekamp-Massey algorithm (see Section 18.2). This algorithm is optimal in two senses: (1) it determines the *smallest* LFSR whose output coincides with $\mathbf{a}$, and (2) it does so with only the first $2 \cdot \text{span}(\mathbf{a})$ bits of the sequence $\mathbf{a}$.

The algorithm is efficient: its time complexity is $O(n^2)$, where n is the number of known symbols of the sequence. Furthermore, the algorithm is adaptive: each time a new bit is determined (say by a known plaintext attack), it can be used to update the previously determined LFSR in linear worst case time. Thus the number of available bits does not need to be known ahead of time.

One might hope that modifying the Berlekamp-Massey algorithm so it uses integer arithmetic rather than polynomial arithmetic would result in an FCSR synthesis algorithm. However, this simple approach fails because the N-ary expansion of the sum of two integers may involve larger powers of N than either of the two numbers. In fact the resulting algorithm fails to converge in general.

The Berlekamp-Massey algorithm is equivalent to finding the continued fraction expansion in the field $\mathbb{Z}/(N)[[X]]$ of formal power series of the element $A(X) = \sum_{i=0}^{\infty} a_i X^i \in \mathbb{Z}/(N)[[X]]$. See Section 18.2.d. (Here N is a prime number.) One might also hope that the continued fraction expansion in the field $\mathbb{Q}_N$ of N-adic numbers of the element $a = \sum_{i=0}^{\infty} a_i N^i$ would exhibit similar optimality properties and lead to an FCSR synthesis algorithm, but this is also false. In fact, the continued fraction expansion may also fail to converge, (see Section 5.7.c).

In what follows we solve the FCSR synthesis problem by finding an optimal rational representation for an N-adic integer. Assume we have consecutive terms $a_0, a_1, \cdots$ of an N-ary sequence $\mathbf{a}$ which we think of as coefficients of an N-adic integer a. We determine a pair of integers $f = (f_1, f_2)$ so that $a = f_1/f_2$ and so that $\Phi(f)$ is minimal among all such pairs of integers. The corresponding FCSR may then be constructed as described in Section 7.3. In Chapter 20 another approach to FCSR synthesis, based on the Berlekamp-Massey algorithm is presented, one that is even applicable to many more general AFSRs.

[1] If you're like me, you were about ready to give up when we got to part 7 or 8.

The FCSR synthesis problem was essentially studied in the context of Hensel and arithmetic codes [71, 113, 128], although no analysis of the aforementioned optimality properties was done in this early work.

19.3.a Lattices and approximation

Several approaches to the rational approximation problem for FCSRs can be best described in terms of integer approximation lattices. This notion is due to Mahler [125] and deWeger [188]. See Section 2.2.k for generalities on lattices.

Definition 19.3.1. *Let $a = a_0 + a_1 N + \cdots \in \mathbb{Z}_N$ be an N-adic integer. Its* kth approximation lattice *is the set*

$$L_k = L_k(\mathbf{a}) = \{(h_1, h_2) \in \mathbb{Z} \times \mathbb{Z} : ah_2 - h_1 \equiv 0 \pmod{N^k}\}.$$

Theorem 19.3.2. *The set L_k is a full lattice, its volume is N^k, and $L_k \supset L_{k+1} \supset \cdots$.*

Proof. The vectors $(a, 1)$ and $(N^k, 0)$ are in L_k. They form a basis, for if $(h_1, h_2) \in L_k$ then $ah_2 - h_1 = tN^k$ for some integer t, so

$$(h_1, h_2) = h_2(a, 1) - t(N^k, 0).$$

It follows that L_k is a full lattice. It contains L_{k+1} because a congruence modulo N^{k+1} implies the same congruence modulo N^k. The volume of L_k is the determinant of the matrix whose rows are $(a, 1)$ and $(N^k, 0)$, and this determinant is N^k. □

If $f = (f_1, f_2) \in L_k$ then $Nf = (Nf_1, Nf_2) \in L_{k+1}$. An element $(f_1, f_2) \in L_k$ with f_2 relatively prime to N represents a fraction f_1/f_2 whose N-adic expansion agrees with that of a in the first k places. Any rational approximation algorithm must therefore attempt to find, for each k, an element $u_k \in L_k$ that minimizes the function Φ. This is usually accomplished by keeping track of two elements, as follows.

Definition 19.3.3. *Let $L \subset \mathbb{R}^2$ be a full 2-dimensional lattice. A pair of elements $u, v \in L$ is called a* pair of successive minima for L *if*

1. $\Phi(u) \leq \Phi(u')$ *for all* $u' \in L - \{(0, 0)\}$ *and*
2. $\Phi(v) \leq \Phi(v')$ *for all* $v' \in L - \mathbf{Z}u$

where $\Phi(f, q) = \max(|f|, |q|)$ is the Weil height of f/q. If $u = (u_1, u_2) \in \mathbb{R}^n$, then we say that u is positive *if $u_1 u_2 > 0$ and* negative *if $u_1 u_2 < 0$. If either coordinate is zero, then u is neither positive nor negative.*

Lemma 19.3.4. *Let $L \subset \mathbb{R}^2$ be a full lattice and let $u = (u_1, u_2)$ and $v = (v_1, v_2)$ be a pair of successive minima for L. Then u and v cannot both be positive, and they cannot both be negative. Moreover, they form a basis for L.*

Proof. The statement about signs is equivalent to $u_1u_2v_1v_2 \leq 0$. So suppose that $u_1u_2v_1v_2 > 0$. By possibly negating one or the other vector, we may assume that $u_1v_1 > 0$ and $u_2v_2 > 0$. but then

$$\Phi(u - v) < \max(|u_1|, |u_2|, |v_1|, |v_2|) = \Phi(v).$$

Since $u - v$ is independent of u, this contradicts the fact that u, v are successive minima for L.

Suppose u, v do not form a basis. Let $z = (z_1, z_2) \in L$ be a vector that is not an integer linear combination of u and v with $\mu(z) = |z_1| + |z_2|$ minimal. Then z is not an integer multiple of u, so $\Phi(z) \geq \Phi(v) \geq \Phi(u)$. It follows from the first part of the lemma that the coordinates of one of the four vectors σu, σv with $\sigma = \pm 1$ do not have opposite signs to the signs of the corresponding coordinates of z (that is, they have the same sign or are zero). Let w be this vector. The vector $z - w$ is not an integer linear combination of u and v, but $\mu(z - w) < \mu(z)$, which is a contradiction. □

A pair (u, v) of successive minima is also called a *minimal basis for L*. Suppose that x and y are any two linearly independent elements of the full 2 dimensional lattice L. Recall from Section 2.2.k that $D_{\{x,y\}}$ is the absolute value of the determinant of the matrix $M_{\{x,y\}}$ whose rows are x and y. Then

$$\mathrm{vol}(L) \leq D_{\{x,y\}} \leq 2\Phi(x)\Phi(y). \tag{19.13}$$

By Theorem 2.2.27, the first inequality is an equality if and only if x and y form a basis. We can now give easily testable conditions to guarantee that we have a minimal basis or a minimal element in a lattice.

Theorem 19.3.5. *Let $L \subset \mathbb{R}^2$ be a full lattice and let $u \in L - \{0\}$ be a Φ-minimizing vector. If $x \in L - \{0\}$ and $\Phi(x) < \sqrt{\mathrm{vol}(L)/2}$ then x is an integer multiple of u, and the only other Φ-minimizing vector in L is $-u$. If $x, y \in L$ are linearly independent, $\Phi(x) < \mathrm{vol}(L)^{1/2}$, and $\Phi(y) < \mathrm{vol}(L)^{1/2}$, then x and y are successive minima for L and any other successive minima for L coincide with (x, y) up to change of sign and order.*

Proof. We prove the second statement first. We claim that x and y have opposite sign and they form a basis of L. Suppose x and y do not have opposite signs, so that $x_1x_2y_1y_2 \geq 0$. Then x_1y_2 and x_2y_1 do not have different signs, so

$$\mathrm{vol}(L) \leq |x_1y_2 - x_2y_1| = ||x_1y_2| - |x_2y_1|| \leq \max(|x_1y_2|, |x_2y_1|) \leq \Phi(x)\Phi(y) < \mathrm{vol}(L)$$

which is a contradiction. So x_1y_2 and x_2y_1 have different signs, hence

$$D_{\{x,y\}} = |x_1y_2 - x_2y_1| = |x_1y_2| + |x_2y_1| \leq 2\Phi(x)\Phi(y) < 2\mathrm{vol}(L).$$

By Theorem 2.2.27, $D_{\{x,y\}}$ is an integral multiple of $\mathrm{vol}(L)$ so it equals $\mathrm{vol}(L)$ which implies (again by Theorem 2.2.27) that $\{x, y\}$ form a basis of L, proving the claim.

Now suppose that $\{u, v\}$ are successive minima for L. Then $\Phi(u) \leq \Phi(x) < \sqrt{\mathrm{vol}(L)}$. The vector u cannot be opposite in sign to both x and y. By symmetry we may assume that u and x do not have opposite sign. The preceding claim (applied to x and u) implies that x and u are linearly dependent: one is a multiple of the other. But they are both vectors in a basis, so Lemma 2.2.29 implies that each is an integral multiple of the other, i.e. $x = \pm u$. Similarly, by the minimality properties of v we have $\Phi(v) \leq \Phi(y) < \sqrt{\mathrm{vol}(L)}$. Thus the same argument gives $y = \pm v$ as claimed.

Consider the first statement. Since $\Phi(u) \leq \Phi(x) < \sqrt{\mathrm{vol}(L)/2}$, part (2) implies that either u and x are linearly dependent, or they are successive minima of L. In the latter case, equation (19.13) gives $\mathrm{vol}(L) \leq 2\Phi(x)\Phi(u) < 2\mathrm{vol}(L)/2$ which is a contradiction. Hence x, u are linearly dependent, and again by Lemma 2.2.29, x is an integer multiple of u. Finally, if v is also Φ-minimizing then the same argument implies that u, v are integer multiples of each other, so $u = \pm v$. □

Corollary 19.3.6. *Let* $\mathbf{a} = a_0, a_1, \cdots$ *be an eventually periodic* N*-ary sequence with* N*-adic complexity* $\varphi_N(\mathbf{a})$*. Then the sequence* $\mathbf{a}$ *is completely determined by the first* k *symbols, where*

$$k = \begin{cases} 2\lceil \varphi_N(\mathbf{a}) \rceil + 1 & \text{if } N > 2 \\ 2\lceil \varphi_2(\mathbf{a}) \rceil + 2 & \text{if } N = 2. \end{cases}$$

Proof. Let $f/q = a_0 + a_1N + \cdots$ be the N-adic integer corresponding to $\mathbf{a}$ with f and q relatively prime. The vector (f, q) is in the approximation lattice $L_k = L_k(\mathbf{a})$ for all $k \geq 1$. Exponentiating the above equation for k gives $2\Phi(f, q)^2 < N^k = \mathrm{vol}(L_k)$. Theorem 19.3.5 then implies that (f, q) is an integer multiple of the unique (up to sign) Φ-minimizing vector in L_k. Hence the quotient f/q is determined by the lattice L_k, which in turn is determined by the first k symbols in the sequence $\mathbf{a}$. □

19.3.b Rational approximation via Euclid's algorithm

Suppose that we are given an N-ary sequence $\mathbf{a}$ and its associated N-adic integer $a = a_0 + a_1N + \cdots$. For each $k \geq 1$ we want to find integers f_k, g_k with $f_k/g_k \equiv a \pmod{N^k}$, $\gcd(g_k, N) = 1$, using Theorem 19.3.5 to guarantee that $\Phi(f_k, g_k) = \max(|f_k|, |g_k|)$ is as small as possible. For $N = 2$ this is accomplished using the lattice approximation algorithm in Section 19.3.c below. For $N > 2$

the idea is to start with a pair of rational approximations and use Euclid's algorithm to find a linear combination of these which is a better approximation and which minimizes Φ.

Unfortunately this is not quite enough. In order that a rational approximation represents an FCSR, we must have the denominator congruent to -1 modulo N, or at least relatively prime to N if we use the slightly more general MWC generator of Section 7.7.a. To help study this question we define

$$L_k^t = \{u = (u_1, u_2) \in L_k : u_2 \equiv t \pmod{N}\},$$

for $t = 0, 1, \cdots, N-1$. Then the sets L_k^t partition L_k. L_k^0 is a sublattice of L_k since it is closed under addition. It contains NL_{k-1}, so it is a full lattice. Suppose that $d = \gcd(N, t) > 1$. If $(r, v) \in L_k^t$, then $d|v$ and $d|r$, so that $(r, v) \in dL_{k-1}$.

Recall the extended Euclidean algorithm from Section 2.2.g. Suppose that the first k symbols $a_0, a_1, \cdots, a_{k-1}$ of an N-ary sequence are available. We will execute the extended Euclidean algorithm with $a = N^k$ and $b = \sum_{i=0}^{k-1} a_i N^i$. We obtain sequences of integers r_i, u_i, and v_i. By equation (2.6), $v_i b - r_i \equiv 0 \pmod{N^k}$ so that $(r_i, v_i) \in L_k$. The $|r_i|$ are decreasing while the $|v_i|$ are increasing, so we just wait until $|r_i|$ is less than $2^{(k-1)/2}$. The resulting algorithm is given in Figure 19.1.

Theorem 19.3.7. *Suppose that N is not a square and the infinite N-adic complexity of the sequence $a_0, a_1, \cdots$ is less than or equal to n, algorithm* EEApprox *is executed with $k \geq 2n+3$, and the algorithm outputs a pair of integers (r_1, y_1). Then*

$$c \stackrel{def}{=} \sum_{i=0}^{\infty} a_i N^i = \frac{r_1}{y_1},$$

r_1 and y_1 are relatively prime, and $\gcd(N, y_1) = 1$.

Proof. Let $c = p/q$ with p and q relatively prime. Then q is relatively prime to N and (p, q) is the minimal nonzero element of the k^{th} approximation lattice L_k for c. By hypothesis $\Phi(p, q) \leq N^n \leq N^{(k-3)/2}$. We may assume that $p > 0$.

Suppose the algorithm outputs (r_1, y_1). By construction we have $0 \leq r_1 \leq N^{k/2} < r_0$. However, $N^{k/2}$ is not an integer, so $r_1 < N^{k/2}$. By Lemma 2.2.17 we have $|r_0 y_1| \leq N^k$. Thus $|y_1| < N^{k/2}$, so that $\Phi(r_1, y_1) < N^{k/2}$.

Suppose that (p, q) and (r_1, y_1) are linearly independent. Then by Theorem 19.3.5, (p, q) and (r_1, y_1) are successive minima of the kth approximation lattice (in some order). In particular, by Lemma 19.3.4 q and y_1 have opposite signs. Therefore

$$N^k = |py_1 - qr_1| = |py_1| + |qr_1| < N^{(k-1)/2} N^{k/2} + N^{(k-1)/2} N^{k/2} = N^k.$$

This is a contradiction, so (p, q) and (r_1, y_1) are linearly dependent.

EEAPPROX$(a_0, a_1, \cdots, a_{k-1})$
begin
if (k is not odd) **then**
$k = k - 1$
fi
$(r_0, x_0, y_0) = (N^k, 1, 0)$
$(r_1, x_1, y_1) = (\sum_{i=0}^{k-1} a_i N^i, 0, 1)$
while $(r_1 > N^{k/2})$ **do**
Let $r_0 = qr_1 + r$
$(x_3, y_3) = (x_0 - qx_1, y_0 - qy_1)$
$(r_0, x_0, y_0) = (r_1, x_1, y_1)$
$(r_1, x_1, y_1) = (r, x_3, y_3)$
od
if ($|y_1| \leq 2^{k/2}$ and $y_1 \equiv 1 \pmod{N}$) **then**
return(r_1, y_1)
else
return(FALSE)
fi
end

Figure 19.1: The Euclidean Rational Approximation Algorithm.

Thus $(r_1, y_1) = \lambda(p, q)$ for some real number λ. By the minimality of $\Phi(p, q)$ in L_k, (p, q) is an element of a basis for L_k. By Lemma 2.2.17, (r_1, y_1) is in L_k, so by Lemma 2.2.29, λ is an integer.

By Lemma 2.2.17 we have

$$u_1 N^k + y_i \sum_{i=0}^{k-1} a_i N^i = r_i.$$

Thus

$$u_1 N^k = \lambda \left(p - q \sum_{i=0}^{k-1} a_i N^i \right).$$

But $\gcd(u_1, y_1) = 1$, so $\lambda | N^k$. It follows that

$$u_1 \frac{N^k}{\lambda} + q \sum_{i=0}^{k-1} a_i N^i = p.$$

We also have a relation

$$wN^k + q\sum_{i=0}^{k-1} a_i N^i = p$$

with $w \in \mathbb{Z}$. Thus $u_1/\lambda = w$ is an integer. Since $\gcd(u_1, y_1) = 1$ and r_1 and p are positive, we must have $\lambda = \pm 1$. This proves the theorem. □

If $N = 2$, then it follows that given $2\lambda(\mathbf{a}) + 3$ bits, the Euclidean rational approximation algorithm outputs a description of the smallest FCSR that generates $\mathbf{a}$. If $N \neq 2$ and N is not a square, then the Euclidean rational approximation algorithm outputs a description of the smallest AFSR that generates $\mathbf{a}$, but this AFSR may not be an FCSR — the connection integer may not be in the form

$$q = \sum_{i=1}^{n} q_i N^i - q_0 \tag{19.14}$$

with $q_0 = 1$. However, if we multiply q as in equation (19.14) by an integer u with $u \equiv q_0^{-1} \pmod{N}$ and $1 \leq |u| < N/2$, then we obtain another connection integer for $\mathbf{a}$. Thus if (p, q) is the output from the Euclidean rational approximation algorithm, then up/uq is the rational number corresponding to an FCSR that outputs $\mathbf{a}$. Moreover, $\Phi(up, uq)$ is within a factor $N/2$ of $\Phi(p', q')$ where p'/q' is the smallest rational representation of an FCSR that outputs $\mathbf{a}$. Thus $\log_N(\Phi(up, uq))$ is at most $1 - \log_N(2)$ plus the N-adic complexity of $\mathbf{a}$.

The time complexity is an immediate consequence of the analysis of the Euclidean algorithm given in Theorem 2.2.16.

Theorem 19.3.8. *The Euclidean rational approximation algorithm runs in time $O(k^2)$ if k elements of $\mathbf{a}$ are used.*

19.3.c Rational approximation via lattice approximation

In this section we present an analog of the Berlekamp-Massey algorithm for FCSRs with $N = 2$ which has both optimality properties: It constructs the smallest FCSR which generates the sequence $\mathbf{a}$ (Theorem 19.3.9), and it does so using only a knowledge of the first $2\varphi_2(\mathbf{a}) + 2\log_2(\varphi_2(\mathbf{a}))$ elements of $\mathbf{a}$, where $\varphi_2(\mathbf{a})$ is the 2-adic span of $\mathbf{a}$ (Theorem 19.3.10). This algorithm is based on 2-adic approximation theory. It is a modification of the procedure outlined by de Weger [188] and Mahler [125], which has the advantage that it is adaptive in the same sense as the Berlekamp-Massey algorithm.

In the rational approximation algorithm, given in Figure 19.3.c, and in the rest of this section, if $f = (f_1, f_2)$ is a pair of integers and if $d \in \mathbb{Z}$ is an integer, write $df = (df_1, df_2)$.

```
RationalApproximation(a_0, a_1, ..., a_{T-1})
    begin
    a = Σ_{i=0}^{T-1} a_i 2^i
    Let t be minimal with a_{t-1} ≠ 0
    f = (0, 2)
    g = (a_{t-1} 2^{t-1}, 1)
    for (k = t, ..., T-1) do
        if (a·g_2 − g_1 ≡ 0 (mod 2^{k+1})) then
            if (Φ(f) < Φ(g)) then
                f = 2f
            else Let d minimize Φ(f + dg)
                ⟨g, f⟩ = ⟨g, 2(f + dg)⟩
        else
            Let f_1 g_2 − f_2 g_1 = ε 2^k, where ε ∈ {1, −1}
            if (Φ(g) < Φ(f)) then
                Let d minimize Φ(f + dg) with d odd
                ⟨g, f⟩ = ⟨f + dg, 2g⟩
            else
                Let d minimize Φ(g + df) with d odd
                ⟨g, f⟩ = ⟨g + df, 2f⟩
        fi fi
    od
    return g
    end
```

Figure 19.2: Rational Approximation Algorithm for 2-Adic integers.

Implementation remarks. The congruence $ag_2 - g_1 \equiv 0 \pmod{2^{k+1}}$ may be checked without performing the full multiplication at each stage, by saving and updating the previous values of $ag_2 - g_1$ and $af_2 - f_1$. Inside the loop, in the second and third cases, the integer d is chosen to minimize $\Phi(f+xg)$ (respectively, $\Phi(g+xf)$) among all possible integers x with a particular residue modulo 2. As observed in [188], it may be computed by division. For example, suppose we are in the second case: $ag_2 - g_1 \not\equiv 0 \pmod{2^{k+1}}$ and $\Phi(g) < \Phi(f)$. If $g_1 \neq \pm g_2$, then d is among the integers with the given residue that are immediately less than or greater than $(f_1 - f_2)/(g_1 - g_2)$ and $-(f_1 + f_2)/(g_1 + g_2)$. Thus it suffices to consider the value of $\Phi(f + dg)$ for these four values of d. When $g_1 = \pm g_2$, one or the other of these quotients is not considered. If $\Phi(g) > \Phi(f)$ then

the roles of f and g are switched.

Theorem 19.3.9. *Let* RationalApproximation$(a_0, \cdots, a_{T-1})$ *output* $g = (g_1, g_2)$. *Then* g_2 *is odd,*

$$(\sum_{i=0}^{T-1} a_i 2^i) \cdot g_2 - g_1 \equiv 0 \pmod{2^T},$$

and any other pair $g' = (g'_1, g'_2)$ *which satisfies these two conditions has* $\Phi(g') \geq \Phi(g)$.

Theorem 19.3.10. *Suppose* $\mathbf{a} = (a_0, a_1, a_2, \cdots)$ *is an eventually periodic sequence with associated 2-adic integer* $a = \sum a_i 2^i = f/q$, *with* $f, q \in \mathbb{Z}$, *and* $\gcd(f, q) = 1$. *If* $T \geq \lceil 2\varphi_2(\mathbf{a}) \rceil + 2$, *then* RationalApproximation$(a_0, \cdots, a_{T-1})$ *outputs* $g = (f, q)$. *Hence also his occurs if* $T \geq 2\lambda_2(\mathbf{a}) + \lceil 2\log_2(\lambda_2(\mathbf{a})) \rceil + 2$.

The proofs of these two optimality results occupy the remainder of this section, The proof of Theorem 19.3.9 is an immediate consequence of the following lemma.

Lemma 19.3.11. *For each* k, *at the top of the loop the following conditions hold:*

1. f *and* g *are in* L_k, f_1 *and* f_2 *are even,* g_2 *is odd, and* f *is not a multiple of* g;
2. $\langle f, g \rangle$ *is a basis for* L_k;
3. $f \notin L_{k+1}$;
4. g *minimizes* $\Phi(h)$ *over all elements* $h \in L_k$ *with* h_2 *odd;*
5. f *minimizes* $\Phi(h)$ *over all elements* $h \in L_k$ *with* h_1 *and* h_2 *even and* h *not a multiple of* g.

Proof. We prove this by induction on k. It is straightforward to check that the conditions (1)-(5) hold initially. Let us suppose that the conditions hold at stage k. We leave it as an exercise to prove that if $g \in L_{k+1}$, then after passing through the loop and returning to the top, the new values of f and g satisfy conditions (1) to (5). Therefore, assume we are at stage k, and that $g \notin L_{k+1}$. Let f' and g' be the new values after updating. We treat the case when $\Phi(g) < \Phi(f)$. The other case is similar.

1. We have

$$\begin{aligned} a \cdot g'_2 - g'_1 &= a \cdot (f_2 + dg_2) - (f_1 + dg_1) \\ &= (a \cdot f_2 - f_1) + d(a \cdot g_2 - g_1) \\ &\equiv 2^k + 2^k \pmod{2^{k+1}} \\ &\equiv 0 \pmod{2^{k+1}}. \end{aligned}$$

Therefore $g' \in L_{k+1}$. Also, g is in L_k, so $f' = 2g$ is in L_{k+1}. The divisibility conditions on f and g are straightforward to check.

2. By Lemma 2.2.27, we have $f_1g_2 - f_2g_1 = \pm 2^k$. Therefore $f_1'g_2' - f_2'g_1' = 2g_1(f_2 + dg_2) - 2g_2(f_1 + dg_1) = 2(f_1g_2 - f_2g_1) = \pm 2^{k+1}$. Again by Lemma 2.2.27, $\langle g', f' \rangle$ is a basis for L_{k+1}.

3. We have $g \notin L_{k+1}$, so $f' = 2g \notin L_{k+2}$.

4. Suppose that minimality fails. Since $\langle f', g' \rangle$ is a basis for L_{k+1}, there are integers a and b so that

$$\Phi(ag' + bf') < \Phi(g') \tag{19.15}$$

and $ag_2' + bf_2'$ is odd. The latter condition is equivalent to a being odd since f_2' is even and g_2' odd. We can assume that a and b are relatively prime since dividing both by a common divisor can only reduce the resulting value of Φ. By possibly negating both a and b, we can assume a is nonnegative. Furthermore, if $a = 1$, then $ag' + bf' = f + (d + 2b)g$ and this contradicts the choice of d in the algorithm. Thus we can assume that $a > 1$. Equation (19.15) can be rewritten

$$\Phi(af + (ad + 2b)g) < \Phi(f + dg). \tag{19.16}$$

Let c be the integer closest to $d + 2b/a$ that is odd. Since a is odd, the quantity $x = c - (d + 2b/a)$ is not an integer, so satisfies $|x| < 1$ hence $|x| \leq (a - 1)/a$. Then

$$\begin{aligned} \Phi(af + acg) &= \Phi(af + (ad + 2b + ax)g) \\ &\leq \Phi(af + (ad + 2b)g) + a|x|\Phi(g) \end{aligned}$$

by the triangle inequality for Φ. The first term may be bounded using (19.16) and the second term is bounded by the induction hypothesis: $\Phi(g) \leq \Phi(f + dg)$. Dividing by a gives

$$\Phi(f + cg) < \frac{1}{a}\Phi(f + dg) + |x|\Phi(f + dg) \leq \Phi(f + dg)$$

which contradicts the choice of d.

5. Suppose there is an element $h' \in L_{k+1}$ with both h_1' and h_2' even, such that $\Phi(h') < \Phi(f') = 2\Phi(g)$. We can write $h' = 2h$ for some $h \in L_k$, so $\Phi(h) < \Phi(g) < \Phi(f)$. If both h_1 and h_2 are even, this contradicts the minimality of f. If h_2 is odd, this contradicts the minimality of g. It is impossible that h_1 is odd and h_2 is even for $k \geq 1$ since $a \cdot h_2 - h_1 \equiv 0 \pmod{2^k}$. □

Remarks. The algorithm runs correctly if we always update g and f by the first method ($\langle g, f\rangle = \langle f + dg, 2f\rangle$), independent of the relation between $\Phi(g)$ and $\Phi(f)$. The relation $\Phi(g) < \Phi(f)$ was only used to verify property (5) above, which is not necessary for rapid convergence of the algorithm. However, property (5) ensures that the size of f remains small, so it leads to better bounds on the complexity of the computations which are involved in the algorithm. Since the algorithm is adaptive there is, of course, no need to assume that the sequence $\mathbf{a}$ is eventually periodic.

Proof of Theorem 19.3.10. By assumption, $a = f/q$ so q is odd and $(f, q) \in L_k$ for all k. The output from the algorithm is a pair $g = (g_1, g_2) \in L_T$ which is Φ-minimal with g_2 odd, so $\Phi(g_1, g_2) \leq \Phi(f, q)$. Hence

$$|g_1 q| = |g_1||q| \leq \Phi(g_1, g_2) \cdot \Phi(f, q) \leq \Phi(f, q)^2 \leq 2^{T-2},$$

since by assumption $T \geq 2\log_2 \Phi(f, q) + 2$. Similarly, $|fg_2| \leq 2^{T-2}$. However, $ag_2 - g_1 \equiv 0 \pmod{2^T}$ so $g_1 q \equiv fg_2 \pmod{2^T}$, which implies that $g_1 q = fg_2$. Therefore (g_1, g_2) is some multiple of (f, q) by an integer that is odd. By Φ-minimality, this integer must be ± 1 which gives $g_1 = f$ and $g_2 = q$ (or else $g_1 = -f$ and $g_2 = -q$). □

Complexity Issues Suppose the rational approximation algorithm is executed with a sequence $\mathbf{a}$ which is eventually periodic, with rational associated 2-adic integer $a = f/q$. Then the rational approximation algorithm takes $T = 2\log_2(\Phi(p, q)) + 2 \leq 2\lambda_2(\mathbf{a}) + 2\lceil \log_2(\lambda_2(\mathbf{a}))\rceil - 2$ steps to converge.

Consider the kth step. If $ag_2 - g_1 \not\equiv 0 \pmod{2^{k+1}}$, then we say that a discrepancy has occurred. The complexity of the algorithm depends on the number of discrepancies. To simplify the computation of ag_2, we maintain af_2 as well. When no discrepancy occurs, these values and the value of f can be updated with k bit operations.

Suppose a discrepancy occurs. The minimization step can be done with two divisions of k bit integers. The remaining steps take time $\mathcal{O}(k)$. Then ag_2 and af_2 can be updated with $\mathcal{O}(k)$ bit operations and two multiplications of k bit integers by d.

Let D be the number of discrepancies, and let M be the maximum time taken by a multiplication or division of T bit integers. The Schönhage-Strassen algorithm [177], gives $M = \mathcal{O}(T \log T \log\log T)$. This can be improved to $M \sim T \log T$ using Pollard's non-asymptotic algorithm and Newton interpolation for $T < 2^{37}$ on a 32 bit machine or $T < 2^{70}$ on a 64 bit machine [160]. These are ranges that are typical in current usage.

The complexity of the algorithm is thus $4DM + \mathcal{O}(T^2)$. Strictly in terms of T, this is $\mathcal{O}(T^2 \log T \log\log T)$. However, in practice the number of discrepancies is often much smaller and the complexity is lower. In particular a cryptographer designing a stream cipher should try to choose sequences for which many discrepancies occur. This is equivalent to the 2-adic complexity profile staying close to the line with slope 1/2.

19.3.d Cryptanalysis of the summation cipher

In the summation cipher [135, 169], several binary m-sequences $\mathbf{a}_1, a_2, \cdots, a_k$ are combined using "addition with carry". The resulting sequence is used as a pseudo-one-time-pad. These sequences have generated great interest since they appear to be resistant to attacks based on the Berlekamp-Massey algorithm. If the constituent sequences $\mathbf{a}_i$ have coprime periods T_i then the resulting sequence has linear span which is close to the period $L = T_1 \cdot T_2 \cdots T_k$ of the combined sequence.

However, by a generalization of Theorem 19.1.5, the 2-adic complexity of the combined sequence is no more than $T_1 + T_2 + \cdots + T_k + \log_2(k)$ so the 2-adic span is no more than

$$\sum T_i + 2\log_2(k) \sum_i \log_2(T_i) + (5 + \log_2(3))k.$$

Thus if the T_i are similar in magnitude, the 2-adic span of the result is bounded by

$$kL^{1/k} + 2\log_2(k)\log_2(L) + O(k)$$

and it may be much less. This throws considerable doubt on the security of these stream ciphers.

Here is a more algorithmic description of the attack:

D1. Determine (perhaps by a known plaintext attack on a stream cipher system) $2 \cdot T$ consecutive bits of the sequence $\mathbf{b}$ formed by applying the summation combiner to $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_k$.
D2. Apply a rational approximation algorithm to this sequence of bits, to find q and f.
D3. Construct the FCSR which outputs the bit stream corresponding to the 2-adic integer $a = f/q$ using the methods in Sections 7.2 and 7.3.

The resulting FCSR uses at most $T = T_1 + T_2 + \cdots + T_k + \mathcal{O}(\log(\sum_i (T_i))) + \mathcal{O}(\log(k))$ bits of storage and outputs the sequence $\mathbf{b}$. The effectiveness of this attack may be minimized by using only $k = 2$ constituent m-sequences, with periods T_i chosen so that $2^{T_i} - 1$ has no small prime factors, but then there is less benefit in the growth of the linear span.

For example, if we combine m-sequences of period $2^L - 1$ with $L = 7$, 11, 13, 15, 16, and 17, then the period and linear span of the resulting sequence are nearly 2^{79}. However, the 2-adic span is less than 2^{18}, so fewer than 2^{19} bits suffice to find a FCSR that generates the sequence. The maximum number of single word arithmetic operations performed by the rational approximation algorithm on a 32 bit machine is about 2^{42}.

The summation combiner is also vulnerable to correlation attacks [145, 146].

19.4 Exercises

1. Prove Proposition 19.1.4.

2. Prove that in the Rational Approximation Algorithm, if $g \in L_{k+1}$, then the new values of f and g satisfy conditions (1) through (5) of Lemma 19.3.11.

3. Suppose that the Berlekamp-Massey Algorithm of Figure 18.1 is modified into a 2-adic version as follows: replace all occurrences of x by 2; define $\Phi(a, b) = \lfloor \log_2(\max(|a|, |b|)) \rfloor$. Where does the proof of correctness in Section 18.2.b (with x replaced by 2 and Φ defined as above) break down?

Chapter 20 AFSR Synthesis

Many of the ingredients that go into the Berlekamp-Massey algorithm exist in the general setting of AFSRs. In particular, we still have the viewpoint of register synthesis as rational approximation, so we can try to solve the rational approximation problem. Fix a ring R and an element $\pi \in R$ so that $R/(\pi)$ is finite. Throughout this section we assume that R is an integral domain. Also fix a complete set $S \subset R$ of representatives for $R/(\pi)$ so the π-adic expansion of an element u/q makes sense, where $u, q \in R$ are coprime.

In this chapter we fix an eventually periodic sequence $\mathbf{a} = a_0, a_1, a_2, \cdots$ of elements $a_i \in R/(\pi)$ and denote the corresponding π-adic element by a, so that $\mathbf{a} = \mathbf{seq}_\pi(a)$, that is,

$$a = a_0 + a_1\pi + a_2\pi^2 + \cdots \in R_\pi.$$

The rational approximation problem in this context is:

Rational Approximation
Instance: A prefix of the eventually periodic sequence $\mathbf{a} = a_0, a_1, \cdots$ over $R/(\pi)$.
Problem: Find $f, q \in R$ such that $\mathbf{a} = \mathbf{seq}_\pi(f/q)$.

We can try to proceed by constructing successively better rational approximations until the exact rational representation of the π-adic number associated with the given sequence is found. As in the Berlekamp-Massey algorithm, at each stage we must deal with the discrepancy, and we can try to do so by forming a linear combination between two previous rational approximations.

The Berlekamp-Massey algorithm works in part because there is a precise algebraic interpretation of the linear span – it is the minimal value of $\max(\deg(f) + 1, \deg(q))$, minimized over all rational representations f/q of $a = a_0 + a_1\pi + \cdots \in R_\pi$. Moreover, the effect of various standard algebraic operations on f and q is well controlled – multiplying by x increases the degree by 1, multiplying two polynomials adds their degrees, and the degree of the sum two polynomials is bounded by the maximum of the degrees of the original polynomials.

In many rings R there is no reasonable analog of degree. In some cases there are candidates for analogs of degree, but these are not so well behaved. For integers $u, \pi \in \mathbb{Z}$, for example (the FCSR case), we might use $\log_\pi(|u|)$ or, if we want an integer, $\lceil \log_\pi(|u|) \rceil$. There are problems with these measures, however. First, they do not exactly give the length of the smallest FCSR over $\mathbb{Z}$ and π that outputs the sequence. As it turns out, this is not a serious problem since the size of the smallest FCSR that generates $\mathbf{a}$ differs from these quantities by an additive log term (see

Proposition 19.1.3). More serious is the fact that this quantity is less well behaved than polynomial degree. For example, $\deg(f+g) \le \max(\deg(f), \deg(g))$, but the best we can say about logs is that $\log(|u+v|) \le \max(\log(|u|), \log(|v|)) + 1$. This introduces additive error terms that affect proofs of convergence.

20.1 Xu's rational approximation algorithm

Despite the objections above, in this section we describe a modification of the Berlekamp-Massey algorithm that works for many AFSRs over pairs (R, π), even if R_π has carry and π is not prime. This is accomplished with two main modifications. First, the linear combination $(h_j, r_j) + d\pi^{j-m}(h_m, r_m)$ is replaced by a more general linear combination $d_1(h_j, r_j) + d_2\pi^{j-m}(h_m, r_m)$, with d_1, d_2 chosen from a fixed small set. Second, we control the growth of the approximations by producing a new approximation that works for several new terms at once, thus compensating for the increase in size due to carry when we form these linear combinations. To make this work we need two structures:

1. a measure of the "size" of elements of R
2. a small subset of R from which we can select the coefficients d_1 and d_2.

We next describe the properties these structures must have. In sections 20.4, 20.5, 20.6, and 20.8 we describe various rings that have these structures. To measure the "size" of elements, assume we have a function $\varphi_{R,\pi} : R \to \mathbb{Z} \cup \{-\infty\}$ satisfying the following properties.

Property 1: There are non-negative integers b and c such that

1. $\varphi_{R,\pi}(0) = -\infty$ and $\varphi_{R,\pi}(x) \ge 0$ if $x \ne 0$;
2. for all $x, y \in R$ we have $\varphi_{R,\pi}(xy) \le \varphi_{R,\pi}(x) + \varphi_{R,\pi}(y) + b$;
3. for all $x, y \in R$, we have $\varphi_{R,\pi}(x \pm y) \le \max\{\varphi_{R,\pi}(x), \varphi_{R,\pi}(y)\} + c$;
4. for all $x \in R$ and $k \ge 0 \in Z$, we have $\varphi_{R,\pi}(\pi^k x) = k + \varphi_{R,\pi}(x)$.

We use the convention that $-\infty + a = -\infty$ for every integer a. Such a function $\varphi_{R,\pi}$ is called an *index function*. From it, define a "height" function $\Gamma_{R,\pi}(x, y)$ for $x, y \in R$, by

$$\Gamma_{R,\pi}(x, y) = \max\{\varphi_{R,\pi}(x), \varphi_{R,\pi}(y)\},$$

and the "π-adic complexity" (which depends on the choices of S and $\varphi_{R,\pi}$) of a sequence,

$$\varphi_\pi(\mathbf{a}) = \inf\left\{\Gamma_{R,\pi}(u, q) : \ \mathbf{a} = \mathbf{seq}_\pi(u/q)\right\}.$$

The next proposition follows immediately from the definition.

Proposition 20.1.1. *For any two pairs (h_1, r_1), $(h_2, r_2) \in R \times R$ and integer $k \ge 0$,*

1. $\Gamma_{R,\pi}(h_1 + h_2, r_1 + r_2) \le \max\{\Gamma_{R,\pi}(h_1, r_1), \Gamma_{R,\pi}(h_2, r_2)\} + c$;
2. $\Gamma_{R,\pi}(h_1 r_2 - r_1 h_2, r_1 r_2) \le \Gamma_{R,\pi}(h_1, r_1) + \Gamma_{R,\pi}(h_2, r_2) + b + c$;
3. $\Gamma_{R,\pi}(\pi^k(h_1, r_1)) = k + \Gamma_{R,\pi}(h_1, r_1)$,

where b and c are the integers appearing in Property 1. □

In many cases $\varphi_\pi(\mathbf{a})$ grows at most linearly with the size of the AFSR that is needed to generate $\mathbf{a}$. Suppose an AFSR over R and π has connection element $q = \sum_{i=0}^{m} q_i \pi^i$ with $q_i \in R$, has an initial state $(a_0, a_1, \cdots, a_{m-1}; z)$, and produces an output sequence $\mathbf{seq}_\pi(u/q)$. Then u is given by equation (8.5). It follows from Property 1 that

$$\varphi_{R,\pi}(u) \le m + c + \max\{2c\lceil \log(m) \rceil + e + f + b, \varphi_\pi(z)\},$$

and

$$\varphi_{R,\pi}(q) \le m + c\lceil \log(m+1) \rceil + e$$

where $e = \max\{\varphi_\pi(q_i) : i = 0, 1, \cdots, m\}$, and $f = \max\{\varphi_\pi(x) : x \in S\}$. Typically $\varphi_{R,\pi}(z)$ is a measure of the space required to store the memory z. If this is the case and if the q_i are chosen from some finite set of allowable coefficients, then the above equations show that $\Gamma_{R,\pi}(u, q)$ is at most linear in the size of the AFSR. Thus if we can bound the execution time of a rational approximation algorithm in terms of $\Gamma_{R,\pi}(u, q)$, then we will have also bounded the execution time in terms of the size of the AFSR.

To control the growth of the size of a new approximation which is a combination of previous ones, we restrict the elements that are used to multiply the previous approximations and make the combination. To do so, we assume we have a subset $P_{R,\pi}$ of R with the following properties.

Property 2: There is an integer $B > 0$ such that

1. $0 \notin P_{R,\pi}$;
2. $b + c + \max\{\varphi_{R,\pi}(s) : s \in P_{R,\pi}\} < B$;
3. for every $h_1, h_2 \ne 0 \in R$, there exist $s, t \in P_{R,\pi}$ such that $\pi^B | (sh_1 + th_2)$;

Such a set $P_{R,\pi}$ is called an *interpolation set*. When there is no risk of ambiguity we drop the subscripts and simply write $\varphi = \varphi_{R,\pi}$, etc.

The first and second conditions together imply that for all $s \in P$, we have $\pi^B \not| s$. Let $d = \max\{\varphi(s) : s \in P\}$ and $C = b + c + d$. It follows from Property 1 that for any two pairs $(h_1, r_1), (h_2, r_2)$ and any $s, t \in P$, we have

$$\Gamma(s(h_1, r_1) + t(h_2, r_2)) \le \max\{\Gamma(h_1, r_1), \Gamma(h_2, r_2)\} + C.$$

If the third condition holds, then it holds even if $h_1, h_2 \in R_\pi$.

With these definitions and properties, Xu's rational approximation algorithm is given in Figure 20.1. The constant B is from Property 2.

$\textsc{Xu}(a_0, \cdots, a_k)$

begin

$a = 1 + \pi \sum_{i=0}^{k} a_i \pi^i$

$(h_0, r_0) = (0, 1)$

$(h_1, r_1) = (1 + a_0\pi + \cdots + a_{B-2}\pi^{B-1}, 1 + \pi^B)$

$m = 0$

for $(i = 1$ to $k - 1)$

  if $((h_i - r_i a) \not\equiv 0 \pmod{\pi^{i+1}})$ {

    if ($\exists s \neq 0 \in P$ with $(\pi^{i+B} \mid s(h_i - r_i a)))$

      $(h_{i+1}, r_{i+1}) = s(h_i, r_i)$

    else {

      Find $s, t \in P$, not both zero, with

        $\pi^{i+B} \mid s(h_i - r_i a) + t\pi^{i-m}(h_m - r_m a)$

      $(h_{i+1}, r_{i+1}) = s(h_i, r_i) + t\pi^{i-m}(h_m, r_m)$

    }

    if $(\Gamma(h_{i+1}, r_{i+1}) > \Gamma(h_i, r_i)$ and

        $\Gamma(h_i, r_i) \leq i - m + \Gamma(h_m, r_m)$ and $t \neq 0)$

      $m = i$

  }

Let $1 + \pi(u/q) = h_k/r_k$

Find the largest power t of π that divides both u and q

output $(u/\pi^t, q/\pi^t)$

end

Figure 20.1: Xu's Rational Approximation Algorithm.

Remarks. The algorithm maintains a rational element h_i/r_i that is an approximation to a, correct for the first i symbols of the π-adic expansion of a. At each stage we check whether this approximation is correct for the next symbol. If not, we make a correction using an earlier approximation. The new approximation is guaranteed to be correct not only for the new symbol but for at least B additional symbols.

At the start of the algorithm, we set $a \leftarrow 1+\pi a$. The purpose is to guarantee that $(h_0 - ar_0) \equiv 0$ modulo π^0 but not modulo π^1, and that there is no element $s \in P$ such that $\pi^B | s(h_0 - ar_0)$.

At the end of the algorithm we have a pair of elements $u, q \in R$ such that, if k is large enough, $u/q = \sum_{i=0}^{\infty} a_i \pi^i$ (this is proved in Theorem 20.3.2). Thus q is the connection element for an AFSR over R that outputs $\mathbf{a} = a_0, a_1, \cdots$. However this may not be the smallest such AFSR.

If R is a Euclidean domain (for example when $R = \mathbb{Z}$, or when R is the ring of integers in a number field with class number one, see Section 3.4.c or [17]) then we can find the greatest common divisor of u and q using the Euclidean algorithm and thus find the smallest such u and q with respect to the Euclidean size function. However we may still not have the smallest AFSR that outputs $\mathbf{a}$. The ring R might have an infinite group of units, so there might be infinitely many connection elements equivalent to q (in the sense that their AFSRs output the same sequences).

Even if R is not a Euclidean domain, Theorem 20.3.2 below states that the size of the AFSR produced by the algorithm is bounded by a constant (depending only on R, π, S, T, and the index function and interpolating set) times the size of the smallest such an AFSR.

20.2 Rational approximation in $\mathbb{Z}$

In this section we consider the case when $R = \mathbb{Z}$. We give an example of the execution of the algorithm that may help in understanding it. If $R = \mathbb{Z}$, then π is an integer (possibly composite). Let $S = \{a : 0 \le a \le \pi - 1\}$. If $x \neq 0$ and $|x| = a_0 + a_1\pi + \cdots + a_t\pi^t$ with $a_i \in S$ and $a_t \neq 0$, then we define $\varphi_{\mathbb{Z},\pi}(x) = t$. Equivalently, $\varphi_{\mathbb{Z},\pi}(x) = t$ if $\pi^t \le |x| < \pi^{t+1}$. Then Property 1 holds with $b = 1$ and $c = 1$. We also define

$$x \in P_{\mathbb{Z},\pi} \text{ if } |x| \le \begin{cases} 5 & \text{if } \pi = 2 \\ 15 & \text{if } \pi = 3 \\ \lfloor \pi^3/2 \rfloor & \text{if } \pi \ge 4. \end{cases}$$

Then Property 2 holds with $B = 5$ and $C = 4$. To see this in the case when $\pi \ge 4$, fix $x, y \neq 0 \in \mathbb{Z}$. By possibly negating x or y, we may assume that $x, y > 0$. Let $P^{\ge 0} = \{s \in P : s \ge 0\}$. Consider the function $\psi : P \times P \to \mathbb{Z}/(\pi^B)$ defined by $\psi(s,t) = sx + ty$. We have $|P^{\ge 0}| \ge (\pi^3 + 1)/2$, so $|P^{\ge 0} \times P^{\ge 0} - \{(0,0)\}| > \pi^6/4 \ge \pi^5 = |\mathbb{Z}/(\pi^5)|$. This implies that there exist $(s_1, t_1), (s_2, t_2) \in P^{\ge 0} \times P^{\ge 0} - \{(0,0)\}$ so that $\psi(s_1, t_1) = \psi(s_2, t_2)$. Hence if $s = s_1 - s_2$ and $t = t_1 - t_2$, then $\psi(s,t) = 0$ and $(s,t) \in P$, as we wanted. A similar argument works if $\pi = 2$ or 3.

Let $\pi = 10$. The sequence

$$\begin{aligned}\mathbf{a} \;=\;& 2\,7\,9\,8\,5\,4\,9\,9\,3\,3\,7\,4\,5\,7\,7\,0\,6\,4\,1\,2\,8\,1\,2\,2\,6\,0\,9\,5\,5\,0\,2\,8\,0\,1\,0\,2\,3\,5\,0\,9\,4\,4\,8\,7\\ & 0\,7\,5\,3\,6\,5\,5\,7\,8\,1\,8\,8\,8\,5\,3\,8\,7\,9\,5\,3\,9\,8\,1\,0\,1\,3\,4\,8\,5\,8\,2\,7\,8\,8\,4\,2\,6\,3\,2\,2\,8\,2\,3\,4\\ & 0\,8\,2\,1\,2\,6\,9\,7\,3\,1\,3\,8\,2\,4\,5\,2\,2\,0\,5\,7\,2\,5\;\cdots\end{aligned}$$

is the 10-adic expansion of the fraction $-52/1109$ with period 1108. For simplicity, we skip the shift step $\mathbf{a} \to 1 + 10\mathbf{a}$. The following steps show how the algorithm is initialized, how approximations are updated, and the simplification at convergence.

Initialization: $m = 0$, $(h_0, r_0) = (0, 1)$, $(h_1, r_1) = (972, 1001)$. The rational number $972/1001$ approximates $\mathbf{a}$ to at least the first 3 symbols.

First updating: Since $972/1001$ only approximates $\mathbf{a}$ to the first 3 symbols, at index $i = 4$ a new approximation is needed. We have $s = -44$ and $t = -39$. Then we have the new pair $(h_4, r_4) = (-42768, -434044)$, and now $(h_m, r_m) = (972, 1001)$. The rational number $42768/434044$ approximates $\mathbf{a}$ to at least the first 6 symbols.

Second updating: Since $42768/434044$ only approximates $\mathbf{a}$ to the first 6 symbols, at index $i = 7$ a new approximation is needed. We have $s = -50$ and $t = 50$. Then we have the new pair $(h_7, r_7) = (50738400, 71752200)$, and now $(h_m, r_m) = (-42768, -434044)$. The rational number $50738400/71752200$ approximates $\mathbf{a}$ to at least the first 9 symbols.

Third updating: Since $50738400/71752200$ only approximates $\mathbf{a}$ to the first 9 symbols, at index $i = 10$ a new approximation is needed. We have $s = -49$ and $t = -42$. Then we have the new pair $(h_{10}, r_{10}) = (-689925600, 14713990200)$, and now $(h_m, r_m) = (50738400, 71752200)$. The rational number $-689925600/14713990200$ approximates $\mathbf{a}$ to at least the first 12 symbols.

Convergence: The rational number $u/q = h_{10}/r_{10} = -689925600/14713990200$ gives $\mathbf{a}$ exactly.

Reduction: $\gcd(-689925600, 14713990200) = 13267800$. After factoring out the gcd, we have the reduced rational number $u/q = -52/1109$, as desired.

20.3 Proof of correctness

In this section we show that the algorithm outputs a correct rational representation of a when enough bits are given. We first show that the output is meaningful. The algorithm computes pairs (h_i, r_i) satisfying $h_i - ar_i \equiv 0 \pmod{\pi^i}$. We want to interpret (h_i, r_i) as a fraction $h_i/r_i \in K$, and hence as defining an AFSR whose output is $a_0, a_1, \cdots$. But this only makes sense if r_i is not zero.

Theorem 20.3.1. *For every j, $r_j \neq 0$.*

It remains to prove that if $\mathbf{a} = a_0, a_1, \cdots$ can be generated by an AFSR, then after some finite number of steps the algorithm outputs a description of such an AFSR. We say the algorithm is *convergent at index i* if $h_i/r_i = a$.

Let $\varphi = \varphi_\pi(\mathbf{a})$ denote the minimum value of $\Gamma(u, q)$ such that $\mathbf{a} = \mathbf{seq}_\pi(u/q)$.

Theorem 20.3.2. *Let $\mathbf{a} = \mathbf{seq}_\pi(u/q)$. In Xu's algorithm for the sequence $\mathbf{a}$, suppose that*

$$i \geq \frac{B(3b + 2c + B + c\lceil \log(B) \rceil + 2f_1)}{B - C} + \frac{2B}{B - C}\varphi_\pi(\mathbf{a}),$$

where $f_1 = \max\{\varphi_\pi(a) : a \in S\} \cup \{\varphi_\pi(1)\}$. Then the algorithm is convergent at i. That is, $h_i/r_i = u/q$.

The following definitions make the proofs of Theorems 20.3.1 and 20.3.2 simpler. For any $i \geq 0$, let $\mu(i) = i - \Gamma(h_i, r_i)$.

Definition 20.3.3. *Define an index to be a* turning point *as follows:*

1. *The initial index $m = 0$ is a turning point.*
2. *If m_1 is a turning point, then m_2 is the turning point following m_1 if it is the smallest integer greater than m_1 satisfying*
   - *(a) $(h_{m_2} - ar_{m_2}) \equiv 0 \ (mod\ \pi^{m_2})$, $\not\equiv 0 \ (mod\ \pi^{m_2+1})$;*
   - *(b) there is no $s \neq 0$ such that $s \in P$ and $\pi^{m_2+B} | s(h_{m_2} - ar_{m_2})$;*
   - *(c) $\Gamma(h_{m_2+1}, r_{m_2+1}) > \Gamma(h_{m_2}, r_{m_2})$;*
   - *(d) $\mu(m_1) \leq \mu(m_2)$.*

Conditions 20.3.3.2.a and 20.3.3.2.b hold with $m_2 = 0$. An index m is a turning point if it is either zero or it is one where the assignment $m \longleftarrow i$ occurs.

At an index i, if $h_i - ar_i \equiv 0 \pmod{\pi^i}$ but $h_i - ar_i \not\equiv 0 \pmod{\pi^{i+1}}$, then (h_{i+1}, r_{i+1}) is obtained either by multiplying (h_i, r_i) by an element $s \in R$, or as a linear combination $s(h_i, r_i) + t(h_m, r_m)$. We call such an i an updating index, with the former a *type 1 updating*, and the latter a *type 2 updating*. If a type 2 updating occurs under the condition $\Gamma(h_i, r_i) \leq i - m + \Gamma(h_m, r_m)$ and $\Gamma(h_{i+1}, r_{i+1}) > \Gamma(h_i, r_i)$, it is called a *turn-updating*. That is, i is the least turning point greater than m.

20.3.a Proof of Theorem 20.3.1

We show here that the pair (h_i, r_i) corresponds to a well defined nonzero rational element.

Lemma 20.3.4. *Let i be a type 2 updating index, and let (s, t) be the pair used in the combination. Then $s \neq 0$ and $t \neq 0$.*

Proof. We have

$$\pi^{i+B}|s(h_i - ar_i) + t(h_m - ar_m)\pi^{i-m}.$$

If $s = 0$, then

$$\pi^{i+B}|t(h_m - ar_m)\pi^{i-m},$$

and so

$$\pi^{m+B}|t(h_m - ar_m).$$

This is impossible because m is a turning-point. Similarly, if $t = 0$, then $\pi^{i+B}|(h_i - ar_i)$. This is also impossible because i is a type 2 updating index. □

Lemma 20.3.5. *Let m be a turning point. For any index $i \geq (m+1)$ before the next turning point, (h_m, r_m) and (h_i, r_i) are R-linearly independent. At any updating index i, $(h_{i+1}, r_{i+1}) \neq (0, 0)$.*

Proof. The proof is by induction. Note that at the initial stage, we have $(h_m, r_m) = (0, 1)$ and

$$(h_{m+1}, r_{m+1}) = (1 + a_0\pi + \cdots + a_{m+B-2}\pi^{m+B-1}, 1 + \pi^{m+B}).$$

It follows that $(h_m, r_m), (h_{m+1}, r_{m+1})$ are R-linearly independent.

Suppose $(h_m, r_m), (h_i, r_i)$ are R-linearly independent and i is an updating index. If i is a type 1 updating index, we have $(h_{i+1}, r_{i+1}) = s(h_i, r_i), s \neq 0$. So $(h_m, r_m), (h_{i+1}, r_{i+1})$ are still R-linearly independent. If i is a type 2 updating index, there are $s \neq 0, t \neq 0$ such that

$$(h_{i+1}, r_{i+1}) = s(h_i, r_i) + t\pi^{i-m}(h_m, r_m).$$

Suppose there are x, y such that $x(h_{i+1}, r_{i+1}) + y(h_m, r_m) = (0, 0)$. Then

$$xs(h_i, r_i) + (xt\pi^{i-m} + y)(h_m, r_m) = (0, 0).$$

This implies that $xs = 0$ and $xt\pi^{i-m} + y = 0$. Since $s \neq 0$, it follows that $x = 0$, so $y = 0$. This shows that $(h_m, r_m), (h_{i+1}, r_{i+1})$ are R-linearly independent. In particular, $(h_{i+1}, r_{i+1}) \neq (0, 0)$. A similar argument shows that if i is a type 2 updating index, then $(h_i, r_i), (h_{i+1}, r_{i+1})$ are R-linearly independent. Since a new turning point is obtained only by a type 2 updating, it follows that if i is a turning-point, then $(h_i, r_i), (h_{i+1}, r_{i+1})$ are updated to the new $(h_m, r_m), (h_{m+1}, r_{m+1})$, and then they are still R-linearly independent. This completes the proof. □

Lemma 20.3.6. *At any updating index i, we have $\varphi_\pi(h_i) < i$.*

Proof. The proof is again by induction on i. At the initial stage, $h_0 = 0$. The next updating index is at least $i = B$, and $h_B = \cdots = h_1$. Hence $\varphi_\pi(h_i) < i$ for the second updating index i. Now we suppose the lemma is true for every updating index $k \leq i$. We prove it is true for the

next updating index. By the induction hypothesis we have $\varphi_\pi(h_i) < i$ and $\varphi_\pi(h_m) < m$. Thus $\varphi_\pi(\pi^{i-m}h_m) < i$. Since

$$h_{i+1} = sh_i \text{ or } h_{i+1} = sh_i + t\pi^{i-m}h_m,$$

it follows from Property 1 that

$$\begin{aligned} \varphi_\pi(h_{i+1}) &\leq \max(\varphi_\pi(h_i), \varphi_\pi(\pi^{i-m}h_m)) + C \\ &< i + C < i + B. \end{aligned}$$

Let j be the next updating index. Then $j \geq i + B$ and $h_{i+1} = h_{i+2} = \cdots = h_j$. Therefore

$$\varphi_\pi(h_j) = \varphi_\pi(h_{i+1}) < i + B \leq j. \qquad \square$$

Now we can complete the proof of Theorem 20.3.1. The result is true initially. Since it is true for a type 1 updating, we only need to consider type 2 updatings with $j = i + 1$, i an updating index. We have

$$(h_{i+1}, r_{i+1}) = s(h_i, r_i) + t\pi^{i-m}(h_m, r_m).$$

Suppose $r_{i+1} = 0$. Then

$$h_{i+1} = h_{i+1} - r_{i+1}a \equiv 0 \pmod{\pi^{i+B}}.$$

If $h_{i+1} \neq 0$, then $\varphi_\pi(h_{i+1}) \geq i + B$. By Lemma 20.3.6, we have $\varphi_\pi(h_i) < i$ and,

$$\varphi_\pi(\pi^{i-m}(h_m, r_m)) < (i - m) + m = i.$$

It follows that

$$\varphi_\pi(h_{i+1}) < i + C < i + B.$$

This contradiction shows that $h_{i+1} = 0$. Therefore $(h_{i+1}, r_{i+1}) = (0, 0)$. This contradicts Lemma 20.3.5 and proves Theorem 20.3.1.

20.3.b Proof of Theorem 20.3.2

The proof of Theorem 20.3.2 uses a series of lemmas that bound the φ values of the various quantities involved. We determine a number of iterations that guarantees convergence. We start with a lower bound on this number in terms of the sizes of the approximations. We then show that the sizes of the approximations grow slowly enough that convergence is guaranteed.

Lemma 20.3.7. *Suppose $a = u/q$ with $\Gamma(u, q)$ minimal for such quotients. If $\mu(i) > \Gamma(u, q)+b+c$, then $h_i/r_i = u/q$.*

Proof. We have

$$\frac{h_i}{r_i} - \frac{u}{q} = \frac{x\pi^i}{qr_i}$$

for some $x \in R$. If $x \neq 0$, then by Property 1 we have $\Gamma(x\pi^i, qr_i) \geq i$. On the other hand, also by Property 1 we have

$$\Gamma(x\pi^i, qr_i) = \Gamma(h_i q - r_i u, r_i q) \leq \Gamma(r, q) + \Gamma(h_i, r_i) + b + c.$$

Thus $i \leq \Gamma(u, q) + \Gamma(h_i, r_i) + b + c$, which is a contradiction. Hence $x = 0$ and $h_i/r_i = u/q = a$. □

Thus if we show that $\Gamma(h_i, r_i)$ grows more slowly than i, then we can show the algorithm converges. Let m and m_1 be consecutive turning points. Let

$$\beta_{m_1} = \Gamma(h_{m_1}, r_{m_1}) - \Gamma(h_{m+1}, r_{m+1})$$

and $\beta_0 = 0$. Let k_m be the number of turning points less than m. Let $f_1 = \max\{\varphi_\pi(a) : a \in S\} \cup \{\varphi_\pi(1)\}$ here and in what follows.

Lemma 20.3.8. *At any turning point m*

$$\Gamma(h_{m+1}, r_{m+1}) \leq (m + B + c\lceil \log(B) \rceil + b + 2f_1) + Ck_m + \sum_{j \leq m} \beta_j - \Gamma(h_m, r_m),$$

and

$$Ck_m + \sum_{j \leq m} \beta_j \leq \frac{Cm}{B}. \tag{20.1}$$

Proof. The proof is by induction. For the base case, $m = 0$, we have

$$k_0 = 0, \quad \beta_0 = 0, \quad \Gamma(h_1, r_1) = B - 1 + c\lceil \log(B) \rceil + b + f_1, \text{ and } \Gamma(h_0, r_0) = \varphi_\pi(1) \leq f_1.$$

Thus the lemma is true at the first turning point.

Suppose the lemma is true at a turning point m and m_1 is the next turning point. Let $w + 1$ be the total number of updatings occurring up to m_1. Then we have

$$m_1 = m + u_0 + u_1 + \cdots + u_w,$$

with $u_i \geq B$ the difference between the i-th and $(i+1)$-st updatings. Since m_1 is a turning point, there exist $s, t \in P$ such that

$$(h_{m_1+1}, r_{m_1+1}) = s(h_{m_1}, r_{m_1}) + t\pi^{m_1 - m}(h_m, r_m).$$

By induction and the fact that $-\Gamma(h_{m+1}, r_{m+1}) = \beta_{m_1} - \Gamma(h_{m_1}, r_{m_1})$, the quantity $\Gamma = \Gamma(h_{m_1+1}, r_{m_1+1})$ satisfies

$$\begin{aligned}
\Gamma &\leq (m_1 - m) + C + \Gamma(h_m, r_m) \\
&\leq (m_1 - m) + C + (m + B + c\lceil \log(B) \rceil + b + 2f_1) + Ck_m + \sum_{j \leq m} \beta_j - \Gamma(h_{m+1}, r_{m+1}) \\
&= (m_1 + B + c\lceil \log(B) \rceil + b + 2f_1) + C(k_m + 1) + \sum_{j \leq m} \beta_j + \beta_{m_1} - \Gamma(h_{m_1}, r_{m_1}) \\
&= (m_1 + B + c\lceil \log(B) \rceil + b + 2f_1) + Ck_{m_1} + \sum_{j \leq m_1} \beta_j - \Gamma(h_{m_1}, r_{m_1}).
\end{aligned}$$

It remains to show the second inequality. It is true at the initial turning point. Assume that at a turning point m,

$$BCk_m + B\left(\sum_{j \leq m} \beta_j\right) \leq Cm \ .$$

We have $\beta_{m_1} = \Gamma(h_{m_1}, r_{m_1}) - \Gamma(h_{m+1}, r_{m+1}) \leq Cw$ and

$$\begin{aligned}
Cm_1 &= Cm + C(u_0 + u_1 + \cdots + u_w) \\
&\geq BCk_m + B\left(\sum_{j \leq m} \beta_j\right) + C(u_0 + u_1 + \cdots + u_w) \\
&\geq BCk_m + B\left(\sum_{j \leq m} \beta_j\right) + BC(w + 1) \\
&\geq BCk_m + B\left(\sum_{j \leq m} \beta_j\right) + BC + B\beta_{m_1} \\
&= BCk_{m_1} + B\left(\sum_{j \leq m_1} \beta_j\right).
\end{aligned}$$

Dividing by B gives equation (20.1): $\quad Ck_{m_1} + \sum_{j \leq m_1} \beta_j \leq \dfrac{Cm_1}{B}$. □

Let φ_m be the smallest $\Gamma(h, r)$ with $h - ar \equiv 0 \pmod{\pi^m}$.

Lemma 20.3.9. *If m is a turning point, then $\varphi_{m+1} + b + c \geq \mu(m)$.*

Proof. Let $h - ar \equiv 0 \pmod{\pi^{m+1}}$ and $\varphi_{m+1} = \Gamma(h, r)$. Then $(h, r) \neq (h_m, r_m)$. We have

$$\frac{h}{r} - \frac{h_m}{r_m} = \frac{hr_m - rh_m}{rr_m} = \frac{x\pi^m}{rr_m}$$

for some $x \neq 0 \in R$. Therefore $\Gamma(hr_m - rh_m, rr_m) \geq m$. On the other hand, we have

$$\Gamma(hr_m - rh_m, rr_m) \leq \Gamma(h, r) + \Gamma(h_m, r_m) + b + c.$$

Consequently, $\varphi_{m+1} + b + c = \Gamma(h, r) + b + c \geq \mu(m)$. □

Proof of Theorem 20.3.2. By Lemma 20.3.7 it suffices to show that $\mu(i) > b + c + \varphi$. Let m be the last turning point before i, let $t = i - m - 1$, and let w be the number of updatings between m and i. Thus $w \leq t/B$. Then

$$\begin{aligned}
b + c + \varphi + \Gamma(h_i, r_i) &\leq b + c + \varphi + \Gamma(h_{m+1}, r_{m+1}) + Cw \\
&\leq b + c + \varphi + m + B + c\lceil \log(B) \rceil + b + 2f_1 + \frac{Cm}{B} - \Gamma(h_m, r_m) + Cw \\
&\leq 2b + c + \varphi + B + c\lceil \log(B) \rceil + 2f_1 + \frac{Cm}{B} + \varphi_{m+1} + b + c + Cw \\
&\leq 3b + 2c + B + c\lceil \log(B) \rceil + 2f_1 + 2\varphi + \frac{Cm}{B} + \frac{Ct}{B} \\
&= 3b + 2c + B + c\lceil \log(B) \rceil + 2f_1 + 2\varphi + \frac{C}{B}(i - 1),
\end{aligned}$$

where the second line follows from Lemma 20.3.8 and the third line follows from Lemma 20.3.9. It follows that $b + c + \varphi < \mu(i)$ if

$$3b + 2c + B + c\lceil \log(B) \rceil + 2f_1 + 2\varphi \leq \frac{B - C}{B}(i - 1).$$

This is equivalent to the hypotheses on i in the statement of the theorem. □

20.3.c Complexity

In this subsection we analyze the computational complexity of the algorithm. At each updating index i we have

$$\Gamma(h_{i+1}, r_{i+1}) \leq \max\{\Gamma(h_i, r_i), \Gamma(h_m, r_m)\} + C$$

by Property 2, where m is the most recent turning point. Furthermore, at most one out of every B consecutive indices can be an updating index. If i is a non-updating index, then $\Gamma(h_{i+1}, r_{i+1}) = \Gamma(h_i, r_i)$. Therefore

$$\Gamma(h_i, r_i) \leq C\lceil i/B \rceil,$$

so $\Gamma(h_i, r_i) \le i$ if $i \ge C(B-1)/(B-C)$. Also, note that we do not have to save all the intermediate values (h_i, r_i), just the current value and the value for the most recent turning point.

Suppose we have a bound $\sigma(m)$ on the time required to add two elements $a, b \in R$ with $\varphi_\pi(a), \varphi_\pi(b) \le m$. Then we have the following.

Corollary 20.3.10. *The worst case time complexity of the Rational Approximation Algorithm is in*

$$O\left(\sum_{m=1}^{\varphi} \sigma(m)\right).$$

The worst case space complexity is in $O(\varphi \log(|S|))$. □

20.4 Rational approximation in function fields

In this section we use Xu's algorithm to solve the rational approximation problem for AFSRs based on function fields. (In Section 20.8 we take a different approach to cryptanalysis of such sequences by transforming them into longer sequences defined over smaller fields, with linear span closely related to the π-adic span.)

As in Section 15.2, let $\mathbf{a}$ be a sequence generated by an AFSR based on a global function field. That is, $F = \mathbb{F}_r$ is a finite field with $r = p^h$, p prime. Assume that $R = F[x_1, \cdots, x_n]/I$ has transcendence degree 1 over $\mathbb{F}_r$, where I is an ideal. Let $\pi \in F[x_1, \cdots, x_n]$ and assume that, as a vector space over F, the ring $K = R/(\pi)$ has finite dimension $e < \infty$. Let $S \subset R$ be a complete set of representatives for K such that hypotheses H1 and H2 of Section 8.3.c hold:

Hypothesis H1: S is closed under addition and contains F.

Hypothesis H2: $R \subset S[\pi]$, in other words, for every $v \in R$ there exist $v_0, v_1, \cdots, v_\ell \in S$ so that

$$v = v_0 + v_1\pi + \cdots + v_\ell \pi^\ell. \tag{20.2}$$

If $v \ne 0$ then define $\varphi_{R,\pi}(v) = \ell$ where ℓ is the smallest integer such that equation (20.2) holds and set $\varphi_{R,\pi}(0) = -\infty$. This is sometimes called the *π-degree of v*. Let

$$L = \max\{\varphi_{R,\pi}(uv) : u, v \in S\}.$$

Theorem 20.4.1. *There is a rational approximation algorithm for R, π, S that is convergent for a sequence $\mathbf{a}$, using only the first $10L^2 + 7L + 1 + (4L+2)\varphi(\mathbf{a})$ symbols of $\mathbf{a}$.*

The proof follows immediately from the next two propositions.

Proposition 20.4.2. *Property 1 holds for $\varphi_{R,\pi}$ with $b = L$ and $c = 0$.*

Proof. Conditions (1), (3), and (4) are immediate from the definition. To see that condition (2) holds, let

$$x = \sum_{i=0}^{k} x_i \pi^i \quad \text{and} \quad y = \sum_{i=0}^{\ell} y_i \pi^i$$

with $x_i, y_i \in S$. For each i, j we have

$$x_i y_j = \sum_{t=0}^{L} z_{i,j,t} \pi^t$$

for some $z_{i,j,t} \in S$. Then

$$xy = \sum_{n=0}^{k+\ell+L} \left(\sum_{i+j+t=n} z_{i,j,t} \right) \pi^n.$$

Thus, $\varphi_{R,\pi}(xy) \leq \varphi_{R,\pi}(x) + \varphi_{R,\pi}(y) + L$. □

Let

$$P_{R,\pi} = \{s \in R : \varphi_{R,\pi}(s) \leq L, s \neq 0\}.$$

It follows that $C = 2L$.

Proposition 20.4.3. *Property 2 holds for $P_{R,\pi}$ with $B = 2L + 1 > C$.*

Proof. Conditions (1) and (2) are immediate from the definition.

Fix $h_1, h_2 \in R$. The set of all pairs (s, t) with $s, t \in V = P_{R,\pi} \cup \{0\}$ has cardinality $|S|^{2L+2}$. The set of such pairs is a vector space over F. The quotient $R/(\pi^B)$ has cardinality $|S|^{2L+1}$. Thus the homomorphism

$$(s, t) \mapsto sh_1 + th_2 \pmod{\pi^B}$$

is not one to one. That is, there is a pair $(s, t) \neq (0, 0)$ so that π^B divides $sh_1 + th_2$. This proves condition (3). □

20.5 Rational approximation in ramified extensions

Let Q be an integral domain, $\tau \in Q$, S a complete set of residues modulo τ, and suppose we have an index function $\varphi_{Q,\tau}$ and interpolation set $P_{Q,\tau}$ with respect to τ. Let b, c, d, B, and $C = b+c+d$ be the constants in Properties 1 and 2 with respect to $\varphi_{Q,\tau}$ and $P_{Q,\tau}$.

Let r be a positive integer and $\epsilon = \pm 1$. Assume that the polynomial $X^r - \epsilon\tau$ is irreducible over Q, and π is a root of this polynomial. In this section we consider the case when

$$R = Q[\pi] = \left\{ \sum_{i=0}^{r-1} a_i \pi^i : a_i \in Q \right\}.$$

We have $R/(\pi) = Q/(\tau)$, so S is a complete set of representatives for R modulo π as well.

For any $x = \sum_{i=0}^{r-1} a_i \pi^i, a_i \in Q$, we define

$$\varphi_{R,\pi}(x) = \max\{r\varphi_{Q,\tau}(a_i) + i : 0 \le i \le r-1\}.$$

This is well defined because, by the irreducibility of $X^r - \epsilon\tau$, this representation of x is unique. Let $c' = cr$ and $b' = cr\lceil\log(r)\rceil + br$. Then for any $x, y \in R$ and nonzero integer k,

1. $\varphi_{R,\pi}(x) \ge 0$ if $x \ne 0$;
2. $\varphi_{R,\pi}(xy) \le \varphi_{R,\pi}(x) + \varphi_{R,\pi}(y) + b'$.
3. $\varphi_{R,\pi}(x \pm y) \le \max\{\varphi_{R,\pi}(x), \varphi_{R,\pi}(y)\} + c'$;
4. $\varphi_{R,\pi}(\pi^k x) = k + \varphi_{R,\pi}(x)$;

Let $e = \max\{\varphi_{Q,\tau}(x) : x \in S\}$. Choose $m \in \mathbb{Z}$ large enough that

$$r(m - c\lceil\log(m)\rceil) \ge b' + 2c' + r(b+e) + 1 - 2r.$$

Let $\ell = m + c\lceil\log(m)\rceil + e + b$,

$$P_0 = \{x \in R : \varphi_{R,\pi}(x) \le r\ell\},$$

and $P = \{x - z : x \ne z \in P_0\}$. Thus $d' = \max\{\phi_{R,\pi}(x) : x \in P\} \le r\ell + c'$. If $x = \sum_{i=0}^{r-1} x_i \pi^i$ with $x_i \in Q$, then $x \in P_0$ if and only if $\varphi_{Q,\tau}(x_i) \le \ell$ for each i. Thus

$$|P_0| = |\{y \in Q : \varphi_{Q,\tau}(y) \le \ell\}|^r.$$

Now suppose that $y = \sum_{i=0}^{m} y_i \tau^i \in Q$ with $y_i \in S$. Then $\varphi_{Q,\tau}(y) \le m + e + b + c\lceil\log(m)\rceil = \ell$, so

$$|\{y \in Q : \varphi_{Q,\tau}(y) \le \ell\}| \ge |S|^{m+1}.$$

It follows that

$$|P_0 \times P_0| \ge |S|^{2rm+r}.$$

For any positive B we have $|R/(\pi^B)| = |S|^B$. Let

$$B' = b' + 2c' + r\ell < 2rm + 2r.$$

Then $|P_0 \times P_0| > |R/(\pi^{B'})|$, so for any $h_1, h_2 \in R$ the map from $P_0 \times P_0$ to $R/(\pi^{B'})$ that takes (s,t) to $sh_1 + sh_2$ is not one to one. It follows that there exists $(s,t) \in P \times P$ such that $\pi^{B'}$ divides $sh_1 + th_2$. This proves the third condition of Property 2.

It follows that there is a Rational Approximation Algorithm for R, π. Suppose any element x of Q can be represented using at most $p\varphi_{Q,\tau}(x)$ bits for some p. Then any element m of R can

be represented using at most $p\varphi_{R,\pi}(x)$ bits. Thus, by the discussion following Proposition 20.1.1, the number of symbols of the output sequence of an AFSR over R, π needed to synthesize an equivalent AFSR is at most linear in the size of the smallest AFSR that generates the sequence.

While the algorithm is guaranteed to find a rational representation for the given sequence, its Γ value may not be minimal. In fact it may be that multiplying both elements in a pair by the same element (thus leaving the corresponding rational element unchanged) decreases Γ. For example, suppose $\tau = 3$ and $r = 2$ so $\pi^2 = 3$. Let $x = 27 - 14\pi$, $y = 28 - 15\pi$, and $z = 1 + \pi$. Then $\varphi_{R,\pi}(x) = \varphi_{R,\pi}(y) = 6$. However, $zx = -15+13\pi$ and $zy = -17+13\pi$ so $\varphi_{R,\pi}(zx) = \varphi_{R,\pi}(zy) = 5$.

20.6 Rational approximation in quadratic extensions

In this section we consider the case of a quadratic extension of a ring Q. Again let Q be an integral domain, $\tau \in Q$, S a complete set of residues modulo τ with $N = |S|$, and suppose we have an index function $\varphi_{Q,\tau}$ and interpolation set $P_{Q,\tau}$ with respect to τ. Let b, c, B, and C be the constants in Properties 1 and 2 with respect to $\varphi_{Q,\tau}$ and $P_{Q,\tau}$.

Let $m, g \in Q$ with $m^a = \tau$ for some $a \geq 1$. Let π be a root of the polynomial $X^2 - 2gmX + m^a$, and assume $\pi \notin Q$. In this section we consider whether there is a rational approximation algorithm for $R = Q[\pi]$. If we let $\Delta = m^a - g^2m^2$, then $\pi = gm + \sqrt{-\Delta}$ and we also have $R = Q[\sqrt{-\Delta}]$. The norm from the field of fractions of R to the field of fractions of Q is given by $\mathsf{N}(u + v\sqrt{-\Delta}) = u^2 + \Delta v^2$. In particular, $\mathsf{N}(\pi) = \tau$. Let

$$\varphi_{R,\pi}(x) = \varphi_{Q,\tau}(\mathsf{N}(x)).$$

It follows immediately that

$$\varphi_{R,\pi}(xy) \leq \varphi_{R,\pi}(x) + \varphi_{R,\pi}(y) + b,$$

and

$$\varphi_{R,\pi}(\pi^k x) = k + \varphi_{R,\pi}(x).$$

However, the additivity condition for an index function does not hold in general. Therefore, we assume at this point that it does hold. That is, we assume that there is a c' such that for any $x_0, x_1, y_0, y_1 \in Q$

$$\varphi_{Q,\tau}((x_0 + y_0)^2 + \Delta(x_1 + y_1)^2) \leq \max\{\varphi_{Q,\tau}(x_0^2 + \Delta x_1^2), \varphi_{Q,\tau}(y_0^2 + \Delta y_1^2)\} + c'. \tag{20.3}$$

At the end of this section we give examples of rings Q for which this condition holds. For now we show that if it holds, then the remaining conditions – the existence of a set $P_{R,\pi}$ satisfying Property 2 – for the existence of a rational approximation algorithm hold.

First we consider the case when $a \geq 2$. First we need a lemma.

Lemma 20.6.1. *For any $k \geq 0$, $\pi^{(k+1)a+1}$ divides $\tau^{a+k(a-1)}$.*

Proof. Let $u, v \in Q$. Then $(2gm - \pi)(mu + \pi v) = m(m(2gu - m^{a-2}v) - u\pi) = m(my - \pi u)$, for some $y \in Q$.

We iterate this a times: For any $u, v \in Q$ there are $y, z \in Q$ such that $(2gm - \pi)^a(mu + \pi v) = m^a(my + \pi z) = \pi(2gm - \pi)(my + \pi z)$. Thus $(2gm - \pi)^{a-1}(mu + \pi v) = \pi(my + \pi z)$. It follows that

$$(2gm - \pi)^{a+k(a-1)} = (2gm - \pi)^{(k+1)(a-1)}(2gm - \pi) = \pi^{k+1}(my + \pi z),$$

for some $y, z \in Q$. Now we have

$$\begin{aligned} \tau^{a+k(a-1)} &= \pm\pi^{a+k(a-1)}(\pi - 2gm)^{a+k(a-1)} \\ &= \pm\pi^{a+k(a-1)}\pi^{k+1}(mf + \pi h) \\ &= \pm\pi^{(k+1)a+1}(mf + \pi h). \qquad \square \end{aligned}$$

Let $e = \max\{\varphi_{Q,\tau}(x) : x \in S\}$ and let $w = \varphi_{Q,\tau}(\Delta) + b' + c' + 3c + 2b + 2e - 1$. Choose $t \in \mathbb{Z}$ large enough that

$$\left\lceil \frac{2t - (a-1)(2c\lceil \log_2(t+1)\rceil + w)}{a} \right\rceil + 1 - a \geq 0.$$

Let

$$k = \left\lceil \frac{2t + 2c\lceil \log_2(t+1)\rceil + w}{a} \right\rceil - 1.$$

Then

$$2t + 1 \geq k(a+1) - k \tag{20.4}$$

and

$$(k+1)a + 1 > 2t + 2c\lceil \log_2(t+1)\rceil + w. \tag{20.5}$$

Take $B' = (k+1)a + 1$,

$$P_0 = \{x_0 + x_1\sqrt{-\Delta} : x_i \in Q, \varphi_{Q,\tau}(x_i) \leq h + c\lceil \log_2(t+1)\rceil + e\},$$

and $P_{R,\pi} = \{x - y : x \neq y \in P_0\}$. As in Section 20.5, if

$$s_i = \sum_{j=0}^{t} s_{i,j}\tau^j$$

with $s_{i,j} \in S$, then $s_0 + s_1\sqrt{-\Delta} \in P_0$. Thus

$$|P_0| \geq |S|^{2t+2}. \tag{20.6}$$

It follows from equation (20.5) that $B' > b' + c' + d'$.

Now let $x+\pi y \in R$, with $x, y \in Q$. We can write $x = x_0+x_1\tau^{a+1+k(a-1)}$ and $y = y_0+y_1\tau^{a+k(a-1)}$. It follows from Lemma 20.6.1 that $\pi^{B'} = \pi^{(k+1)a+1}$ divides both $\tau^{a+k(a-1)}$ and $\pi\tau^{a+k(a-1)-1}$. Thus

$$|R/\pi^{B'}| \le |Q/\tau^{a+k(a-1)}| \cdot |Q/\tau^{a+k(a-1)-1}| = |S|^{2a+2k(a-1)-1}.$$

It then follows from equations (20.4) and (20.6) that $|P_0^2| > |R/\pi^{B'}|$. Thus we can use earlier arguments to show that P satisfies Property 2.

Now consider the case when $a = 1$. Then $\tau = \pi(2\tau - \pi) = \pi^2(4\tau - 2\pi - 1)$ so π^2 divides τ. In this case we can choose $t \in \mathbb{Z}$ so that $t - c\lceil\log_2(t+1)\rceil \ge e + b + c + c' + \varphi_Q(\Delta)/2 - 1$. Then we can take $B' = 2t + c\lceil\log_2(t+1)\rceil + 2e + 2b + 2c + 2c' + \varphi(\Delta)$, $P_0 = \{x_0 + x_1\sqrt{-\Delta} : \varphi(x_i) \le t + c\lceil\log_2(t+1)\rceil + e\}$, and $P_{R,\pi} = \{x - y : x \ne y \in P_0\}$. We omit the details.

We have proved the following theorem.

Theorem 20.6.2. *If equation (20.3) holds, then there is a rational approximation algorithm for R with respect to π.*

Remarks:
(1) We have shown the existence of constants b, c, B, but have not attempted to optimize them. We know the algorithm converges after a linear number of iterations. In many cases the convergence may be more rapid than indicated by the results here.
(2) Rational approximation algorithms exist for extensions by roots of other quadratic polynomials. For instance, $\pi = 3 + \sqrt{-3}$ is a root of the equation $X^2 - 6X + 12 = 0$. Let $N = 12$. Then $\pi^4 = \pi^7(\pi - 6)$. In this case we can choose $b' = 1$. Since this is an imaginary quadratic extension, the additivity condition on the index function holds, in this case with $c' = 1$. We can also take $B' = 7$ to establish a rational approximation algorithm. The task of completely characterizing those quadratic extensions for which there is a rational approximation algorithm remains.

20.6.a Imaginary quadratic extensions of $\mathbb{Z}$

In this subsection we assume $R = \mathbb{Z}[\pi]$ is an imaginary quadratic extension of the integers, with $\pi^2 - 2gm\pi + N = 0$ and $N = m^a$.

In this case Δ is a positive integer. We carry out the above construction with $Q = \mathbb{Z}$, $\tau = N$, and index function and interpolation set as in Section 20.2. It suffices to show equation (20.3) holds. Let $x = x_0 + x_1\sqrt{-\Delta}$ and $y = y_0 + y_1\sqrt{-\Delta}$ with $x_0, x_1, y_0, y_1 \in \mathbb{Z}$. We then have $\mathsf{N}(x+y) = (x_0+y_0)^2 + \Delta(x_1+y_1)^2$. Notice that $(c+d)^2 \le 2(c^2+d^2)$ for any real numbers c and d. This implies that

$$\begin{aligned} \mathsf{N}(x+y) &\le 2(x_1^2+y_1^2) + 2\Delta(x_2^2+y_2^2) \\ &= 2\mathsf{N}(x) + 2\mathsf{N}(y). \end{aligned}$$

Let $w_1 = \varphi_{R,\pi}(x) = \varphi_{\mathbb{Z},N}(\mathsf{N}(x))$ and $w_2 = \varphi_{R,\pi}(y) = \varphi_{\mathbb{Z},N}(\mathsf{N}(y))$. Then we have

$$\mathsf{N}(x) \leq N^{w_1+1} - 1,$$

$$\mathsf{N}(y) \leq N^{w_2+1} - 1,$$

and

$$\mathsf{N}(x+y) \leq 4(N^{max(w_1,w_2)+1} - 1).$$

Since $N \geq 2$, we have $\varphi_{R,\pi}(x+y) = \varphi_{\mathbb{Z},N}(\mathsf{N}(x+y)) \leq \max\{w_1, w_2\} + 2$. We have proven the following corollary.

Corollary 20.6.3. *If $R = \mathbb{Z}[\pi]$ is an imaginary quadratic extension of the integers, with $\pi^2 - 2gm\pi + N = 0$ and $N = m^a$, then R has a rational approximation algorithm with respect to π.*

Any element $m = x_0 + x_1\sqrt{-\Delta}$ can be represented using $\varphi_{\mathbb{Z},N}(x_0) + \varphi_{\mathbb{Z},N}(x_1) \leq \varphi_{\mathbb{Z},N}(x_0^2 + \Delta x_1^2) = \varphi_{R,\pi}(m)$ elements of $\{0, 1, \cdots, N-1\}$. Thus, by the discussion following Proposition 20.1.1, the number of symbols of the output sequence of an AFSR over R, π needed to synthesize an equivalent AFSR is at most linear in the size of the smallest AFSR that generates the sequence.

20.6.b Quadratic extensions of $\mathbb{Z}[\sqrt{N}]$

In this subsection we let N be a positive integer which is not a perfect square, let $\tau^2 = N$, and let $Q = \mathbb{Z}[\tau]$. Let $\pi^2 - 2gm\pi + \tau = 0$ with $\tau = m^a$ and $g, m \in Q$, and let $R = Q[\pi]$. Thus $Q = \mathbb{Z} + \tau\mathbb{Z}$ and $R = Q + \pi Q$. Let $\Delta = m^a - g^2m^2 = \Delta_0 + \Delta_1\tau$ with $\Delta_0 > 0$, $\Delta_1 \neq 0$ in $\mathbb{Z}$, and $\Delta_0^2 > N\Delta_1^2$. That is, we assume the norm from the fraction field of Q to the rational numbers of Δ is positive. In this section we show that there is a rational approximation algorithm in the setting of R and π.

By Section 20.5 with $r = 2$ we have the ingredients for a rational approximation algorithm with respect to Q, τ, and S: a function

$$\varphi_{Q,\tau}(x_0 + x_1\tau) = \max\{2\varphi_{\mathbb{Z},N}(x_0), 2\varphi_{\mathbb{Z},N}(x_1) + 1\} = \max\{2\lfloor \log_N |x_0| \rfloor, 2\lfloor \log_N |x_1| \rfloor + 1\},$$

a set $P_{Q,\tau}$, and constants b, c, and B satisfying Properties 1 and 2. The proof of the following lemma is straightforward.

Lemma 20.6.4. *If $u \in Q$, then $2\varphi_{Q,\tau}(u) - 2 \leq \varphi_{Q,\tau}(u^2)$.* □

Lemma 20.6.5. *Let $\Delta = \Delta_0 + \Delta_1\tau$ with $\Delta_0, \Delta_1 \in \mathbb{Z}$, $\Delta_0 > 0$, $\Delta_1 \neq 0$, and $\Delta_0^2 > N\Delta_1^2$. If $u, v \in Q$, then $2\varphi_{Q,\tau}(u) \leq \varphi_{Q,\tau}(u^2 + \Delta v^2) + 2$ and $2\varphi_{Q,\tau}(v) \leq \varphi_{Q,\tau}(u^2 + \Delta v^2) + 2$.*

Proof. Let $u = u_0 + \tau u_1$ and $v = v_0 + \tau v_1$ with $u_0, u_1, v_0, v_1 \in \mathbb{Z}$. Then

$$\begin{aligned} u^2 + \Delta v^2 &= u_0^2 + N u_1^2 + \Delta_0 v_0^2 + \Delta_0 N v_1^2 + 2\Delta_1 N v_0 v_1 \\ &\quad + (2u_0 u_1 + 2\Delta_0 v_0 v_1 + \Delta_1 v_0^2 + \Delta_1 N v_1^2)\tau. \end{aligned} \tag{20.7}$$

We have

$$\begin{aligned} \Delta_0 v_0^2 + \Delta_0 N v_1^2 + 2\Delta_1 N v_0 v_1 &= \Delta_0 (v_0 + \sqrt{N} v_1)^2 + 2 v_0 v_1 \sqrt{N}(\Delta_1 \sqrt{N} - \Delta_0) \qquad (20.8) \\ &= \Delta_0 (v_0 - \sqrt{N} v_1)^2 + 2 v_0 v_1 \sqrt{N}(\Delta_1 \sqrt{N} + \Delta_0). \qquad (20.9) \end{aligned}$$

Suppose that $\Delta_1\sqrt{N} - \Delta_0$ and $\Delta_1\sqrt{N} + \Delta_0$ have the same sign. Then $\Delta_1^2 N - \Delta_0^2 > 0$, which is false by hypothesis. Thus one of is positive and one is negative. Whatever the sign of $v_0 v_1$ is, either expression (20.8) or expression (20.9) is nonnegative. It follows from equation (20.7) that

$$\begin{aligned} \varphi_{Q,\tau}(u^2 + \Delta v^2) &\geq 2\varphi_{\mathbb{Z},N}(u_0^2 + N u_1^2) \\ &\geq \max\{4\varphi_{\mathbb{Z},N}(u_0) - 2, 4\varphi_{\mathbb{Z},N}(u_1)\} \\ &= 2\varphi_{Q,\tau}(u) - 2. \end{aligned}$$

It also follows that

$$\begin{aligned} \varphi_{Q,\tau}(u^2 + \Delta v^2) &\geq 2\varphi_{\mathbb{Z},N}(\Delta_0 v_0^2 + \Delta_0 N v_1^2 + 2\Delta_1 N v_0 v_1) \\ &\geq 2 \max\left\{ \left\lfloor \log_N |(v_0 \pm \sqrt{N} v_1)^2| \right\rfloor, \left\lfloor \log_N |2\sqrt{N} v_0 v_1| \right\rfloor \right\}. \end{aligned}$$

If $2N^{1/2} v_0 \geq v_1 \geq v_0/(2N^{3/2})$, then $\left\lfloor \log_N |2\sqrt{N} v_0 v_1| \right\rfloor \geq \max\{2\lfloor \log_N |v_0| \rfloor - 1, 2\lfloor \log_N |v_1| \rfloor\}$. If $2N^{1/2} v_0 \leq v_1$ or $v_1 \leq v_0/(2N^{3/2})$, then $\left\lfloor \log_N |(v_0 \pm \sqrt{N} v_1)^2| \right\rfloor \geq \max\{2\lfloor \log_N |v_0| \rfloor - 1, 2\lfloor \log_N |v_1| \rfloor\}$. In every case it follows that

$$\begin{aligned} \varphi_{Q,\tau}(u^2 + \Delta v^2) &\geq 2(\max\{2\varphi_{\mathbb{Z},N}(v_0), 2\varphi_{\mathbb{Z},N}(v_1) + 1\} - 1) \\ &= 2\varphi_{Q,\tau}(v) - 2. \end{aligned}$$

□

Let $x = x_0 + \pi x_1$ and $y = y_0 + \pi y_1$. We have

$$\begin{aligned} \varphi_{R,\pi}(x + y) &= \varphi_{Q,\tau}((x_0 + y_0)^2 + \Delta(x_1 + y_1)^2) \\ &\leq \max\{\varphi_{Q,\tau}((x_0 + y_0)^2), \varphi_{Q,\tau}((x_1 + y_1)^2) + \varphi_{Q,\tau}(\Delta) + b\} + c \\ &\leq \max\{2\varphi_{Q,\tau}(x_0 + y_0), 2\varphi_{Q,\tau}(x_1 + y_1) + \varphi_{Q,\tau}(\Delta) + b\} + c + 4 \\ &\leq \max\{2\varphi_{Q,\tau}(x_0), 2\varphi_{Q,\tau}(y_0), 2\varphi_{Q,\tau}(x_1) + \varphi_{Q,\tau}(\Delta) + b, \\ &\quad 2\varphi_{Q,\tau}(y_1) + \varphi_{Q,\tau}(\Delta) + b\} + 3c + 4. \end{aligned}$$

By Lemma 20.6.5, both $2\varphi_{Q,\tau}(x_0)$ and $2\varphi_{Q,\tau}(x_1)$ are bounded by $\varphi_{Q,\tau}(x_0^2+\Delta x_1^2)+2$, and similarly for y. It follows that

$$\begin{aligned}\varphi_{R,\pi}(x+y) &\leq \max\{\varphi_{Q,\tau}(x_0^2+\Delta x_1^2), \varphi_{Q,\tau}(y_0^2+\Delta y_1^2)\} + b + 3c + 6 \\ &= \max\{\varphi_{R,\pi}(x), \varphi_{R,\pi}(y)\} + b + 3c + 6.\end{aligned}$$

We have proved the following.

Corollary 20.6.6. *Let N be a positive integer which is not a perfect square, let $\tau^2 = N$, and let $Q = \mathbb{Z}[\tau]$. Let $\pi^2 - 2gm\pi + \tau = 0$ with $\tau = m^a$ and $g, m \in Q$, and let $R = Q[\pi]$. If $m^a - g^2m^2 = \Delta_0 + \Delta_1\tau$ with $\Delta_0 > 0$, $\Delta_1 \neq 0$, and $\Delta_0^2 > N\Delta_1^2$, then R has a rational approximation algorithm with respect to π.*

Any element $m = x_0 + x_1\tau + x_2\sqrt{-\Delta} + x_3\tau\sqrt{-\Delta} \in R$, with $x_i \in \mathbb{Z}$, can be represented using $\sum_{i=0}^{3} \varphi_{\mathbb{Z},N}(x_i)$ elements, plus four sign bits. We have

$$\begin{aligned}\sum_{i=0}^{3} \varphi_{\mathbb{Z},N}(x_i) &\leq 4\max\{\varphi_{\mathbb{Z},N}(x_i) : i = 0, \cdots, 3\} \\ &\leq 2\max\{\varphi_{Q,\tau}(x_0 + x_1\tau), \varphi_{Q,\tau}(x_2 + x_3\tau)\} \\ &\leq \varphi_{Q,\tau}((x_0 + x_1\tau)^2 + \Delta(x_2 + x_3\tau)^2) + 2 \\ &= \varphi_{R,\pi}(m) + 2.\end{aligned}$$

Thus, by the discussion following Proposition 20.1.1, the number of symbols of the output sequence of an AFSR over R, π needed to synthesize an equivalent AFSR is at most linear in the size of the smallest AFSR that generates the sequence.

20.7 Rational approximation by interleaving

In some cases we may be able to synthesize an AFSR over R, π, and S for a given sequence $\mathbf{a}$ by decomposing it as as a sum or interleaving of sequences over subrings. We illustrate this phenomenon with an example.

Let Q be an integral domain with $\tau \in Q$ not a unit and let S be a complete set of residues for Q modulo τ. Assume that $\cap_{i=1}^{\infty}(\tau^i) = (0)$ so that $Q \subseteq Q_\tau$. Suppose that

$$f(x) = x^k g(x) - \tau h(x)$$

is an irreducible polynomial over Q, where $g(x)$ and $h(x)$ are polynomials whose constant terms are units in Q, $g(x)$ is monic of degree $d \geq 0$, $k \geq 2$, and $\deg(h) < k + \deg(g)$.

Let $R = Q[\pi]$ with π a root of $f(x)$. Then $(\pi) \cap Q = (\tau)$. Thus there is an induced homomorphism from $Q/(\tau)$ to $R/(\pi)$ and this homomorphism is one to one. Moreover

$$R = \left\{ \sum_{i=0}^{k+d-1} c_i \pi^i : c_i \in Q \right\},$$

so every element of R differs from an element of Q by a multiple of π. Thus this homomorphism is an isomorphism. In particular S is a complete set of residues for R modulo π as well.

We want to express a sequence over S in terms of several related sequences over S so that the π-adic complexity of the original sequence is related to the τ-adic complexities of the new sequences and so that the original sequence and the new sequences can be efficiently computed from each other. Algebraically, we want to express an element of R_π in terms of elements of Q_τ.

Theorem 20.7.1. *Under the hypotheses above,*

$$R_\pi = Q_\tau[\pi] = \left\{ \sum_{i=0}^{k-1} c_i \pi^i : c_i \in Q_\tau \right\}.$$

Proof. Since their constant terms are units, the polynomial $h(x)$ has a multiplicative inverse $h'(x)$ in R_π. Thus

$$\tau = \pi^k g(\pi) h'(\pi) \in R_\pi,$$

and the leading coefficient in the π-adic expansion of τ is a unit. Let $\mathbf{a} = a_0, a_1, \cdots$ be a sequence of elements of S and

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

We claim that there are elements $b_{i,j} \in S$, $i = 0, 1, \infty$ and $j = 0, 1, \cdots, k-1$, so that

$$a = \sum_{i=0}^{\infty} \sum_{j=0}^{k-1} b_{j,i} \pi^j \pi^{ik} (g(\pi) h'(\pi))^i.$$

The $b_{j,i}$ can be found recursively as follows. First, $b_{0,0} = a_0$, $b_{1,0} = a_1, \cdots, b_{k-1,0} = a_{k-1}$. Now suppose that we have found $\{b_{j_i} \in S : 0 \le j \le k-1, i = 0, \cdots, \ell - 1\}$ so that

$$a \equiv \sum_{i=0}^{\ell-1} \sum_{j=0}^{k-1} b_{j,i} \pi^j \pi^{ik} (g(\pi) h'(\pi))^i \pmod{\pi^{\ell k}}$$

Let

$$a - \sum_{i=0}^{\ell-1}\sum_{j=0}^{k-1} b_{j,i}\pi^j\pi^{ik}(g(\pi)h'(\pi))^i = \sum_{i=0}^{\infty} c_i\pi^i$$

with $c_i \in S$. Let $b_{0,\ell} = c_0, b_{1,\ell} = c_1, \cdots, b_{k-1,\ell} = c_{k-1}$. Then

$$a \equiv \sum_{i=0}^{\ell}\sum_{j=0}^{k-1} b_{j,i}\pi^j\pi^{ik}(g(\pi)h'(\pi))^i \pmod{\pi^{(\ell+1)k}}$$

This proves the claim. It follows that

$$\begin{aligned}
a &= \sum_{i=0}^{\infty}\sum_{j=0}^{k-1} b_{j,i}\pi^j\pi^{ik}(g(\pi)h'(\pi))^i \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{k-1} b_{j,i}\tau^i\pi^j \\
&= \sum_{j=0}^{k-1}\sum_{i=0}^{\infty} b_{j,i}\tau^i\pi^j \\
&\in \left\{\sum_{i=0}^{k-1} c_i\pi^i : c_i \in Q_\tau\right\} \\
&\subseteq Q_\tau[\pi]. \qquad \square
\end{aligned}$$

Suppose we are given the first m elements of a sequence $\mathbf{a} = a_0, a_1, \cdots$ with $a_i \in S$. We want to find an approximation to the coefficients $c_j \in Q_\tau$ in the representation

$$a = \sum_{i=0}^{\infty} a_i\pi^i = \sum_{j=0}^{k-1} c_j\pi^j. \tag{20.10}$$

The procedure in the proof of Theorem 20.7.1 can be used to compute the first $\lfloor m/k \rfloor$ coefficients of the τ-adic representation of each c_j. The time complexity is m times the time it takes to compute the first m coefficients of a difference of π-adic numbers. Typically the resulting complexity is $O(m^2)$.

Now suppose a can be represented as in equation (20.10) and that the sequence $\mathbf{c}_j$ of coefficients in S in the τ-adic expansion each c_j can be generated by an AFSR over Q, τ, S. Let $c_j = f_j/q_j$ with $f_j, q_j \in Q$ and q_j a unit modulo τ. If q is a common multiple of $q_0, \cdots, q_{k-1}$ and $r_j = q/q_j$, $j = 0, \cdots, k-1$, then

$$a = \frac{f_0r_0 + f_1r_1\pi + \cdots f_{k-1}r_{k-1}\pi^{k-1}}{q}.$$

If there is a rational approximation algorithm for Q, τ, S, then we immediately obtain a rational approximation algorithm for R, π, S:

1. From the sequence $\mathbf{a}$ obtain the component sequences $\mathbf{c}_j$;
2. Apply a rational approximation algorithm over Q, τ, S to each $\mathbf{c}_j$ to obtain the approximation f'_j/q'_j;
3. Construct a rational approximation to a by combining the f'_j/q'_j into a single rational element f'/q'.

If each rational approximation f'_j/q'_j is correct modulo τ^ℓ, then the rational approximation to a will be correct modulo $\pi^{\ell k}$.

It is not clear in general how close this method comes to finding the best rational approximation to a, even if the underlying rational approximation algorithm for Q finds the best rational approximation. There are several complications. Q may not be a GCD ring, so q' may not be the unique least common multiple of the q'_j. Even if Q is a GCD ring, it may not be possible to compute the least common multiple of the q'_j efficiently. However, we can always take the product of the q'_j as a (not necessarily least) common multiple. Even if Q is a GCD ring and we can efficiently compute a least common multiple of the q'_j, this may not give us the connection element of the smallest AFSR that outputs a. For example, multiplying q' and f' by a common unit gives another AFSR that outputs a but may have a different size. The unit group of Q may be infinite, so finding the unit u that minimizes the size of the AFSR corresponding to $uf'/(uq')$ may not be feasible.

20.8 Rational function fields: π-adic vs. linear span

Suppose $R = \mathbb{F}_r[x]$ with $r = p^n$, p prime, $\pi \in R$ is irreducible, and $S \subseteq R$ is a complete set of representatives for $K = R/(\pi)$. Then K is a field and is an extension of $L = \mathbb{F}_r$ of degree $d = \deg(\pi)$. Hence K has r^d elements, so it is $\mathbb{F}_{r^d}$. Also, the cardinality of S is r^d. We would like to relate the π-adic span or complexity of sequences over K to the linear span of related sequences over L.

Throughout this section we assume we are given an AFSR A over (R, π, S) with connection element

$$q = -q_0 + q_1\pi + \cdots + q_r\pi^m, \tag{20.11}$$

with $q_i \in R$, q_0 invertible modulo π. We let m be the size of A. Let

$$U = \max\{\deg(u) : u \in S\} \quad \text{and} \quad V = \max\{\deg(q_i) : i = 0, \cdots, d\}.$$

We also let $\mathbf{a}$ be an output sequence from A with initial memory z. That is, the initial state is $(a_0, \cdots, a_{m-1}; z)$.

One standard correspondence between sequences over K and sequences over L is the following. If we choose a basis $c_1, \cdots, c_d$ for K over L, then every element of K can be treated as a d-tuple of elements of L. For an element $e \in K$, let $\sigma_i(e)$ denote the coefficient of c_i in the representation of e using this basis. Let the function σ be defined by

$$\sigma(\mathbf{a}) = \sigma_1(a_0), \sigma_2(a_0), \cdots, \sigma_d(a_0), \sigma_1(a_1), \cdots,$$

or equivalently

$$\sigma(\mathbf{a})_{id+j} = \sigma_j(a_i),$$

with $i = 0, 1, \cdots$ and $j = 1, 2, \cdots, d$. Then σ is a one to one correspondence. If $\mathbf{a}$ has period t then $\sigma(\mathbf{a})$ has period dt. However, this is not adequate for our purposes, so in this section we use a more general construction: the value of any symbol of $\sigma(\mathbf{a})$ will depend on more than one symbol of $\mathbf{a}$.

We show that there is a sequence $\mathbf{b}$ over L such that $\mathbf{a}$ can be transformed into a sequence $\mathbf{b}$ (and vice versa) by a finite state "filter" that depends only on R, π, S, and the q_i. The linear span of $\mathbf{a}$ is at most rd plus the maximum of $\deg(z)$ and a constant that depends only on R, π, S, and the degrees of the q_i (but not the size of the AFSR). First we treat a special case.

Proposition 20.8.1. *If S is closed under addition and is closed under multiplication by elements of L, then the state transition function of A is L-linear.*

Proof. Addition and multiplication by fixed elements (the q_i) in $L[x]$ are L-linear operations. By the closure properties of S, if $w_1, w_2, w_1', w_2' \in L[x]$, $a_1, a_2 \in S$, and $u, v \in L$ satisfy $w_i = a_i + \pi w_i'$, then $uw_1 + vw_2 = (ua_1 + va_2) + \pi(uw_1' + vw_2')$ with $ua_1 + va_2 \in S$. Thus the entire state change operation is linear. □

It follows that such an A is equivalent to a linear feedback (not necessarily shift) register. By Theorem 10.6.1, however, the linear span of the output from such a register is at most its size.

Corollary 20.8.2. *Let σ be defined as above. If S is closed under addition and is closed under multiplication by elements of L, then the linear span of $\sigma(\mathbf{a})$ is at most*

$$md + \max(U + V - d, \deg(z)).$$

Proof. Under these hypotheses the degree of the memory is bounded by $\max(U + V - d, \deg(z))$. The result follows by representing the memory by its list of coefficients. □

One example of a set of representatives satisfying the closure properties is

$$S_0 = \{t(x) : deg(t) < d\}.$$

In general the AFSR A does not have a linear state change function. Let $a = a_0 + a_1\pi + \cdots$ be the π-adic number associated with $\mathbf{a}$ with coefficients $a_i \in S$. In the ring R_π of π-adic numbers, a can also be represented by a series $a = a'_0 + a'_1\pi + \cdots$ with the $a'_i \in S_0$. We relate $\mathbf{a}$ to a sequence over L in two stages. First we show that there is a finite state device that depends only on R, π, and S that takes $\mathbf{a} = a_0, a_1, \cdots$ as input and outputs $\mathbf{a}' = a'_0, a'_1, \cdots$. Second, the sequence $\mathbf{a}'$ can be generated by an AFSR defined over (R, π, S_0) whose length is at most r and whose memory is small.

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

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

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 8.2.2, if $a = u/q$, then

$$\deg(u) \le \max((m-1)d + U + V, md + \deg(z)).$$

Also, we have equality if $\deg(z) > U + V - d$. Now consider the AFSR over (R, π, S_0) that generates $\mathbf{a}'$ using the same representation of q. The size of this AFSR is m. If z' is the initial memory of this AFSR, then

$$u = \sum_{i=0}^{m-1} \sum_{j=0}^{m-i-1} q_i a'_j \pi^{i+j} - z'\pi^m.$$

Thus $\deg(z') + md \le \max(V + d - 1 + (m-1)d, (m-1)d + U + V, md + \deg(z))$ so $\deg(z') \le \max(V - 1, U + V - d, \deg(z))$. Combining this with Corollary 20.8.1 we have proved the following.

Theorem 20.8.3. *If* $\mathbf{a}$ *can be generated by an AFSR over* (R, π, S) *of length* m *with initial memory of degree* e*, then there is a sequence* $\mathbf{a}'$ *so that* $\sigma(\mathbf{a}')$ *has linear span at most* $md + \max(V - 1, U + V - d, e)$ *over* L *and so that* $\mathbf{a}'$ *can be transformed into* $\mathbf{a}$ *and* $\mathbf{a}$ *can be transformed into* $\mathbf{a}'$ *by finite state devices depending only on* R*,* π*, and* S *with* $p^{n(U-d)}$ *states.*

20.9 Exercises

1. Let F be a finite field, let I be an ideal in $F[x_1, \cdots, x_n]$, and let $R = F[x_1, \cdots, x_n]/I$. Let $\pi \in R$ and assume that $R/(\pi)$ is finite. Show that there exists a subset S in R that is a complete set of representatives modulo π and satisfies Hypothesis H1.

Chapter 21 Average and Asymptotic Behavior of Security Measures

Among all periodic sequences with a fixed period T, the m-sequences are those with the smallest linear span, and the ℓ-sequences are those with the smallest N-adic complexity. In this chapter we consider the opposite question: what is the linear and N-adic complexity of a random sequence? It turns out that, on average, these complexities are high.

There are two probability models in which this question can be made precise. Let Σ be a finite alphabet and let Σ^* and Σ^{ev} be the set of finite length sequences, and the set of eventually periodic sequences (respectively) over Σ. Assume that $\lambda : \Sigma^* \cup \Sigma^{ev} \to \mathbb{R}_{\geq 0}$ is a "complexity function" that satisfies the following properties.

S1: If $\mathbf{a} \in \Sigma^*$ (that is, $\mathbf{a}$ has finite length), then $\lambda(\mathbf{a}) \leq \text{length}(\mathbf{a})$.
S2: If $\mathbf{a} \in \Sigma^{ev}$ is periodic, then $\lambda(\mathbf{a}) \leq \text{period}(\mathbf{a})$.
S3: If $\mathbf{a} \in \Sigma^*$, $\mathbf{b} \in \Sigma^* \cup \Sigma^{ev}$, and $\mathbf{a}$ is a prefix of $\mathbf{b}$ then $\lambda(\mathbf{a}) \leq \lambda(\mathbf{b})$.

Consider the set P_n of sequences over Σ of length n. Make this into a probability space by assigning the same probability, $1/|\Sigma|^n$, to each sequence. In the first model we are interested in the expectation and variance of $\lambda(\mathbf{a})$, $\mathbf{a} \in P_n$.

If $\mathbf{a} \in P_n$ let $\mathbf{a}^\infty$ be the periodic sequence obtained by repeating $\mathbf{a}$ infinitely, and let $\lambda^\infty(\mathbf{a}) = \lambda(\mathbf{a}^\infty)$. For the second model we are interested in the expectation and variance of $\lambda^\infty(\mathbf{a})$, for $\mathbf{a} \in P_n$.

There is a third probability model that will be used in Chapter 21.3. It is relevant to the study of the asymptotic properties of security measures for sequences that are not eventually periodic. For this model, let Σ^∞ be the set of all infinite sequences over Σ and make this into a probability space using the following infinite product measure. Fix any finite sequence $a_0, a_1, \cdots, a_{k-1}$. Then the probability of the set of infinite sequences whose first k symbols are $a_0, a_1, \cdots, a_{k-1}$ is $1/|\Sigma|^k$. There is a unique way to extend this to a probability measure on Σ^∞, but we omit the details.

21.1 Average behavior of linear complexity

In this section we assume that our security measure λ is linear complexity. We let $\Sigma = \mathbb{F}_r$, where r is a power of a prime number p. In this section we let $\lambda(\mathbf{a})$ denote the linear span of the sequence $\mathbf{a}$. That is, $\lambda(\mathbf{a})$ is the least integer k such that there are elements $q_1, \cdots, q_k \in \mathbb{F}_r$ such that for all $n \geq k$ we have $a_n = q_1 a_{n-1} + \cdots + q_k a_{n-k}$. We refer to k as the length of the recurrence. When $k = 0$ we interpret the empty sum on the right as 0.

21.1.a Averaging for finite length sequences

In this section we compute the average linear span of fixed finite length sequences. That is, the linear span of a sequence $\mathbf{a} = a_0, a_1, \cdots, a_{n-1}$ is the smallest k so that there exist $q_1, q_2, \cdots, q_k$ with $a_m = q_1 a_{m-1} + \cdots + q_k a_{m-k}$ for all $m = k, k+1, \cdots, n-1$.

We begin by counting the the number of sequences of length n whose linear span is a given integer L with $0 \le L \le n$. We denote this quantity by $N_n(L)$. Thus the expected linear span of finite length sequences of length n is

$$E_n^{\text{lin}} = \frac{1}{r^n} \sum_{L=0}^{n} L N_n(L).$$

Theorem 21.1.1. *For $0 \le L \le n$ we have*

$$N_n(L) = \begin{cases} 1 & \text{if } L = 0 \\ (r-1) r^{\min(2L-1, 2n-2L)} & \text{if } 1 \le L \le n. \end{cases}$$

Proof. By definition the only sequence whose linear span is zero is the all zero sequence, so $N_n(0) = 1$. The length 1 sequence whose only element is $a \neq 0 \in \mathbb{F}_r$ satisfies every length 1 recurrence, so $N_1(1) = r - 1$ as claimed. Now we proceed by induction on n.

Suppose that $\mathbf{a}$ is a sequence of length $n \ge 1$ over $\mathbb{F}_r$. We consider the effect of adding a symbol to $\mathbf{a}$. There are two cases.

Suppose $\lambda(\mathbf{a}) \ge (n+1)/2$. Then by Lemma 18.2.4 adding a symbol to $\mathbf{a}$ leaves the linear span unchanged.

Suppose that $\lambda(\mathbf{a}) \le n/2$. Let $f(x)/q(x)$ be the optimal rational approximation to $\mathbf{a}$. Thus $\deg(q(x)) \le n/2$ and $\deg(f(x)) \le n/2 - 1$. If we add the next symbol in the power series expansion of $f(x)/q(x)$ to $\mathbf{a}$, then the linear span is unchanged. On the other hand, suppose we add a different symbol to $\mathbf{a}$. Let $f'(x)/q'(x)$ be the optimal rational approximation to the lengthened sequence. Then

$$\frac{f(x)}{q(x)} - \frac{f'(x)}{q'(x)} \equiv u x^n \pmod{x^{n+1}}$$

for some $u \neq 0 \in \mathbb{F}_r$. Thus $f(x)q'(x) - f'(x)q(x) \equiv v x^n \pmod{x^{n+1}}$ for some $v \neq 0 \in \mathbb{F}_r$. Let $L = \Phi(f'(x), q'(x))$. Then

$$n \le \deg(f(x)q'(x) - f'(x)q(x)) \le \frac{n}{2} - 1 + L.$$

It follows that $L \ge n/2 + 1$, so the linear span changes. By Lemma 18.2.4 the new linear span is $n + 1 - L$.

Combining these results we see that

$$N_{n+1}(L) = \begin{cases} N_n(L) & \text{if } L \leq n/2 \\ rN_n(L) & \text{if } L = (n+1)/2 \\ (r-1)N_n(n+1-L) + rN_n(L) & \text{if } L > (n+1)/2. \end{cases}$$

It is then straightforward to check by induction that

$$N_{n+1}(L) = \begin{cases} 1 & \text{if } L = 0 \\ (r-1)r^{\min(2L-1, 2n+2-2L)} & \text{if } 1 \leq L \leq n. \end{cases}$$ □

This is equivalent to saying that

$$N_n(L) = \begin{cases} 1 & \text{if } L = 0 \\ (r-1)r^{2L-1} & \text{if } 1 \leq L \leq n/2. \\ (r-1)r^{2n-2L} & \text{if } (n+1)/2 \leq L \leq n. \end{cases} \tag{21.1}$$

We can use these values to compute the expectation.

Theorem 21.1.2. *The expected value of the linear span of sequences with length n and with terms in $\mathbb{F}_r$ is given by*

$$E_n^{lin} \in \frac{n}{2} + O(1/r).$$

Proof. We have

$$\begin{aligned} E_n^{\text{lin}} &= \frac{1}{r^n} \sum_{L=0}^{n} L N_n(L) \\ &= \frac{1}{r^n} \left(\sum_{L=1}^{\lfloor n/2 \rfloor} L(r-1)r^{2L-1} + \sum_{L=\lceil (n+1)/2 \rceil}^{n} L(r-1)r^{2n-2L} \right). \end{aligned}$$

First we evaluate the first sum:

$$\begin{aligned} A &\stackrel{def}{=} \sum_{L=1}^{\lfloor n/2 \rfloor} L r^{2L-1} \\ &= r \sum_{L=1}^{\lfloor n/2 \rfloor} L (r^2)^{L-1} \\ &= r \frac{\lfloor n/2 \rfloor \, r^{2(\lfloor n/2 \rfloor + 1)} - (\lfloor n/2 \rfloor + 1) r^{2\lfloor n/2 \rfloor} + 1}{(r^2-1)^2} \end{aligned}$$

$$= \begin{cases} r\dfrac{nr^{n+2}-(n+2)r^n+2}{2(r^2-1)^2} & \text{if } n \text{ is even} \\ r\dfrac{(n-1)r^{n+1}-(n+1)r^{n-1}+2}{2(r^2-1)^2} & \text{if } n \text{ is odd.} \end{cases}$$

Also,

$$\begin{aligned} B &\stackrel{def}{=} \sum_{L=\lceil (n+1)/2 \rceil}^{n} L r^{2n-2L} \\ &= \left\lceil \frac{n-1}{2} \right\rceil r^{2\lfloor (n-1)/2 \rfloor} \sum_{i=0}^{\lfloor (n-1)/2 \rfloor} (r^{-2})^i + r^{2\lfloor (n-1)/2 \rfloor} \sum_{i=1}^{\lfloor (n+1)/2 \rfloor} i (r^{-2})^{i-1} \\ &= \left\lceil \frac{n-1}{2} \right\rceil r^{2\lfloor (n-1)/2 \rfloor} \frac{r^{-2\lfloor (n+1)/2 \rfloor}-1}{r^{-2}-1} \\ &\quad + r^{2\lfloor (n-1)/2 \rfloor} \frac{\lfloor (n+1)/2 \rfloor r^{-2\lfloor (n+3)/2 \rfloor} - \lfloor (n+3)/2 \rfloor r^{\lfloor (n+1)/2 \rfloor} + 1}{(r^{-2}-1)^2} \\ &= \left\lceil \frac{n-1}{2} \right\rceil \frac{r^{2\lfloor (n+1)/2 \rfloor}-r}{r^2-1} \\ &\quad + \frac{r^{2\lfloor (n+3)/2 \rfloor} - \lfloor (n+3)/2 \rfloor r^3 + \lfloor (n+1)/2 \rfloor r^2}{(r^2-1)^2} \\ &= \frac{\left\lceil \frac{n+1}{2} \right\rceil r^{2\lfloor (n+3)/2 \rfloor} - \left\lceil \frac{n-1}{2} \right\rceil r^{2\lfloor (n+1)/2 \rfloor} - (n+1)r^3 + \lfloor (n+1)/2 \rfloor r^2 + \left\lceil \frac{n-1}{2} \right\rceil}{(r^2-1)^2} \\ &= \begin{cases} \dfrac{(n+2)r^{n+2} - nr^n - 2(n+1)r^3 + nr^2 + n}{2(r^2-1)^2} & \text{if } n \text{ is even} \\ \dfrac{(n+1)r^{n+3} - (n-1)r^{n+1} - 2(n+1)r^3 + (n+1)r^2 + n - 1}{2(r^2-1)^2} & \text{if } n \text{ is odd.} \end{cases} \end{aligned}$$

Combining these, when n is even we have

$$\begin{aligned} E_n^{\text{lin}} &= \frac{r-1}{r^n}(A+B) \\ &= \frac{nr^3 + (n+2)r^2 - (n+2)r - n - 2(n+1)r^{3-n} + nr^{2-n} + 2r^{1-n} + nr^{-n}}{2(r^2-1)(r+1)} \\ &\in \frac{n}{2} + O(1/r) \end{aligned}$$

Similarly, when n is odd we have

$$E_n^{\text{lin}} \in \frac{n}{2} + O(1/r).$$

□

21.1.b Averaging for periodic sequences

In this section we compute the average linear span of periodic sequences with a fixed period n (that is, whose least period divides n). Let $n = p^v m$ for some integers v and m with $\gcd(m, p) = 1$. If $\mathbf{a}$ is a sequence with period n, then we let $a(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}$. Recall from Section 18.3 that we defined the generalized discrete Fourier transform (GDFT) of a polynomial. It is a matrix with p^v rows, indexed by $0, 1, \cdots, p^v - 1$, and n columns, indexed by $0, 1, \cdots, n-1$. The entry in the ith row and jth column is $a^{[i]}(b^j)$ where b is a primitive mth root of 1 in some extension field of $\mathbb{F}_r$, and $a^{[i]}(x)$ is the ith Hasse derivative of $a(x)$,

$$a^{[i]}(x) = \sum_{i=j}^{n-1} \binom{i}{j} a_i x^{i-j}.$$

In Theorem 18.4.4 we showed that we can also characterize the linear span as the Günther weight — the number of entries that are nonzero or below a nonzero entry — of the GDFT of $a(x)$. It is this descriptstyleion that we make use of here.

We also make use of the notion of cyclotomic cosets modulo m relative to r from Section 3.2.c,

$$C_k = \{k, rk \pmod m, r^2 k \pmod m, \cdots\}.$$

Let $\ell_k = |C_k|$.

Lemma 21.1.3. *Let $0 \le k < m$ and let $j \equiv kr^t \pmod m \in C_k$. Then for $0 \le i < p^v$ we have*

$$a^{[i]}(b^k) \in \mathbb{F}_{r^{\ell_k}} \text{ and } a^{[i]}(b^j) = (a^{[i]}(b^k))^{r^t}.$$

Proof. We have $kr^{\ell_k} \equiv k \pmod m$. Since b is a primitive mth root of unity, we have

$$(a^{[i]}(b^k))^{r^{\ell_k}} = a^{[i]}(b^{kr^{\ell_k}}) = a^{[i]}(b^k).$$

In particular, $a^{[i]}(b^k) \in \mathbb{F}_{r^{\ell_k}}$. Also,

$$a^{[i]}(b^j) = a^{[i]}(b^{kr^t}) = a^{[i]}(b^j)^{r^t}.$$

□

Each column of the GDFT corresponds to a power of b. Equivalently, each column corresponds to the integer exponent of b. Thus we can partition the columns of the GDFT into sets whose corresponding exponents are all in the same cyclotomic coset. We refer to such a set as a *cyclotomic*

coset of columns. It follows from Lemma 21.1.3 that in any row of the GDFT, the entries in a common cyclotomic coset of columns are either all zero or are all nonzero. Thus all columns in a given cyclotomic coset of columns contribute the same amount to the Günther weight of the GDFT. For a given k, let t_k be the smallest row index so that

$$a^{[t_k]}(b^k) \neq 0,$$

if such a row exists. Then C_k contributes $(p^v - t_k)\ell_k$ to the Günther weight of the GDFT. In particular, if $C_{k_1}, \cdots, C_{k_s}$ are the distinct cyclotomic cosets, then the linear span is a linear combination

$$\lambda(\mathbf{a}) = \sum_{i=1}^{s} w_i \ell_{k_i},$$

for some integers w_i.

Moreover, the GDFT is determined by the s columns with indices $k_1, \cdots, k_s$. Each column has height p^v and the ith of these columns has entries in

$$\mathbb{F}_{r^{\ell_{k_i}}}.$$

Let $\mathcal{M}$ denote the set of all p^v by s matrices such that the ith column has values in this field. The cardinality of $\mathcal{M}$ is

$$\prod_{i=1}^{s} r^{p^v \ell_{k_i}} = r^{\sum_{i=1}^{s} p^v \ell_{k_i}} = r^n,$$

which is exactly the number of sequences with period n. By Theorem 18.4.3 the GDFT is invertible, so there is a unique sequence with any given GDFT. Thus each matrix of this form occurs as the GDFT of a sequence with period n. This allows us to count the sequences with a given linear span.

Theorem 21.1.4. *Let $n = p^v m$. Let the cardinalities of the distinct cyclotomic cosets modulo m relative to r be $\ell_{k_1}, \cdots, \ell_{k_s}$. Then the expected value of the linear span of periodic sequences with period n and with terms in $\mathbb{F}_r$ is given by*

$$E_n^{lin,per} = n - \sum_{i=1}^{s} \frac{\ell_{k_i}(1 - r^{-p^v \ell_{k_i}})}{r^{\ell_{k_i}} - 1}.$$

Proof. We have seen that $E_n^{\text{lin,per}}$ equals the expected value of the Günther weight of a p^v by m matrix K in GDFT format. Let M be the associated matrix in $\mathcal{M}$ (consisting only of the columns with indices $k_1, \cdots, k_s$). Let z_j denote the jth column of M (that is, the k_jth column of M),

and let $\mu(z_j)$ denote the least positive integer u so that the uth coordinate is not zero. Then the Günther weight of K is

$$g(M) = \sum_{\substack{j=1 \\ z_j \neq 0}}^{s} \ell_{k_j}(p^v - \mu(z_j) + 1).$$

Thus

$$\begin{aligned}
E_n^{\text{lin,per}} &= \frac{1}{r^n} \sum_{M \in \mathcal{M}} \sum_{\substack{j=1 \\ z_j \neq 0}}^{s} \ell_{k_j}(p^v - \mu(z_j) + 1) \\
&= \frac{1}{r^n} \sum_{j=1}^{s} \ell_{k_j} \sum_{\substack{M \in \mathcal{M} \\ z_j \neq 0}}^{s} (p^v - \mu(z_j) + 1) \\
&= \frac{p^v}{r^n} \sum_{j=1}^{s} \ell_{k_j} \sum_{\substack{M \in \mathcal{M} \\ z_j \neq 0}}^{s} 1 - \frac{1}{r^n} \sum_{j=1}^{s} \ell_{k_j} \sum_{\substack{M \in \mathcal{M} \\ z_j \neq 0}} (\mu(z_j) - 1) \\
&= T_1 - T_2.
\end{aligned}$$

For T_1 we have

$$\begin{aligned}
T_1 &= \frac{p^v}{r^n} \sum_{j=1}^{s} \ell_{k_j}(r^{p^v \ell_{k_j}} - 1) r^{n - p^v \ell k_j} \\
&= p^v \sum_{j=1}^{s} \ell_{k_j}(1 - r^{-p^v \ell_{k_j}}) \\
&= n - p^v \sum_{j=1}^{s} \frac{\ell_{k_j}}{r^{p^v \ell_{k_j}}}.
\end{aligned}$$

For T_2 we have

$$\begin{aligned}
T_2 &= \frac{1}{r^n} \sum_{j=1}^{s} \ell_{k_j} \sum_{u=1}^{p^v} (u-1) \sum_{\substack{M \in \mathcal{M} \\ \mu(z_j) = u}} 1 \\
&= \frac{1}{r^n} \sum_{j=1}^{s} \ell_{k_j} \sum_{u=1}^{p^v} (u-1)(r^{\ell_{k_j}} - 1)(r^{\ell_{k_j}})^{p^v - u} r^{n - p^v \ell_{k_j}}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{s} \ell_{k_j} \sum_{u=1}^{p^v} (u-1)(r^{\ell_{k_j}}-1) r^{-u\ell_{k_j}} \\
&= \sum_{j=1}^{s} \ell_{k_j} \left(1 - \frac{1}{r^{\ell_{k_j}}}\right) \sum_{u=0}^{p^v-1} u r^{-u\ell_{k_j}} \\
&= \sum_{j=1}^{s} \frac{\ell_{k_j}(r^{\ell_{k_j}}-1)}{r^{\ell_{k_j}}} \frac{(p^v-1)r^{-(p^v+1)\ell_{k_j}} - p^v r^{-p^v\ell_{k_j}} + r^{-\ell_{k_j}}}{(r^{-\ell_{k_j}}-1)^2} \\
&= \sum_{j=1}^{s} \frac{\ell_{k_j}}{r^{\ell_{k_j}}-1} \left((p^v-1)r^{-p^v\ell_{k_j}} - p^v r^{-(p^v-1)\ell_{k_j}} + 1\right) \\
&= \sum_{j=1}^{s} \frac{\ell_{k_j}(1-r^{-p^v\ell_{k_j}})}{r^{\ell_{k_j}}-1} - p^v \sum_{j=1}^{s} \frac{\ell_{k_j}}{r^{p^v\ell_{k_j}}}.
\end{aligned}
$$

Taking the difference between the expressions for T_1 and T_2 proves the theorem. □

Corollary 21.1.5. *The expectation of the linear span of periodic sequences with period $n = p^v m$ is lower bounded by*

$$E_n^{lin,per} > n - \frac{m}{r-1}.$$

Proof. The remaining cosets all have at least two elements. By Theorem 21.1.4 we have

$$
\begin{aligned}
E_n^{\text{lin,per}} &= n - \sum_{i=1}^{s} \frac{\ell_{k_i}(1-r^{-p^v\ell_{k_i}})}{r^{\ell_{k_i}}-1} \\
&> n - \sum_{i=1}^{s} \frac{\ell_{k_i}}{r^{\ell_{k_i}}-1} \\
&\geq n - \sum_{i=1}^{s} \frac{\ell_{k_i}}{r-1} \\
&= n - \frac{m}{r-1}
\end{aligned}
$$

as claimed. □

The variance can be derived similarly, but we omit the details of the proof.

Theorem 21.1.6. *Let $n = p^v m$. Let the cardinalities of the distinct cyclotomic cosets modulo n relative to r be $\ell_{k_1}, \cdots, \ell_{k_s}$. Then the variance of the linear span of periodic sequences with period*

n and with terms in $\mathbb{F}_r$ is given by

$$V_n^{lin,per} = n - \sum_{i=1}^{s} \frac{\ell_{k_i}(1 - r^{-p^v \ell_{k_i}})}{q^{\ell_{k_i}} - 1}.$$

21.2 Average behavior of N-adic complexity

In this section we consider N-adic complexity, where $N \geq 2$ is a natural number. We let $\Sigma = \mathbb{Z}/(N)$. We let $\lambda(\mathbf{a})$ denote the N-adic complexity of the sequence $\mathbf{a}$. That is, $\lambda(\mathbf{a})$ is the least real number $\log(\min(|f|, |q|))$ with

$$\frac{f}{q} = \sum_{i=0}^{\infty} a_i N^i.$$

If $\mathbf{a}$ is periodic, then $\lambda(\mathbf{a}) = \log |q|$.

21.2.a Average behavior of N-adic complexity for periodic sequences

Let t be a positive integer. Then a sequence has period t if and only if it is the coefficient sequence of the N-adic expansion of a rational number $-f/(N^t - 1)$ with $0 \leq f \leq N^n - 1$. This happens if and only if the connection integer of the sequence divides $N^t - 1$. Thus to consider the set of sequences with period t we need to consider the set of sequences whose connection integer divides $N^t - 1$. More generally, let q be any odd integer. We want to consider the set of sequences whose minimal connection integers divide q. A sequence has minimal connection integer q if and only if it is the coefficient sequence of the N-adic expansion of a rational number $-f/q$ with $0 \leq f \leq q$ and $\gcd(q, f) = 1$. Let $M(q)$ denote the number of such sequences. Then $M(q) = \phi(q)$ (Euler's totient function) if $q \neq 1$, and $M(1) = 2$ (f can be 0 or 1 if $q = 1$). There are $q + 1$ periodic sequences whose minimal connection integer divides q, and if the minimal connection integer of a sequence is d, then its N-adic complexity is $\log_N(d)$.

It follows that the average N-adic complexity of sequences with connection integer dividing q is

$$\begin{aligned} E^{N\text{-adic,per}}(q) &= \frac{1}{q+1}\left(\sum_{d|q, d\neq 1} \log_N(d)\phi(d) + 2\right) \\ &= \frac{1}{q+1}\left(\sum_{d|q} \log_N(d)\phi(d) + 2\right). \end{aligned}$$

Similarly, the second moment is

$$E_2^{N\text{-adic,per}}(q) = \frac{1}{q+1}\left(\sum_{d|q}\log_N(d)^2\phi(d)+4\right).$$

Thus the variance is

$$\begin{aligned} V^{N\text{-adic,per}}(q) &= E_2^{N\text{-adic,per}}(q) - E^{N\text{-adic,per}}(q)^2 \\ &= \frac{1}{q+1}\left(\sum_{d|q}\log_N(d)^2\phi(d)+4\right) - \frac{1}{(q+1)^2}\left(\sum_{d|q}\log_N(d)\phi(d)+2\right)^2. \end{aligned}$$

For example, if q is prime, then

$$E^{N\text{-adic,per}}(q) = \frac{2+\log_N(q)(q-1)}{q+1} = \log_N(q) - \frac{2(\log_N(q)-1)}{q+1} \sim \log_N(q)$$

and

$$V^{N\text{-adic,per}}(q) \sim \frac{\log_N(q)^2}{q}.$$

More generally, suppose that $q = \prod_{i=1}^k f_i^{t_i}$ where $f_i \in \mathbb{Z}$ is a prime number and $t_i \geq 1$, $i = 1, \cdots, k$. Then

$$\begin{aligned} E^{N\text{-adic,per}}(q) &= \frac{1}{q+1}\left(\sum_{\substack{J=(j_1,\cdots,j_k)\\ 0\leq j_i\leq t_j\\ j_i \text{not all } 0}} \log_N\left(\prod_{i=1}^k f_i^{j_i}\right)\phi\left(\prod_{i=1}^k f_i^{j_i}\right)+2\right) \\ &= \frac{1}{q+1}\left(\sum_{\substack{J=(j_1,\cdots,j_k)\\ 0\leq j_i\leq t_i}} \sum_{i=1}^k j_i\log_N(f_i)\prod_{i=1}^k\phi(f_i^{j_i})\right)+\frac{2}{q+1} \\ &= \frac{1}{q+1}\sum_{\ell=1}^k\prod_{\substack{i=1\\ i\neq\ell}}^k\left(\sum_{j_i=0}^{t_i}\phi(f_i^{j_i})\right)\sum_{j_\ell=0}^{t_\ell} j_\ell\log_N(f_\ell)\phi(f_\ell^{j_\ell})+\frac{2}{q+1} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{q+1}\sum_{\ell=1}^{k}\prod_{\substack{i=1\\ i\neq\ell}}^{k} f_i^{t_i}\sum_{j_\ell=0}^{t_\ell} j_\ell \log_N(f_\ell)\phi(f_\ell^{j_\ell}) + \frac{2}{q+1}\\
&= \frac{1}{q+1}\sum_{\ell=1}^{k}\frac{q}{f_\ell^{t_\ell}}\sum_{j_\ell=0}^{t_\ell} j_\ell \log_N(f_\ell)\phi(f_\ell^{j_\ell}) + \frac{2}{q+1}\\
&= \frac{q}{q+1}\sum_{\ell=1}^{k}\frac{\log_N(f_\ell)}{f_\ell^{t_\ell}}\sum_{j_\ell=0}^{t_\ell} j_\ell \phi(f_\ell^{j_\ell}) + \frac{2}{q+1}\\
&= \frac{q}{q+1}\sum_{\ell=1}^{k}\frac{\log_N(f_\ell)}{f_\ell^{t_\ell}}\sum_{j_\ell=1}^{t_\ell} j_\ell (f_\ell^{j_\ell} - f_\ell^{j_\ell-1}) + \frac{2}{q+1}\\
&= \frac{q}{q+1}\sum_{\ell=1}^{k}\frac{\log_N(f_\ell)}{f_\ell^{t_\ell}}\left(t_\ell f_\ell^{t_\ell} - \frac{f_\ell^{t_\ell}-1}{f_\ell-1}\right) + \frac{2}{q+1}\\
&= \frac{q}{q+1}\sum_{\ell=1}^{k}\log_N(f_\ell)\left(t_\ell - \frac{1-f_\ell^{-t_\ell}}{f_\ell-1}\right) + \frac{2}{q+1}\\
&= \frac{q}{q+1}\log_N(q) - \frac{q}{q+1}\sum_{\ell=1}^{k}\log_N(f_\ell)\left(\frac{1-f_\ell^{-t_\ell}}{f_\ell-1}\right) + \frac{2}{q+1}.
\end{aligned}
$$

Now we want to bound the summation in the last line. For any $f \geq 2$ we have

$$\frac{\log_N(f)}{f-1} \leq \frac{2\log_N(f)}{f} \leq \frac{2\log_2(f)}{f},$$

and if $t \geq 1$ then $1 - f^{-t} < 1$. Thus it suffices to bound

$$A \stackrel{def}{=} \sum_{\ell=1}^{k}\frac{\log_2(f_\ell)}{f_\ell}$$

in terms of q. In bounding A, if we replace q by a smaller integer while leaving A unchanged (that is, replace q by a smaller integer with the same prime factors), or replace A by a larger number

while leaving q unchanged, this can only weaken our bound. So first we replace q by

$$\prod_{\ell=1}^{k} f_\ell.$$

That is, we may assume that each $t_\ell = 1$. The function $\log_2(x)/x$ is decreasing for integers $x \geq 3$. Thus decreasing an f_ℓ will increase A while decreasing q. Thus we may assume that the set of f_ℓs consists of the first k prime numbers.

Let n be a positive integer and consider the primes between 2^n and 2^{n-1}. By the prime number theorem there are asymptotically

$$\frac{2^n}{\ln(2^n)} = \frac{2^n}{n\ln(2)}$$

such primes. Each contributes at most $n/2^n$ to A, so the total contribution to A is at most $1/\ln(2)$. It also follows from the prime number theorem that the largest 2^n we must consider is asymptotically $k\ln(k)$. It follows that $A \in O(\log(k))$.

Now consider q. The contribution to q from the primes between 2^n and 2^{n+1} is asymptotically at least

$$(2^n)^{2^n/(n\ln(2))} = 2^{2^n/\ln(2)}.$$

The largest n we need is at most $\log(k\ln(k)) < 2\log(k)$. Thus

$$\begin{aligned} q &\in \Omega\left(\prod_{n=1}^{\log(k)} 2^{2^n/\ln(2)}\right) \\ &= \Omega(2^{\sum_{n=1}^{\log(k)} 2^n/\ln(2)}) \\ &= \Omega(2^{2^{\log(k)+1}/\ln(2)}) \\ &= \Omega(2^{2k/\ln(2)}). \end{aligned}$$

Therefore $A \in O(\log\log(q))$.

Theorem 21.2.1. *We have*

$$E^{N\text{-}adic,per}(q) \in \log_N(q) - O(\log\log(q)).$$

If we take $q = N^t - 1$, then we obtain the expectation and variance of the N-adic complexity for sequences of period t. However, even if we specialize to this case we do not know how to say more precisely what the expectation is. In general we do not know a precise description of the factorization of $N^t - 1$. We do obtain good asymptotic bounds, however.

Corollary 21.2.2. *The expected N-adic complexity of periodic sequences of period t is*

$$E^{N\text{-}adic,per}(N^t - 1) \in t - O(\log(t)).$$

We leave it to an exercise to show that this bound is tight.

The derivation of the second moment is essentially the same up to the seventh step of the above derivation. This gives

$$E_2^{N\text{-adic,per}}(q) = \frac{q}{q+1}\sum_{\ell=1}^{k}\log_N(f_\ell)t_\ell^2 - \frac{q}{q+1}\sum_{\ell=1}^{k}\log_N(f_\ell)\left(\frac{2t_\ell+1}{f_\ell-1} - \frac{2+f_\ell^{1-t_\ell}-3f_\ell^{-t_\ell}}{(f_\ell-1)^2}\right)$$

21.3 Asymptotic behavior of security measures

We next consider the asymptotic behavior of security or randomness measures for infinite sequences. The kinds of measures we are interested in arise in the following manner. As in Section 5.1.c, a sequence generator is a 4-tuple (U, Σ, f, g) where U and Σ are sets, $f : U \to U$, and $g : U \to \Sigma$. For a given initial state $\mathbf{s} \in U$, the sequence generator outputs the infinite sequence $U(\mathbf{s}) = g(\mathbf{s}), g(f(\mathbf{s})), g(f^2(\mathbf{s})), \cdots$. As in Section 18.1, we fix a class $\mathcal{G}$ of sequence generators with a common finite output alphabet Σ, such that every eventually periodic sequence over Σ is generated by at least one element of $\mathcal{G}$. We also assume there is a notion of the size of a generator in $\mathcal{G}$ with a given initial state, a positive real number. In general this measure should be close to the number, n, of elements of Σ needed to represent a state of the generator. Typically "close" means differing from n by at most $O(\log(n))$. For example, consider an LFSR over a finite ring R. Then $\Sigma = R$ and the size of an LFSR is the number of cells. For an N-ary FCSR, $\Sigma = \mathbb{Z}/(N)$ and size is number of cells in the basic FCSR plus the maximum number of elements Σ needed to represent the memory given a particular initial state. Alternatively, if the output from a given initial state is the N-adic expansion of p/q, then we can define the size to be the log base N of the maximum of $|p|$ and $|q|$. These two definitions are not equal, but as seen in Chapter 19, they differ by an additive log factor.

We denote by $\lambda^{\mathcal{G}}(\mathbf{a})$ the minimum size of a generator in $\mathcal{G}$ that outputs the sequence $\mathbf{a}$, if there is such a generator. For a sequence $\mathbf{a}$ that is the output from no generator in $\mathcal{G}$, $\lambda^{\mathcal{G}}(\mathbf{a})$ is undefined. However, we can apply the measure to the various prefixes of $\mathbf{a}$ and try to understand the asymptotic behavior. For $n > 0$, let $\lambda_n^{\mathcal{G}}(\mathbf{a})$ denote the minimum size of a generator from $\mathcal{G}$ that outputs the first n symbols of $\mathbf{a}$ as its first n outputs (and after that we don't care what it outputs). The sequence of numbers $(\lambda_n^{\mathcal{G}}(\mathbf{a}) : n = 1, 2, \cdots)$ is called the *$\mathcal{G}$-complexity profile of* $\mathbf{a}$. For a sequence that is not eventually periodic, the limit of the $\lambda_n^{\mathcal{G}}(\mathbf{a})$ is infinite, so we normalize the measure by letting $\delta_n^{\mathcal{G}}(\mathbf{a}) = \lambda_n^{\mathcal{G}}(\mathbf{a})/n$. For the typical measures we are interested in we have

$$\lambda_n^{\mathcal{G}}(\mathbf{a}) \leq n + O(\log(n)),$$

so that

$$0 \leq \delta_n^G(\mathbf{a}) \leq 1 + o(n).$$

In general the $\delta_n^{\mathcal{G}}(\mathbf{a})$ do not have a single limit, but rather may vary over a range of values. More precisely, there are various infinite sequences of integers $m_1 < m_2 < \cdots$ so that

$$\lim_{n\to\infty} \delta_{m_n}^{\mathcal{G}}(\mathbf{a})$$

exists. When the limit exists it is called an *accumulation point*. the set of accumulation points is denoted by $\Upsilon(\mathbf{a})$. Our goal is to determine what sets of accumulation points are possible.

It is immediate for such a measure $\lambda_n^{\mathcal{G}}(\mathbf{a})$ that

$$\lambda_{n+1}^{\mathcal{G}}(\mathbf{a}) \geq \lambda_n^{\mathcal{G}}(\mathbf{a})$$

for all $n \geq 1$, so that

$$\delta_{n+1}^{\mathcal{G}}(\mathbf{a}) \geq \frac{n}{n+1}\delta_n^{\mathcal{G}}(\mathbf{a}) \geq \delta_n^{\mathcal{G}}(\mathbf{a}) - \frac{1}{n+1}. \tag{21.2}$$

This allows us to show that the set of accumulation points is a closed interval.

Theorem 21.3.1. *Let $\{\lambda_n : n = 1, 2, \infty\}$ be a sequence of real numbers, $1 \leq \lambda_n \leq n$, satisfying $\lambda_n \leq \lambda_{n+1}$ for all $n = 1, 2, \cdots$. Let $\delta_n = \lambda_n/n \in [0, 1]$. Then the set Υ of accumulation points of the δ_n is a closed interval $[B, C]$.*

Proof. First note that under these hypotheses, for any n we have

$$\begin{aligned}
\delta_n - \delta_{n+1} &= \frac{\lambda_n}{n} - \frac{\lambda_{n+1}}{n+1} \\
&\leq \lambda_n(\frac{1}{n} - \frac{1}{n+1}) \\
&= \frac{\lambda_n}{n(n+1)} \\
&< \frac{1}{n}.
\end{aligned}$$

Let B and C be the least and largest accumulation points of Υ, respectively. Then there are sequences of integers $n_1 < n_2 < \cdots$ and $m_1 < m_2 < \cdots$ so that

$$\lim_{i\to\infty} \delta_{n_i} = B \text{ and } \lim_{i\to\infty} \delta_{m_i} = C.$$

By deleting some integers, we may assume that $m_1 < n_1 < m_2 < n_2 < \cdots$ and that

$$|B - \delta_{n_i}| > |B - \delta_{n_{i+1}}|$$

and

$$|C - \delta_{m_i}| > |C - \delta_{m_{i+1}}|$$

for all i. Let $B < D < C$. Take i sufficiently large that

$$\delta_{n_i} < D \text{ and } \delta_{m_i} > D.$$

Let k_i be the largest index between m_i and n_i so that

$$\delta_{k_i} \geq D.$$

Then

$$\delta_{k_i+1} < D$$

and

$$\delta_{k_i} - \delta_{k_i+1} \leq 1/k_i$$

by equation (21.2). It follows that

$$\lim_{i \to \infty} \delta_{k_i} = D.$$

That is, every real number in the interval $[B, C]$ is an accumulation point, which proves the theorem. □

21.4 Asymptotic linear complexity

In this section we consider asymptotic linear complexity. Here $\mathcal{G}$ is the set of LFSRs over a finite field F. This case was first considered by Dai, Jiang, Imamura, and Gong with $F = \mathbb{F}_2$ using the theory of continued fractions [34]. An easier approach is similar to (but much simpler than) the proof of Theorem 21.5.4. The key fact, which follows from Lemma 18.2.4, is that that for every n, $\Phi(f_{n+1}, q_{n+1}) = \max(\Phi(f_n, q_n), n + 1 - \Phi(f_n, q_n))$, where for a pair of polynomials f, q we have $\Phi(f, q) = \max(\deg(f) + 1, \deg(q))$. We omit the details of the proof.

Theorem 21.4.1. *Let* $\mathbf{a} = a_0, a_1, \cdots$ *be a sequence over a finite field* F*, and suppose that* $\mathbf{a}$ *is not eventually periodic. Then* $\Upsilon(\mathbf{a}) = [B, 1 - B]$ *for some real number* B*.*

Dai, Imamura, and Yang, and Feng and Dai also studied this problem for vector valued non-periodic sequences [33, 46]. In this setting, however, there are much more limited results. They showed that there is a specific number associated with the generalized continued fraction expansion of a multisequence that is a lower bound for the maximum accumulation point and an upper bound for the minimum accumulation point. One might ask whether $\Upsilon(\mathbf{a})$ is a closed interval centered at this number. Theorem 21.3.1 shows that the set is a closed interval but leaves open the remainder of this question.

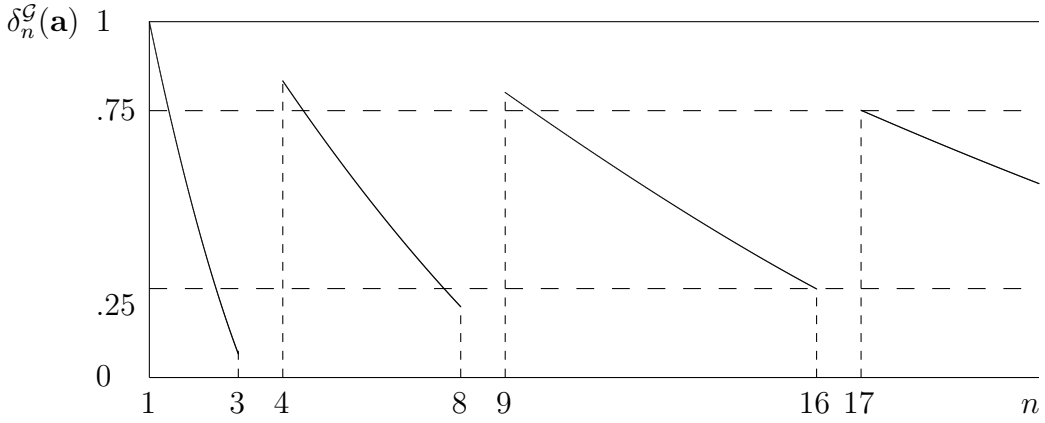


Figure 21.1: A Plot of $\delta_n^{\mathcal{G}}(\mathbf{a})$ for $B = .25$.

21.4.a All balanced intervals occur as $\Upsilon(\mathbf{a})$s

We denote by β the function that associates the real number B with the sequence $\mathbf{a}$, where $\Upsilon(\mathbf{a}) = [B, 1-B]$. That is,

$$\beta : \{\mathbf{a} = a_0, a_1, \cdots, a_i \in F\} \to [0, 1/2]$$

and $\beta(\mathbf{a})$ is the least accumulation points of the set of linear complexities of prefixes of $\mathbf{a}$. In this section we see that β is surjective.

Theorem 21.4.2. *For every $B \in [0, 1/2]$ there is a sequence $\mathbf{a}$ with entries in F with $\Upsilon(\mathbf{a}) = [B, 1-B]$.*

Proof. Let $0 \leq B < 1/2$. We build $\mathbf{a}$ with $\beta(\mathbf{a}) = B$ in stages. The idea is to let $\delta_n^{\mathcal{G}}(\mathbf{a})$ decrease by keeping $\lambda_n^{\mathcal{G}}(\mathbf{a})$ constant, until $\delta_n^{\mathcal{G}}(\mathbf{a})$ is first less than B, then make $\lambda_{n+1}^{\mathcal{G}}(\mathbf{a})$ be different from $\lambda_n^{\mathcal{G}}(\mathbf{a})$, so that it jumps to a much larger value. Then repeat this process. The effect is described by Figure 21.1.

Suppose that for some $m > 0$ we have chosen $n_1 < n_2 \leq \cdots < n_{k-1} < m \in \mathbb{Z}^+$ and $a_0, \cdots, a_{m-1} \in \{0, 1, \cdots, N-1\}$ so that for each $i = 1, 2, \cdots, k-2$ we have

$$\lambda_{n_i}^{\mathcal{G}}(\mathbf{a}) < \lambda_{n_i+1}^{\mathcal{G}}(\mathbf{a}) = \lambda_{n_i+2}^{\mathcal{G}}(\mathbf{a}) = \cdots = \lambda_{n_{i+1}}^{\mathcal{G}}(\mathbf{a}) \tag{21.3}$$

and

$$B > \delta_{n_i}^{\mathcal{G}}(\mathbf{a}) \geq \frac{1}{n_i}, \tag{21.4}$$

and we have

$$\lambda_{n_{k-1}}^{\mathcal{G}}(\mathbf{a}) < \lambda_{n_{k-1}+1}^{\mathcal{G}}(\mathbf{a}) = \lambda_{n_{k-1}+2}^{\mathcal{G}}(\mathbf{a}) = \cdots = \lambda_m^{\mathcal{G}}(\mathbf{a}).$$

Suppose that f and q are polynomials satsifying

$$\frac{f}{q} \equiv \sum_{i=0}^{m-1} a_i x^i \pmod{x^m}.$$

and that $\lambda_m^{\mathcal{G}}(\mathbf{a}) = \Phi(f, q)$. Moreover, let

$$\frac{f}{q} \equiv \sum_{i=0}^{m-1} a_i x^i + a'_m x^m \pmod{x^{m+1}},$$

for some a'_m. If $\delta_m^{\mathcal{G}}(\mathbf{a}) \geq B$, then we choose $a_m = a'_m$, so that $\lambda_{m+1}^{\mathcal{G}}(\mathbf{a}) = \lambda_m^{\mathcal{G}}(\mathbf{a})$. Otherwise we choose a_m to be any value other than a'_m. We claim that in the second case $\lambda_{m+1}^{\mathcal{G}}(\mathbf{a}) \neq \lambda_m^{\mathcal{G}}(\mathbf{a})$.

Suppose to the contrary that $\lambda_{m+1}^{\mathcal{G}}(\mathbf{a}) = \lambda_m^{\mathcal{G}}(\mathbf{a})$. Let f' and q' be polynomials satsifying

$$\frac{f'}{q'} \equiv \sum_{i=0}^{m} a_i x^i \pmod{x^{m+1}}.$$

Then

$$\frac{f}{q} \equiv \frac{f'}{q'} + (a'_m - a_m) x^m \pmod{x^{m+1}},$$

so that $\deg(fq' - f'q) \geq m$. But $\delta_m^{\mathcal{G}}(\mathbf{a}) < B$ implies that $\lambda_m^{\mathcal{G}}(\mathbf{a}) < mB \leq m/2$, so that $\deg(fq' - f'q) < m$, which is a contradiction. In this case we also let $n_k = m$.

It follows that

$$\delta_{n_k+1}^{\mathcal{G}}(\mathbf{a}) = 1 - \frac{n_k}{n_k + 1} \delta_{n_k}^{\mathcal{G}}(\mathbf{a}) > \delta_{n_k}^{\mathcal{G}}(\mathbf{a}).$$

Therefore equations (21.3) and (21.4) hold for all i. Hence

$$B = \lim_{i \to \infty} \delta_{n_i}^{\mathcal{G}}(\mathbf{a})$$

is the least accumulation point. This proves the theorem. □

Note that if we modify this construction so that the linear complexity changes when $\delta_m^{\mathcal{G}}(\mathbf{a}) < \delta_{m-1}^{\mathcal{G}}(\mathbf{a}) < B < \delta_{m-2}^{\mathcal{G}}(\mathbf{a})$, then B is still the least accumulation point. In fact, at each phase we can either use this method or the one in the proof to determine when to change the linear complexity. Since there are infinitely many phases, this gives uncountably many sequences for which B is the least accumulation point.

Corollary 21.4.3. *Every balanced interval $[B, 1 - B]$ occurs uncountably often as a $\Upsilon(\mathbf{a})$ with $\mathbf{a}$ a sequence over F.*

21.5 Asymptotic N-adic complexity

The goal of this section is to do a similar analysis for N-adic complexity, where N is an integer greater than 1. In fact we show that if N is a power of 2 or 3, then exactly the same theorem holds. For more general N we obtain weaker results.

In this section $\mathcal{G}$ is the set of FCSRs with entries in $\Sigma = \{0, 1, \cdots, N-1\}$. We have $\Phi_N(f, q) = \max(|f|, |q|)$ and we take $\lambda_N(f, q) = \log(\Phi_N(f, q))$ as the size of the FCSR with connection integer q that outputs the N-adic expansion of f/q. For an N-ary sequence $\mathbf{a} = a_0, a_1, \cdots, a_i \in \{0, 1, \cdots, N-1\}$, we let $\Phi_n^{\mathcal{G}}(\mathbf{a})$ denote the least $\Phi_N(f, q)$ so that

$$\frac{f}{q} \equiv \sum_{i=0}^{\infty} a_i N^i \pmod{N^n}$$

and let $\lambda_n^{\mathcal{G}}(\mathbf{a}) = \log_N(\Phi_n(\mathbf{a}))$. Let $\delta_n^{\mathcal{G}}(\mathbf{a}) = \lambda_n^{\mathcal{G}}(\mathbf{a})/n$. Also, we let $\Phi^{\mathcal{G}}(\mathbf{a})$ denote the least $\Phi_N(f, q)$ so that

$$\frac{f}{q} = \sum_{i=0}^{\infty} a_i N^i,$$

if such a pair f, q exists. We let $\Phi^{\mathcal{G}}(\mathbf{a}) = \infty$ otherwise. Let $\lambda^{\mathcal{G}}(\mathbf{a}) = \log_N(\Phi^{\mathcal{G}}(\mathbf{a}))$, the *$N$-adic complexity of* $\mathbf{a}$. The set of numbers $\{\lambda_n^{\mathcal{G}}(\mathbf{a}) : n = 1, 2, \cdots\}$ is the *N-adic complexity profile of* $\mathbf{a}$. Note that $\Phi_n^{\mathcal{G}}(\mathbf{a}) \leq N^n - 1$ since we can take

$$f = \sum_{i=0}^{n-1} a_i N^i, \quad q = 1.$$

Thus $\lambda_n^{\mathcal{G}}(\mathbf{a}) < n$ and $0 \leq \delta_n^{\mathcal{G}}(\mathbf{a}) < 1$.

Our goal is to show that the set $\Upsilon(\mathbf{a})$ of accumulation points of the normalized N-adic complexity profile of a sequence $\mathbf{a}$ that is not eventually periodic is a closed interval centered at $1/2$. In fact we see that this holds when $N \leq 4$, and in general when the least accumulation point is at most $\log_N(2)$. If the least accumulation point is greater than $\log_N(2)$, then it is not known what can happen. The analysis is somewhat more complicated than in the case of LFSRs since we do not have as precise a description of what can happen when the complexity changes from one length to the next.

21.5.a A Useful lemma

In this section we develop some technical tools needed analyze the set of accumulation points of the N-adic complexity of a sequence.

Our goal is to understand the asymptotic behavior of $\delta_n^{\mathcal{G}}(\mathbf{a})$. If $\mathbf{a}$ is eventually periodic, then $\Phi_n^{\mathcal{G}}(\mathbf{a})$ is constant for n sufficiently large, so the limit of the $\delta_n^{\mathcal{G}}(\mathbf{a})$ exists and is zero. Therefore from here on in this section we assume that $\mathbf{a}$ is not eventually periodic.

We need a lemma to bound $\delta_{n+1}^{\mathcal{G}}(\mathbf{a})$ in terms of $\delta_n^{\mathcal{G}}(\mathbf{a})$.

Lemma 21.5.1. *Suppose that* $\Phi_{n+1}^{\mathcal{G}}(\mathbf{a}) > \Phi_n^{\mathcal{G}}(\mathbf{a})$. *Then*

1. *For all* $N \geq 2$ *we have*
$$\frac{N^n}{2\Phi_n^{\mathcal{G}}(\mathbf{a})} \leq \Phi_{n+1}^{\mathcal{G}}(\mathbf{a}) \leq \frac{N\Phi_n^{\mathcal{G}}(\mathbf{a})}{2} + \frac{N^{n+1}}{2\Phi_n^{\mathcal{G}}(\mathbf{a})}.$$
2. *For all* $N \geq 2$ *we have*
$$\frac{n-1}{n+1} - \frac{n}{n+1}\delta_n^{\mathcal{G}}(\mathbf{a}) \leq \delta_{n+1}^{\mathcal{G}}(\mathbf{a}).$$
3. *For all* $N \geq 2$ *and* $\epsilon > 0$, *if* n *is sufficiently large and* $\delta_n^{\mathcal{G}}(\mathbf{a}) > \max(1/2+\epsilon, 1-\log_N(2/(1+\epsilon)))$, *then* $\delta_{n+1}^{\mathcal{G}}(\mathbf{a}) < \delta_n^{\mathcal{G}}(\mathbf{a})$.

Proof. Let
$$\frac{f}{q} \equiv \sum_{i=0}^{\infty} a_i N^i \pmod{N^n} \tag{21.5}$$
with $\gcd(q, N) = 1$ and $\Phi(f, q) = \Phi_n^{\mathcal{G}}(\mathbf{a})$. By the assumption that $\Phi_{n+1}^{\mathcal{G}}(\mathbf{a}) > \Phi_n^{\mathcal{G}}(\mathbf{a})$, equation (21.5) does not hold modulo N^{n+1}. Thus for some integer v with v not divisible by N we have
$$\frac{f}{q} \equiv vN^n + \sum_{i=0}^{\infty} a_i N^i \pmod{N^{n+1}}.$$
Suppose also that
$$\frac{g}{r} \equiv \sum_{i=0}^{\infty} a_i N^i \pmod{N^{n+1}}$$
with $\gcd(r, N) = 1$ and $\Phi(g, r) = \Phi_{n+1}^{\mathcal{G}}(\mathbf{a})$. Then
$$\frac{f}{q} \equiv \frac{g}{r} + vN^n \pmod{N^{n+1}}.$$
It follows that $fr - qg \equiv qrvN^n \pmod{N^{n+1}}$. Since v is nonzero, we have
$$N^n \leq |fr - qg| \leq 2\Phi_n^{\mathcal{G}}(\mathbf{a})\Phi_{n+1}^{\mathcal{G}}(\mathbf{a}).$$

This implies the lower bound in the first assertion. The lower bound on $\delta^{\mathcal{G}}_{n+1}(\mathbf{a})$ in the second assertion follows by taking logarithms and dividing by $n+1$.

To obtain an upper bound on $\Phi^{\mathcal{G}}_{n+1}(\mathbf{a})$ we construct a "pretty good" rational approximation modulo N^{n+1}. Then $\Phi^{\mathcal{G}}_{n+1}(\mathbf{a})$ is upper bounded by the value of Φ on this approximation. Note that f and q are relatively prime: if they weren't, then we could factor out a common factor and reduce $\Phi(f,q)$.

First assume that $0 < |q| < |f|$. Let u be an integer so that $u \equiv qv \pmod{N}$ and $|u| \leq N/2$. Since fN and q are relatively prime, there exist integers h and s so that

$$fNs - qh = uN^n - f. \tag{21.6}$$

Let $t = 1 + Ns$, so that $t \equiv 1 \pmod{N}$ and

$$ft - qh = uN^n. \tag{21.7}$$

It then follows that $ft - qh = uN^n \equiv qvtN^n \pmod{N^{n+1}}$, so that

$$\frac{f}{q} - \frac{h}{t} \equiv vN^n \pmod{N^{n+1}}.$$

Thus

$$\frac{h}{t} \equiv \sum_{i=0}^{\infty} a_i N^i \pmod{N^{n+1}},$$

and $\gcd(g, N) = 1$, so $\Phi(h,t)$ is an upper bound for $\Phi^{\mathcal{G}}_{n+1}(\mathbf{a})$. In fact there are many choices for (h,t) satisfying equation (21.6). By the relative primality of fN and q, the solutions to equation (21.6) with $t \equiv 1 \pmod{N}$ are exactly the pairs $(h,t) = (h_0,t_0) + (zfN, zqN)$ where (h_0,t_0) is a fixed solution and $z \in \mathbb{Z}$. In particular, we can take $|h| \leq |f|N/2 = N\Phi^{\mathcal{G}}_n(\mathbf{a})/2$. We then have $ft = qh + uN^n$ so that

$$\begin{aligned} |t| &\leq \frac{|qh|}{|f|} + \frac{|u|N^n}{|f|} \\ &\leq \frac{N|q|}{2} + \frac{|u|N^n}{|f|} \\ &< \frac{N\Phi^{\mathcal{G}}_n(\mathbf{a})}{2} + \frac{N^{n+1}}{2\Phi^{\mathcal{G}}_n(\mathbf{a})}. \end{aligned}$$

Therefore

$$\Phi^{\mathcal{G}}_{n+1}(\mathbf{a}) \leq \frac{N\Phi^{\mathcal{G}}_n(\mathbf{a})}{2} + \frac{N^{n+1}}{2\Phi^{\mathcal{G}}_n(\mathbf{a})}. \tag{21.8}$$

This proves the upper bound in the first assertion when $|q| < |f|$.

Now let $0 < |f| < |q|$. As in the previous case there are integers $t = 1 + Ns$ and h with $ft - qh = uN^n$. By adding a multiple of (qN, fN) to the pair (t, h), we may assume that $|t| \leq |q|N/2$. It follows that

$$|h| < \frac{N\Phi_n^{\mathcal{G}}(\mathbf{a})}{2} + \frac{N^{n+1}}{2\Phi_n^{\mathcal{G}}(\mathbf{a})}.$$

Finally we prove the third assertion. Let $\epsilon > 0$ and suppose that $\delta_n^{\mathcal{G}}(\mathbf{a}) > \max(1/2 + \epsilon, 1 - \log_N(2/(1+\epsilon)))$. Take n large enough that

$$\frac{1}{2} + \epsilon > \frac{1}{2} + \frac{\log(\epsilon^{-1})}{2n}. \tag{21.9}$$

From $\delta_n^{\mathcal{G}}(\mathbf{a}) > (1/2) + \epsilon$ and equation (21.9) it then follows that

$$\frac{N^{n+1}}{2\Phi_n^{\mathcal{G}}(\mathbf{a})} < \epsilon \frac{N\Phi_n^{\mathcal{G}}(\mathbf{a})}{2}.$$

Also, from $\delta_n^{\mathcal{G}}(\mathbf{a}) > 1 - \log_N(2/(1+\epsilon))$ we have

$$(1+\epsilon)\frac{N\Phi_n^{\mathcal{G}}(\mathbf{a})}{2} < \Phi_n^{\mathcal{G}}(\mathbf{a})^{(n+1)/n}.$$

Therefore

$$\begin{aligned} \Phi_{n+1}^{\mathcal{G}}(\mathbf{a}) &\leq \frac{N\Phi_n^{\mathcal{G}}(\mathbf{a})}{2} + \frac{N^{n+1}}{2\Phi_n^{\mathcal{G}}(\mathbf{a})} \\ &\leq (1+\epsilon)\frac{N\Phi_n^{\mathcal{G}}(\mathbf{a})}{2} \\ &< \Phi_n^{\mathcal{G}}(\mathbf{a})^{(n+1)/n}. \end{aligned}$$

Taking logarithms and dividing by $n+1$ then gives $\delta_{n+1}^{\mathcal{G}}(\mathbf{a}) < \delta_n^{\mathcal{G}}(\mathbf{a})$ as desired. □

This will suffice to characterize sets $[B, C]$ of accumulation points of normalized N-adic complexities of sequences when

$$1 - B \geq \lim_{\epsilon \to 0} \max\left(\frac{1}{2} + \epsilon, 1 - \log_N\left(\frac{2}{1+\epsilon}\right)\right) = \max\left(\frac{1}{2}, 1 - \log_N(2)\right).$$

This is equivalent to having $B \leq \min(1/2, \log_N(2))$. If $N \leq 4$, then $1 - \log_N(2) \leq 1/2$, so this suffices to characterize all sets of accumulation points. For larger N the characterization is incomplete.

21.5.b Sets of accumulation points

Let $\mathbf{a}$ be a N-ary sequence that is not eventually periodic. In this section we show that in many cases the set of accumulation points $\Upsilon(\mathbf{a})$ satisfies $\Upsilon(\mathbf{a}) = [B, 1-B]$ for some B. Let $\Upsilon(\mathbf{a}) = [B, C]$. Let $m_1, m_2, \cdots$ be a sequence of indices such that $B = \lim_{n\to\infty} \delta^{\mathcal{G}}_{m_n}(\mathbf{a})$. If $\lambda^{\mathcal{G}}_{n+1}(\mathbf{a}) = \lambda^{\mathcal{G}}_n(\mathbf{a})$, then $\delta^{\mathcal{G}}_{n+1}(\mathbf{a}) < \delta^{\mathcal{G}}_n(\mathbf{a})$. If we replace each m_n by the next index $j \geq m_n$ so that $\lambda^{\mathcal{G}}_j(\mathbf{a}) < \lambda^{\mathcal{G}}_{j+1}(\mathbf{a})$, then any accumulation point D of the resulting $\delta^{\mathcal{G}}_{m_n}(\mathbf{a})$s will satisfy $D \leq B$. Since B is the minimal accumulation point of the $\delta^{\mathcal{G}}_i(\mathbf{a})$, $D = B$. Therefore we may assume that $\lambda^{\mathcal{G}}_{m_n}(\mathbf{a}) < \lambda^{\mathcal{G}}_{m_n+1}(\mathbf{a})$.

Lemma 21.5.2. *Let $N \geq 2$ and $B < 1/2$. Then*

$$\lim_{n\to\infty} \delta^{\mathcal{G}}_{m_n+1} = 1 - B.$$

Proof. Let $\epsilon > 0$. Take n large enough that $m_n \geq 4/\epsilon$ and $|B - \delta^{\mathcal{G}}_{m_n}(\mathbf{a})| < \min(\epsilon/2, (1-2B)/4)$. Then $\delta^{\mathcal{G}}_{m_n}(\mathbf{a}) < 1/2$ and by Lemma 21.5.1.2

$$\begin{aligned} 1 - B - \delta^{\mathcal{G}}_{m_n+1}(\mathbf{a}) &\leq 1 - B - \left(\frac{m_n - 1}{m_n + 1} - \frac{m_n}{m_n + 1}\delta^{\mathcal{G}}_{m_n}(\mathbf{a})\right) \\ &= (\delta^{\mathcal{G}}_{m_n}(\mathbf{a}) - B) + \frac{2}{m_n + 1} \\ &\leq \epsilon. \end{aligned}$$

Also, $\delta^{\mathcal{G}}_{m_n}(\mathbf{a}) < 1/2$ implies that $\Phi^{\mathcal{G}}_{m_n}(\mathbf{a}) < N^{m_n}/\Phi^{\mathcal{G}}_{m_n}(\mathbf{a})$, so by Lemma 21.5.1.1 we have $\Phi^{\mathcal{G}}_{m_n+1}(\mathbf{a}) < N^{m_n+1}/\Phi^{\mathcal{G}}_{m_n}(\mathbf{a})$. Thus $\lambda^{\mathcal{G}}_{m_n+1}(\mathbf{a}) < m_n + 1 - \lambda^{\mathcal{G}}_{m_n}(\mathbf{a})$, so

$$\begin{aligned} \delta^{\mathcal{G}}_{m_n+1}(\mathbf{a}) - (1 - B) &\leq \frac{m_n + 1}{m_n + 1} - \frac{m_n}{m_n + 1}\delta^{\mathcal{G}}_{m_n}(\mathbf{a}) - (1 - B) \\ &< (B - \delta^{\mathcal{G}}_{m_n}(\mathbf{a})) + \frac{1}{m_n + 1}\delta^{\mathcal{G}}_{m_n}(\mathbf{a}) \\ &\leq \epsilon. \end{aligned}$$

Thus $|1 - B - \delta^{\mathcal{G}}_{m_n+1}(\mathbf{a})| < \epsilon$ for n sufficiently large, proving the lemma. □

Corollary 21.5.3. *In general $1/2 \leq C$.*

Proof. By Lemma 21.5.2, if $B < 1/2$, then $1 - B > 1/2$ is an accumulation point. If $B \geq 1/2$, then $C \geq B \geq 1/2$. In either case $C \geq 1/2$. □

We can now prove our main result on sets of accumulation points of the normalized N-adic complexity.

Theorem 21.5.4. *Let* $\mathbf{a}$ *be an* N*-ary sequence and suppose that the set of accumulation points of the set of* $\delta_n^{\mathcal{G}}(\mathbf{a})$ *is the interval* $[B, C]$*. Then* $B \leq \max(1/2, 1 - \log_N(2))$*. If* $B < \log_N(2)$ *then* $C = 1 - B$*.*

Proof. There is a sequence of integers $\ell_1 < \ell_2 < \cdots$ such that

$$\lim_{n\to\infty} \delta_{\ell_n}^{\mathcal{G}}(\mathbf{a}) = C$$

and we can assume that

$$|C - \delta_{\ell_n}^{\mathcal{G}}(\mathbf{a})| > |C - \delta_{\ell_{n+1}}^{\mathcal{G}}(\mathbf{a})|$$

for all n. By possibly deleting some of the ℓ_n and m_n, we can assume that $m_n < \ell_n < m_{n+1}$ for all $n \geq 1$. For n sufficiently large we have $\delta_{m_n}^{\mathcal{G}} < \delta_{\ell_n}^{\mathcal{G}}$, so we can assume this holds for all $n \geq 1$. Thus there is an $\ell \leq \ell_n$ so that $\delta_{\ell-1}^{\mathcal{G}}(\mathbf{a}) < \delta_{\ell}^{\mathcal{G}}(\mathbf{a})$. If we replace ℓ_n by the largest such ℓ, then we still have a sequence whose limit is C. So we can assume that $\delta_{\ell_n-1}^{\mathcal{G}}(\mathbf{a}) < \delta_{\ell_n}^{\mathcal{G}}(\mathbf{a})$ for all n. In particular, $\Phi_{\ell_n-1}^{\mathcal{G}}(\mathbf{a}) < \Phi_{\ell_n}^{\mathcal{G}}(\mathbf{a})$. Then by Lemma 21.5.1.3, for every $\epsilon > 0$ there is an n so that

$$\delta_{\ell_n-1}^{\mathcal{G}}(\mathbf{a}) < \max(1/2 + \epsilon, 1 - \log_N(2/(1+\epsilon))).$$

This implies that there is an accumulation point of the $\delta_n^{\mathcal{G}}(\mathbf{a})$ that is less than or equal to $\max(1/2, 1 - \log_N(2))$, so

$$B \leq \max(1/2, 1 - \log_N(2)) = 1 - \min(1/2, \log_N(2)).$$

This proves the first statement.

To prove the second statement, let us assume to the contrary that $1 - B < C$ and that $B \leq \log_N(2)$. Thus $C > 1 - \log_N(2)$. Also, $B \leq 1/2$ (since when $N = 2$ or 3, $\max(1/2, 1 - \log_N(2)) = 1/2$ and when $N \geq 4$, $\log_N(2) \leq 1/2$), so $1/2 \leq 1 - B < C$. By Lemma 21.5.1.1,

$$\Phi_{n+1}^{\mathcal{G}}(\mathbf{a}) \leq \frac{N\Phi_n^{\mathcal{G}}(\mathbf{a})}{2} + \frac{N^{n+1}}{2\Phi_n^{\mathcal{G}}(\mathbf{a})} \leq \max(N\Phi_n^{\mathcal{G}}(\mathbf{a}), N^{n+1}/\Phi_n^{\mathcal{G}}(\mathbf{a})).$$

Thus

$$\delta_{\ell_n}^{\mathcal{G}}(\mathbf{a}) \leq \max\left(\frac{1}{\ell_n} + \frac{\ell_n - 1}{\ell_n}\delta_{\ell_n-1}^{\mathcal{G}}(\mathbf{a}), 1 - \frac{\ell_n - 1}{\ell_n}\delta_{\ell_n-1}^{\mathcal{G}}(\mathbf{a})\right). \tag{21.10}$$

Suppose that $\delta_{\ell_n-1}^{\mathcal{G}}(\mathbf{a}) \leq 1/2$. Then the right hand side of equation (21.10) equals the second term, so

$$\delta_{\ell_n-1}^{\mathcal{G}}(\mathbf{a}) \leq \frac{\ell_n}{\ell_n - 1} - \frac{\ell_n}{\ell_n - 1}\delta_{\ell_n}^{\mathcal{G}}(\mathbf{a}).$$

If this occurs for infinitely many n, then the set $\{\delta^{\mathcal{G}}_{\ell_n-1}(\mathbf{a}) : n \geq 1\}$ has an accumulation point less than or equal to

$$\lim_{n\to\infty} \frac{\ell_n}{\ell_n - 1} - \frac{\ell_n}{\ell_n - 1}\delta^{\mathcal{G}}_{\ell_n}(\mathbf{a}) = 1 - \lim_{n\to\infty} \delta^{\mathcal{G}}_{\ell_n}(\mathbf{a}) = 1 - C < B.$$

This is a contradiction, so (by possibly deleting finitely many ℓ_ns) we may assume that $\delta^{\mathcal{G}}_{\ell_n-1}(\mathbf{a}) > 1/2$ for every n. Thus the right hand side of equation (21.10) equals the first term, and

$$C = \lim_{n\to\infty} \delta^{\mathcal{G}}_{\ell_n}(\mathbf{a}) \leq \lim_{n\to\infty} \frac{1}{\ell_n} + \frac{\ell_n - 1}{\ell_n}\delta^{\mathcal{G}}_{\ell_n-1}(\mathbf{a}) = \lim_{n\to\infty} \delta^{\mathcal{G}}_{\ell_n-1}(\mathbf{a}).$$

But C is the maximum accumulation point of the $\delta^{\mathcal{G}}_i(\mathbf{a})$, so in fact $\lim_{n\to\infty} \delta^{\mathcal{G}}_{\ell_n-1}(\mathbf{a}) = C$.

Again using Lemma 21.5.1.3 and taking limits, we see that $C \leq \max(1/2, 1 - \log_N(2))$, which is a contradiction. □

Corollary 21.5.5. *Let $N = 2$, 3, or 4. Let $\mathbf{a}$ be an eventually non-periodic N-ary sequence. Then $\Upsilon(\mathbf{a}) = [B, 1 - B]$ for some real number B.*

Proof. For these values of N we have $\max(1/2, 1 - \log_N(2)) = 1/2$ and $\log_N(2) \geq 1/2$, so the first assertion of Theorem 21.5.4 says that $B \leq 1/2$ for all such $\mathbf{a}$ and the second assertion then says $C = 1 - B$ for all such $\mathbf{a}$. □

Now fix a positive integer k. Consider a sequence $\mathbf{a} = a_0, a_1, \cdots$ with each $a_i \in \{0, 1, \cdots, N-1\}$. For each i, let

$$b_i = \sum_{j=0}^{k-1} a_{ki+j} N^j \in \{0, 1, \cdots, N^k - 1\}$$

and let $\mathbf{b} = b_0, b_1, \cdots$. Then the function $\Gamma : \mathbb{Z}_N \to \mathbb{Z}_{N^k}$ defined by

$$\Gamma\left(\sum_{i=0}^{\infty} a_i N^i\right) = \sum_{i=0}^{\infty} b_i (N^k)^i$$

is a ring isomorphism.

Theorem 21.5.6. *Let $\mathbf{a} = a_0, a_1, \cdots$ with each $a_i \in \{0, 1, \cdots, N-1\}$. Then the set of accumulation points of $\{\delta^{\mathcal{G}}_n(\mathbf{a})\}$ is identical to the set of accumulation points of $\{\delta^{\mathcal{G}}_n(\Gamma(\mathbf{a}))\}$ (where we use N^k-adic complexity to define $\delta^{\mathcal{G}}_n(\Gamma(\mathbf{a}))$).*

Proof. First observe that if R is a ring, then there is a unique ring homomorphism from $\mathbb{Z}$ into R, defined by mapping 1 to 1. This is called the *canonical map* from $\mathbb{Z}$ to R. If U is any subring of the rational numbers, then there is at most one ring homomorphism from U to R, since any such homomorphism is still determined by mapping 1 to 1. In fact such a homomorphism exists if and only if the image of every invertible integer in U under the canonical map is a unit in R.

When $R = \mathbb{Z}_N$, the maximal subring of the rational numbers that maps to R is $U = \{f/q : \gcd(N, q) = 1\}$. In this case the map is an injection. Since $\gcd(N, q) = 1$ if and only if $\gcd(N^k, q) = 1$, U is also the maximal subring of the rational numbers that maps to $\mathbb{Z}_{N^k}$. By the uniqueness of these maps, they must be identified under the identification of $\mathbb{Z}_N$ and $\mathbb{Z}_{N^k}$.

Now suppose that D is an accumulation point of the $\delta_n^{\mathcal{G}}(\mathbf{a})$, say

$$D = \lim_{i \to \infty} \delta_{n_i}^{\mathcal{G}}(\mathbf{a}).$$

There is some $j \in \{0, \cdots, k-1\}$ so that $\{n_i \equiv j \pmod{k}\}$ is infinite. Thus D is a limit of $\delta_n^{\mathcal{G}}(\mathbf{a})$s with all n congruent to j.

Let $a = \sum_{i=0}^{\infty} a_i N^i$. Suppose that k divides n. Then

$$a \pmod{N^n} = \Gamma(a) \pmod{(N^k)^{n/k}}.$$

Thus $\Phi_n^{\mathcal{G}}(\mathbf{a})$ — the minimal $\Phi(a, b)$ among rational approximations of a modulo N^n — is the same as $\Phi_{n/k}^{\mathcal{G}}(\Gamma(\mathbf{a}))$ — the minimal $\Phi(a, b)$ among rational approximations of $\Gamma(a)$ modulo $(N^k)^{n/k} = N^n$. Thus $\lambda_n^{\mathcal{G}}(\mathbf{a}) = k\lambda_{n/k}^{\mathcal{G}}(\Gamma(\mathbf{a}))$, and

$$\delta_n^{\mathcal{G}}(\mathbf{a}) = \frac{\lambda_n^{\mathcal{G}}(\mathbf{a})}{n} = \frac{k\lambda_{n/k}^{\mathcal{G}}(\Gamma(\mathbf{a}))}{n} = \frac{\lambda_{n/k}^{\mathcal{G}}(\Gamma(\mathbf{a}))}{n/k} = \delta_{n/k}^{\mathcal{G}}(\Gamma(\mathbf{a})).$$

Thus the set of accumulation points of the $\delta_n^{\mathcal{G}}(\mathbf{a})$ that are limits of $\delta_n^{\mathcal{G}}(\mathbf{a})$s with k dividing n coincides exactly with the set of accumulation points of the $\delta_n^{\mathcal{G}}(\Gamma(\mathbf{a}))$. Thus it remains only to show that if $1 \leq j \leq k-1$, then every accumulation point of $\{\delta_n^{\mathcal{G}}(\mathbf{a}) : n \equiv j \pmod{k}\}$ is also an accumulation point of $\{\delta_n^{\mathcal{G}}(\Gamma(\mathbf{a}))\}$. For the remaining details we refer the reader to [98].

□

Corollary 21.5.7. *Let N be a power of 2 or 3. Let $\mathbf{a}$ be an eventually non-periodic N-ary sequence. Then $\Upsilon(\mathbf{a}) = [B, 1-B]$ for some real number B.*

Proof. The proof is immediate from Corollary 21.5.5 and Theorem 21.5.6. □

Finally, we mention that for every $N \geq 2$ and every $B \in [0, 1/2]$ there are uncountably many N-ary sequences $\mathbf{a}$ with $\Upsilon(\mathbf{a}) = [B, 1-B]$. The proof is similar to the proof of Theorem 21.4.2 and Corollary 21.4.3. See [98] for details.

21.6 Consequences and questions

Suppose a stream cipher uses an infinite non-periodic sequence $\mathbf{a}$ as a keystream, and suppose that the set of accumulation points of the normalized N-adic or linear complexity is $[B, 1-B]$. Now imagine a cryptanalyst who has observed a prefix of $\mathbf{a}$ and wants to predict the next symbol. If the normalized complexity up to this point is close to B, the next symbol is likely to change the complexity so the normalized complexity increases. Likewise, if the normalized complexity up to this point is close to $1-B$, then the next symbol is likely to leave the complexity unchanged so the normalized complexity decreases. In this sense sequences for which the set of accumulation points is $[0, 1]$ are the most random sequences.

Various questions remain. If N is not a power of 2 or 3, and $\beta(\mathbf{a}) > \log_N(2)$, is $\Upsilon(\mathbf{a}) = [\beta(\mathbf{a}), 1-\beta(\mathbf{a})]$? Is there any (perhaps measure theoretic) sense in which some least accumulation points are more likely than others? In the case of linear complexity over a field $\mathbb{F}_q$, Niederreiter showed that $\Upsilon(\mathbf{a}) = [1/2, 1/2]$ with probability 1 (where the set of infinite sequences is endowed with the infinite product measure arising from the uniform measure on $\mathbb{F}_q$) [150]. We do not have such a result for N-adic complexity. For a fixed $\mathbf{a}$, what can be said about the distribution of the $\delta_n^{\mathcal{G}}(\mathbf{a})$ in $[B, 1-B]$? If the distribution is non-uniform, this might lead to better prediction methods than those outlined in the first paragraph of this section.

Finally, we have treated, in varying detail, a number of generalizations of feedback with carry shift registers and linear feedback shift registers [64, 95, 107]. Do the same results hold in these settings?

21.7 Exercises

1. Show that the bound in Corollary 21.1.5 can be improved to

$$E_n^{\text{lin,per}} > n - \frac{m + r^2 - r}{r^2 - 1}.$$

2. Simplify the expression for $E_n^{\text{lin,per}}$ in Theorem 21.1.4 in the case when $v = 0$.

3. Simplify the expression for $E_n^{\text{lin,per}}$ in Theorem 21.1.4 in the case when $m = 1$.

4. We can compute the expected linear span of periodic sequences without the use of GDFTs. This problem explores the direct approach. Let r be a power of the prime p. We consider periodic sequences over $\mathbb{F}_r$ with period $n = p^v m$, where $\gcd(m, p) = 1$

   a. Let $C_{k_1}, \cdots, C_{k_s}$ be the distinct cyclotomic cosets modulo m. Show that every divisor of $x^n - 1$ is of the form $\prod_{i=1}^{s} q_i(x)^{t_i}$ where $q_1, q_2, \cdots, q_s$ is a fixed set of irreducible polynomials with $\deg(q_i) = |C_{k_i}| = d_i$ and $0 \le t_i \le p^v$.

b. Recall that the number of periodic sequences with connection polynomial $q(x)$ is $N_q = r^{\deg(q)}$. Let M_q be the number of periodic sequences with minimal connection polynomial q. Use the inclusion/exclusion formula to show that if $q = \prod_{i=1}^{s} q_i(x)^{t_i}$, then

$$M_q = r^{\sum_{i=1}^{s} t_i d_i} \prod_{t_i > 0} (1 - \frac{1}{r^{d_i}}).$$

c. Use the results of part (b) to obtain Theorem 21.1.4.

5. Show that the bound in Theorem 21.2.1 is tight in the sense that there is an infinite family of connection integers q for which $\log_N(q) - E^{N\text{-adic,per}}(q) \in \Omega(\log\log(q))$.

6. Let $R = \mathbb{Z}[\pi]$ where $\pi^2 = -2$. Let $E = \mathbb{Q}[\pi]$ be the fraction field of R and let $N = \mathsf{N}_{\mathbb{Q}}^{E}$ be the norm. Let $\mathbf{a} = a_0, a_1, \cdots$ be a sequence over S. Let $\Phi_n(\mathbf{a})$ be the minimum $\max(N(u), N(q))$ with $u/q = \sum_{i=0}^{\infty} a_i \pi^i$. Let $\lambda_n(\mathbf{a}) = \log_2(\Phi_n(\mathbf{a}))$, let $\delta_n(\mathbf{a}) = \lambda_n(\mathbf{a})/n$, and let $\Upsilon = [B, C]$ denote the set of accumulation points of $\mathbf{a}$.

a. Prove that for any $a, b \in R$ we have
   i. $N(a) = 1$ if and only if a is a unit in R if and only if $a \in \{1, -1\}$.
   ii. $N(\pi) = p$.
   iii. $N(a+b) \leq N(a) + N(b) + 2(N(a)N(b))^{1/2} = (N(a)^{1/2} + N(b)^{1/2})^2 \leq 4\max(N(a), N(b))$.

b. Prove that for every $a, b \in R$ with $b \neq 0$, there exits $q, r \in R$ so that $a = qb + r$ and $N(r) < (3/4)N(b)$.

c. Suppose that $\Phi_{n+1}(\mathbf{a}) > \Phi_n(\mathbf{a})$. Prove that
   i.

$$\frac{2^{n-2}}{\Phi_n(\mathbf{a})} \leq \Phi_{n+1}(\mathbf{a}) \leq \frac{3\Phi_n(\mathbf{a})}{2} + \frac{2^n}{\Phi_n(\mathbf{a})} + \sqrt{6} \cdot 2^{n/2}.$$

   ii.

$$\frac{n-2}{n+1} - \frac{n}{n+1}\delta_n(\mathbf{a}) \leq \delta_{n+1}(\mathbf{a}).$$

   iii. For all $\epsilon > 0$, if n is sufficiently large and $\delta_n(\mathbf{a}) > \max(1/2 + \epsilon, \log_2(3(1+2\epsilon)/2))$, then $\delta_{n+1}(\mathbf{a}) < \delta_n(\mathbf{a})$.

d. Prove that $B \leq \log_2(3/2)$ and that if $B < \log_2(4/3)$, then $C = 1 - B$. That is, $\Upsilon(S) = [B, 1-B]$.

7. Let the notation be as in Exercise 6. Let $\mathbf{a}$ be a binary sequence generated by an AFSR D over R and π with connection element q. Let μ denote the number of bits required to represent the states of D in its infinite execution that outputs $\mathbf{a}$. More precisely, if the memory is represented as $\sum_{i=0}^{m-1} z_i \pi^i$ with $z_i \in \{0, 1\}$, then μ is the length of the AFSR plus the maximum m in the infinite execution. Let the rational representation of the π-adic number associated with $\mathbf{a}$ be u/q. Show that

$$|\mu - \log_2(\max(N(u), N(q)))| \in O(\log\log(\max(N(u), N(q)))).$$

8. Prove that for every $N \geq 2$ and every $B \in [0, 1/2]$ there are uncountably many N-ary sequences $\mathbf{a}$ with $\Upsilon(\mathbf{a}) = [B, 1 - B]$.

Bibliography

[1] F. Annexstein, Generating de Bruijn sequences: an efficient implementation, *IEEE Trans. on Computers* **46** (1997), 198–200. 279

[2] F. Arnault, T. Berger, C. Lauradoux, M. Minier, and B. Pousse, A New Approach to FCSRs, Cryptology ePrint Archive: Report 2009/167, http://eprint.iacr.org/2009/167. 268

[3] M. F. Atiyah and I. G. MacDonald, *Introduction to Commutative Algebra*, Addison–Wesley, Reading MA, 1969. 38

[4] E. Bach, Efficient prediction of Marsaglia–Zaman random number generators, (draft, Univ. of Wisconsin, 1993).

[5] J. T. Barrows, Jr., *A new method for constructing multiple error correcting linear residue codes*, **Rep. R–277**, Coordinated Sci. Lab., Univ. of Illinois, Urbana, 1966. 360

[6] E. R. Berlekamp, *Algebraic Coding Theory* ch. 7, McGraw–Hill, New York, 1968 395

[7] E. R. Berlekamp, *Algebraic Coding Theory*, second ed., Aegean Park Press, Laguna Hills Ca, 1984. 279

[8] J. Bernasconi and C. G. Günther, Analysis of a nonlinear feedforward logic for binary sequence generators, in E. Pichler, ed., *Advances in Cryptology – Eurocrypt 1985*, Lecture Notes in Computer Science **219**, Springer–Verlag, Berlin, 1986. 411

[9] T. Beth, and F. Piper, The stop–and–go generator, in T. Beth, N. Cot, and I. Ingemarsson, ed., *Advances in Cryptology – Eurocrypt '84*, Lecture Notes in Computer Science **209**, Springer–Verlag, Berlin, 1986, pp. 88–92. 414

[10] R. E. Blahut, *Theory and Practice of Error Control Codes*, Addison–Wesley, Reading MA, 1983. 402, 507

[11] S. R. Blackburn, A generalisation of the discrete Fourier transform: determining the minimal polynomial of a periodic sequence, *IEEE Trans. Info. Theory* **40** (1994), 1702–1704. 404

[12] S. R. Blackburn, A Note on sequences with the shift and add property. *Designs, Codes, and Crypt.* **9** (1996), pp. 251–256. 295, 298

[13] L. Blum, M. Blum, and M. Shub, A simple unpredictable pseudo–random number generator, *Siam J. Comput.* **15** (1986), 364–383. 360

[14] A. Blumer and J. Blumer, Linear size finite automata for the set of all subwords of a word: an outline of results, *Bulletin of the European Association for Theoretical Computer Science* **21** (1983), 68–77. 425

[15] E. Bombieri, personal communication. 361

[16] E. Bombieri and A. van der Poorten, Continued fractions of algebraic numbers, in W. Bosma and A. van der Poorten, ed., *Computational Algebra and Number Theory, Sydney, 1992*, Kluwer, 1995, pp. 137–152. 141

[17] Z. I. Borevich and I. R. Shefarevich, *Number Theory*, Academic Press, Orlando, Fl, 1966. 48, 62, 92, 216, 218, 447

[18] J. Bourgain, Personal correspondence, 2007.

[19] J. Bourgain, T. Cochrane, J. Paulhus, and C. Pinner, Decimations of ℓ-sequences and permutations of even residues mod p, *SIAM Journal on Discrete Mathematics* **23** (2009), 842–857 368

[20] N. G. de Bruijn, A Combinatorial Problem, *Koninklijke Nederlandse Akademie v. Wetenschappen* **49** (1946), 758–764. 277

[21] L. Bryniellson, On the linear complexity of combined shift registers, in E. Pichler, ed., *Advances in Cryptography – Eurocrypt '85*, Lecture Notes in Computer Science **219**, Springer–Verlag, Berlin, 1986, pp. 156–166. 409

[22] L. Carlitz and S. Uchiyama, Bounds for exponential sums, Duke Math J. **24** (1957), 37–41. 74

[23] A. Chan and R. Games, On the quadratic span of de Bruijn sequences, *IEEE Trans. Info. Theory* **36** (1990), 822–829.

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

[25] A. H. Chan, M. Goresky, and A. Klapper, Cross–correlations of quadratically decimated geometric sequences and GMW sequences, *Discr. Appl. Math.* **46** (1993), 1–20. 335

[26] U. Cheng, On the continued fraction and Berlekamp's algorithm, *IEEE Trans. Info. Theory* **30** (1984), 541–544. 395

[27] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer–Verlag, Berlin, 1993. 360

[28] R. Couture and P. L'Écuyer, On the lattice structure of certain linear congruential sequences related to AWC/SWB generators, *Math. Comp.* **62** (1994), 799–808. 16

[29] R. Couture and P. L'Écuyer, Distribution properties of multiply–with–carry random number generators, *Math. Comp.* **66** (1997), 591–607. 16, 192, 204, 205

[30] R. Coveyou and R. MacPherson, Fourier analysis of uniform random number generators, *J. ACM* **14** (1967), 100–119. 204

[31] T. W. Cusick, C. Ding, and A. Renvall, *Stream Ciphers and Number Theory*, Elsevier (North–Holland), Amsterdam, 1998.

[32] Z. Dai, Binary sequences derived from ML-sequences over rings I: Periods and minimal polynomials, J. Crypt. **5** (1992), 193–207. 341, 344

[33] Z. Dai, K. Imamura, and J. Yang, *Asymptotic behavior of normalized linear complexity of multi–sequences*, in T. Helleseth, D. Sarwate, H.-Y. Song, K. Yang, eds., *Sequences and Their Applications – SETA 2004*, Springer–Verlag Lecture Notes in Computer Science, **3486** (2005), pp. 126–142. 483

[34] Z. Dai, S. Jiang, K. Imamura, and G. Gong, *Asymptotic behavior of normalized linear complexity of ultimately non–periodic sequences*, *IEEE Trans. Info. Theory* **50** (2004), 2911–2915. 483

[35] Z. D. Dai and K. C. Zeng, Continued fractions and the Berlekamp–Massey algorithm, in J. Seberry, J. Pieprzyk, eds., *Advances in Cryptology – AUSCRYPT '90*, Lecture Notes in Computer Science **453**, Springer–Verlag, Berlin, 1990, pp. 24-31. 395

[36] P. Deligne, La conjecture de Weil, I, *Publ. Math. IHES* **43** (1974), 273–307. 75

[37] L. E. Dickson, *History of the Theory of Numbers, vol. 1,* Carnegie Inst., Washington D.C., 1919 (reprinted by Chelsea and published by American Mathematical Society). 15, 312

[38] J. Dillon, Elementary Hadamard difference sets, in: **Proceedings of the Sixth SE Conference Combinatorics Graph Theory and Computers, Winnipeg**, F. Hoffman et al. (Ed), Utilitas Math., 1975, 237–249. 317

[39] C. Ding, Algebraic construction of optimal frequency-hopping sequences. IEEE. Trans. Info. Theory **53** (2007), 2606–2610.

[40] C. Ding, R. Fuji-Hara, Y. Fujiwara, M. Jimbo and M. Mishima, Sets of frequency hopping sequences: bounds and optimal constructions. IEEE Trans. Info. Theory **55** (2009), 3297–3304. 293

[41] H. D. Ebbinghaus et al., *Numbers,* Graduate Texts in Mathematics **123**, Springer–Verlag, Berlin, 1990.

[42] P. Fan and M. Darnell, *Sequence Design for communications and applications*, John Wiley and Sons, New York, 1996. 408

[43] P. L'Écuyer, Maximally equidistributed combined Tausworthe generators, *Math. Comp.* **65** (1996), 8–30.

[44] B. Elspas, The theory of autonomous linear sequential networks, *IRE Trans. Circuit Theory* **CT–6** (1959), 45–60. 14

[45] D. Everett, Periodic digital sequences with pseudonoise properties, *G.E.C. Journal* **33** (1966),115–126. 14

[46] X. Feng and X. Dai, *The expected value of the normalized linear complexity of 2–dimensional binary sequences*, in T. Helleseth, D. Sarwate, H.-Y. Song, K. Yang, eds., *Sequences and Their Applications – SETA 2004*, Lecture Notes in Computer Science **3486**, Springer–Verlag, Berlin, 2005, pp. 113–128. 483

[47] J. P. Fillmore and M. L. Marx, Linear recursive sequences, *SIAM Rev.* **10** (1968), 342–353 173, 175

[48] J. von zur Gathen and J. Gerhard, *Modern Computer Algebra*, second edition, Cambridge Univ. Press, Cambridge, 2003 141

[49] C. F. Gauss, *Disquisitiones Arithmeticae*, 1801; reprinted in English translation by Yale Univ. Press, New Haven, CT., 1966. 360

[50] J. v. z. Gathen and J. Gerhard, *Modern Computer Algebra,* Cambridge University Press, Cambridge, U.K., 1999. 164

[51] R. Gold, Optimal binary sequences for spread spectrum multiplexing, *IEEE Trans. Info. Theory* **13** (1967), 619–621. 331

[52] S. Golomb, *Sequences with randomness properties*, Terminal Progress Report under Contract 639498, Glenn L. Martin Co., Baltimore, 1955. 14

[53] S. Golomb, *Shift Register Sequences.* Holden–Day, San Francisco, CA, 1967. 15, 213, 273, 312, 501

[54] S. Golomb, *Shift Register Sequences.* Aegean Park Press, Laguna Hills CA, 1982. (Reprint of [53].) 273, 283, 312

[55] S. Golomb and G. Gong, *Signal Design for Good Correlation* Cambridge Univ. Press, New York, 2005. 15, 408

[56] G. Gong, A. Di Porto, and W. Wolfowicz, Galois linear group sequences, *La Comm., Note Rec. Not.* **XLII**(1993), 83–89. 295, 298, 313, 354

[57] B. Gordon, W. H. Mills, and L. R. Welch, Some new difference sets, *Canad. J. Math.* **14** (1962), 614–625. 336

[58] M. Goresky and A. Klapper, Feedback registers based on ramified extensions of the 2–adic numbers, in A. De Santos, ed., *Advances in Cryptology – Eurocrypt '94*, Lecture Notes in Computer Science **950**, Springer–Verlag, Berlin, 1995, pp. 215-222. 192, 228, 240, 242

[59] M. Goresky and A. Klapper, Arithmetic Cross–Correlations of FCSR Sequences, *IEEE Trans. Info. Theory* **43** (1997) pp. 1342–1346. 368

[60] M. Goresky and A. Klapper, Periodicity, correlation, and distribution properties of d–FCSR sequences, Designs, Codes, and Cryptography 33 (2004), 123-148.

[61] M. Goresky and A. Klapper, Fibonacci and Galois Mode Implementation of Feedback with Carry Shift Registers, *IEEE Trans. Info. Theory* **48** (2002), 2826–2836. 240, 266

[62] M. Goresky and A. Klapper, Efficient multiply–with–carry random number generators with optimal distribution properties, *ACM TOMACS* **13** (2003), pp. 1–12. 192, 204, 208

[63] M. Goresky and A. Klapper: Periodicity and Correlations of of d–FCSR Sequences. Designs, Codes, and Crypt. **33** (2004), 123–148. 240, 243

[64] M. Goresky and A. Klapper, Polynomial pseudo–noise sequences based on algebraic shift register sequences. *IEEE Trans. Info. Theory* **53** (2006), 1649–1662. 354, 494

[65] M. Goresky, A. Klapper, and R. Murty, On the distinctness of decimations of ℓ-sequences, in T. Helleseth, P. V. Kumar, and K. Yang, eds., *Sequences and Their Applications – SETA '01,* Springer-Verlag, Berlin, 2002, pp. 197-208. 368

[66] M. Goresky, A. Klapper. R. Murty, and I. Shparlinski, On decimations of ℓ-sequences, *SIAM J. Disc. Math* **18** (2004), 130-140. 368

[67] D. Gollman and W. Chambers, Clock–controlled shift registers: a review. *IEEE J. Selected Areas Commun.* **7** (1989), 525–533. 414

[68] R. Göttfert and H. Niederreiter, Hasse–Teichmüller derivatives and products of linear recurring sequences, *Contemp. Math.* **168** (1994), 117–125. 403

[69] F. Q. Gouvêa, *p–Adic Numbers, an Introduction* Universitext, Springer–Verlag, Berlin, 2003.

[70] C. G. Günther, A finite field Fourier transform for vectors of arbitrary length, in *Communications and Cryptography: Two Sides of One Tapestry* (Blahut, Costello, Maurer and Mittelholzer, ed.), Kluwer Academic, 1994, Norwell MA, pp. 141–153. 403, 404

[71] R. T. Gregory and E. V. Krishnamurthy, *Methods and Applications of Error–Free Computation*, Springer–Verlag, N. Y., 1984. 431

[72] M. Hall, A survey of difference sets, Proc. Amer. Math. Soc. **7** (1956), 975–986. 15

[73] M. Hall, *Combinatorial Theory*, second ed., John Wiley & Sons, New York, 1986. 277

[74] G. H. Hardy and J. E. Littlewood, Some problems of 'Partitio Numerorum'; III: on the expression of a number as a sum of primes. *Acta Mathematica* **44** (1922), 1–70. 361

[75] G. Hardy and E. Wright, *An Introduction to the Theory of Numbers.* Oxford University Press, Oxford UK, 1979. 141, 145, 361

[76] H. Hasse, Theorie der höheren Differentiale in einem algebraischen Funktionenkörper mit vollkommenen Konstantenkörper bei beliebiger Charakteristik, *J. Reine u. Ang. Math.*, **. 175** (1936), 50–54. 403

[77] T. Herlestam, On function of linear shift register sequences, in E. Pichler, ed., *Advances in Cryptology – Eurocrypt '85*, Lecture Notes in Computer Science **219**, Springer–Verlag, Berlin, 1986, pp. 119–129. 412

[78] I.N. Herstein, *Topics in Algebra*, 2nd ed., 1975: Xerox College Publ., Lexington, MA.

[79] T. Honda, Isogeny classes of Abelian varieties over finite fields, J. Math. Soc. Japan, **20**(1968), 83–95. 326

[80] C. Hooley, On Artin's conjecture. *J. Reine Angew. Math.* **22** (1967), 209–220. 360, 368

[81] D. A. Huffman, The synthesis of linear sequential coding networks, Proc. 3rd London Symp. on Info. Theory, Butterworths, 1956, pp. 77–95. 14

[82] C. W. Hungerford, *Algebra*, Graduate Texts in Mathematics **73**, Springer–Verlag, Berlin,

[83] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, Springer–Verlag, N.Y., 1990. 34, 245

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

[85] N. Jacobson, *Basic Algebra II.* W.H. Freeman, San Francisco, 1980. 140

[86] C. J. A. Jansen, *Information theory of shift registers* in *Proceedings of the Tenth Symposium on Information Theory in the Benelux*, A. M. Barbe, ed., Werkgemeenschap voor Inf.– & Communicatietheorie, Enschede, Netherlands, 1989, pp. 153 –160. 425

[87] C. J. A. Jansen and D. E. Boekee, The shortest feedback shift register that can generate a given sequence, in G. Brassard, ed. *Advances in Cryptology – CRYPTO '89*, Lecture Notes in Computer Science **435**, Springer–Verlag, Berlin, 1990, pp. 90–99. 425

[88] C. J. A. Jansen and D. E. Boekee, On the significance of the directed acyclic word graph in cryptology, in J. Seberry and J. Pieprzyk, eds., *Advances in Cryptology–AUSCRYPT '90*, Lecture Notes in Computer Science **453**, Springer–Verlag, Berlin, 1990, pp. 318–326. 425

[89] E. Key, An Analysis of the structure and complexity of nonlinear binary sequence generators, *Trans. Info. Theory* **IT–22** (1976), 732–736. 409, 410, 413

[90] A. Y. Khinchin, *Continued Fractions* (Russian), 1935. English translation: University of Chicago Press, Chicago, 1961; reprinted by Dover Publications, Mineola New York, 1997. 141, 145

[91] A. Klapper, The Vulnerability of Geometric Sequences Based on Fields of Odd Characteristic, *J. Cryptology*, **7** (1994), 33–51. 412

[92] A. Klapper, Feedback with carry shift registers over finite fields, in B. Preneel, ed., *Fast Software Encryption. Second Internatinoal Workshop,* Lecture Notes in Computer Science **1008**, Springer–Verlag, Berlin, 1995, pp. 170–178.

[93] A. Klapper, d–form Sequences: Families of Sequences with Low Correlation Values and Large Linear Span, *IEEE Trans. Info. Theory* **41** (1995), 1–9. 338

[94] A. Klapper, Crosscorrelations of quadratic form sequences in odd characteristic, Designs, Codes, and Cryptography **11** (1997), 1–17. 331

[95] A. Klapper, Distributional properties of d–FCSR sequences, *J. Complexity* **20** (2004), 305–317. 494

[96] A. Klapper, Randomness and Register Synthesis for AFSRs based on Function Fields, *Sequences and Their Applications – SETA 2004*, Lecture Notes in Computer Science **3486**, Springer–Verlag, 2005, pp. 282–297. 220

[97] A. Klapper, Linear Complexity of Sequences under Different Interpretations, *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* **E89–A** (2006), 2254–2257.

[98] A. Klapper, The Asymptotic Behavior of 2–Adic Complexity, *Advances in Mathematics of Communications* **1** (2007), 307–319. 493

[99] A. Klapper, A. H. Chan and M. Goresky, Cross correlation of linearly and quadratically related geometric sequences and GMW sequences *Discr. Appl. Math.* **46** (1993), 1–20. 334, 335

[100] A. Klapper, A. H. Chan, and M. Goresky, Cascaded GMW Sequences, *Trans. Info. Theory* **39** (1993), 177–183. 336

[101] A. Klapper and M. Goresky, 2–adic shift registers, in R. Anderson, ed., *Fast Software Encryption 1993,* Lecture Notes in Computer Science **809**, Springer–Verlag, Berlin, 1994, pp. 174–178. 16

[102] Klapper, A. and Goresky, M. 1993. Feedback shift registers, combiners with memory, and arithmetic codes. *Univ. of Kentucky Dept. of Comp. Sci. Tech. Rep. No. 239–93.* 16, 192

[103] A. Klapper and M. Goresky, Large period nearly deBruijn FCSR sequences, in L. Guillou, J.-J. Quisquater, eds., *Advances in Cryptology – Eurocrypt 1995,* Lecture Notes in Computer Science **921**, Springer–Verlag, Berlin, 1995, pp. 263–273.

[104] A. Klapper and M. Goresky, Cryptanalysis based on 2–adic rational approximation, in D. Coppersmith, ed., *Advances in Cryptology – CRYPTO 95.* Lecture Notes in Computer Science **963**, Springer–Verlag, Berlin, 1995, pp. 262–273.

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

[106] A. Klapper and M. Goresky, Feedback Shift Registers, Combiners with Memory, and 2–Adic Span, *Journal of Cryptology* **10** (1997), pp. 111–147. 16, 192, 194, 196, 200, 201, 203, 206, 240

[107] A. Klapper and J. Xu, Algebraic feedback shift registers, *Theoretical Computer Science* **226** (1999), pp. 61–93. 205, 210, 213, 216, 217, 494

[108] A. Klapper and J. Xu, Feedback with carry shift registers over $\mathbf{Z}/(N)$, in C. Ding, T. Helleseth, and H. Niederreiter, eds., *Proceedings of International Conference on Sequences and their Application (SETA), Singapore, December 1998*, Springer-Verlag, 1999.

[109] A. Klapper and J. Xu: Register synthesis for algebraic feedback shift registers based on non–primes. Designs, Codes, and Crypt. **31** (2004), 227–25.

[110] C. Koç, Recurring with carry sequences, *J. Appl. Prob.* **32** (1995), 966–971 192

[111] D. Knuth, *The Art of Computer Programming, Vol 2. Seminumerical Algorithms,* Addison–Wesley, Reading MA, 1981. 190, 204, 360

[112] N. Koblitz, *p–Adic Numbers, p–Adic Analysis, and Zeta Functions.* Graduate Texts in Mathematics Vol. 58, Springer–Verlag, Berlin, 1984.

[113] E. V. Krishnamurthy and R. T. Gregory, Mapping integers and Hensel codes onto Farey fractions, *BIT* **23** (1983), 9–20. 431

[114] V. Kumar and R. Scholtz, Bounds on the linear span of bent sequences, *IEEE Trans. Info. Theory* **IT–29** (1983), 854–862. 317, 411

[115] V. Kumar, R. Scholtz and L. Welch, Generalized bent functions and their properties, *J. Comb. Theory A* **40** (1985), 90–107. 317

[116] V. Kurakin, A. Kuzmin, A. Mikhalev and A. Nechaev, Linear recurring sequences over rings and modules, *Journal of Mathematical Sciences* **76** (1995), pp. 2793–2915. 341, 344

[117] A. Kuzmin and A. Nechaev, Linear recurring sequences over Galois rings, Algebra and Logic **34** (1995), 87–100 (translated from Russian, Algebra i Logika, **34** (1995), 169–189). 341, 344

[118] G. Lachaud and J. Wolfmann, The weights of the orthogonals of the extended quadratic binary Goppa codes, IEEE Trans. Info. Theory **36** (1990), 686–692. 325

[119] S. Lang, *Algebra*, 2nd ed., Addison–Wesley, Reading, MA, 1984. 27, 62, 132

[120] A. Lempel, On a homomorphism of the de Bruijn graph and its applications to the design of feedback shift registers, *IEEE Trans. Computers* **19** (1970), 1204–1209. 279

[121] A. Lempel and H. Greenberger, Families of sequences with optimal Hamming correlation properties. IEEE. Trans. Info. Theory **20** (1974), 90–94. 289, 290, 292, 340, 341

[122] A. Lempel, M. Cohn, and W. Eastman, A class of balanced binary sequences with optimal autocorrelation properties, *IEEE Trans. Info. Theory* **23** (1977), 38–42. 292

[123] R. Lidl and H. Niederreiter, *Finite Fields*, 2nd ed., Cambridge University Press, Cambridge, UK, 1997. 63, 68, 74, 77, 78, 79, 403, 417

[124] F. J. MacWilliams, The structure and properties of binary cyclic alphabets, *Bell Systems Technica Journal* **44** (1965), 303–332. 14

[125] K. Mahler, On a geometrical representation of p–adic numbers, *Ann. of Math.* **41** (1940), 8–56. 431, 436

[126] D. Mandelbaum, Arithmetic codes with large distance, *IEEE Trans. Info. Theory,* **13** (1967), 237–242. 360

[127] D. Mandelbaum, A method for decoding generalized Goppa codes, *IEEE Trans. Info. Theory* **23** (1977), 137–140. 145

[128] D. Mandelbaum, An approach to an arithmetic analog of Berlekamp's algorithm, *IEEE Trans. Info. Theory* **30** (1984), 758–762. 431

[129] H. K. Markey and G. Antheil, **Secret communication systems**, U.S. Patent 2-292-387, Aug 11, 1942 (filed June 10, 1941). 289

[130] G. Marsaglia, The mathematics of random number generators, *The Unreasonable Effectiveness of Number Theory*, Amer. Math. Soc., Providence R. I., 1992, pp. 73–90. 16, 192

[131] G. Marsaglia, Yet another RNG, posted to Sci. Stat. Math, Aug 1 1994. 192, 205

[132] G. Marsaglia and A. Zaman, A new class of random number generators, *Annals of Applied Probability* **1** (1991), 462–480. 16, 192, 204, 205

[133] J.L. Massey, Shift register synthesis and BCH decoding, *IEEE Trans. Info. Theory* **IT–15** (1969), 122–127. 395

[134] J. L. Massey, Book Review of [10], *IEEE Trans. Info. Theory* **31** (1985), 553–554. 402

[135] J. L. 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. 441

[136] J. L. Massey and S. Serconek, A fourier transform approach to the linear complexity of nonlinearly filtered sequences, in Y. G. Desmedt, ed., *Advances in Cryptology–CRYPTO '94*, Lecture Notes in Computer Science **839**, Springer–Verlag, Berlin, 1994, pp. 332–340. 402, 403, 404, 409, 410

[137] J. L. Massey and S. Serconek, Linear complexity of periodic sequences: A general theory, in N. Koblitz, ed., *Advances in Cryptology–CRYPTO'96*, Lecture Notes in Computer Science **1109**, Springer–Verlag, Berlin, 1996, pp. 358–371. 403, 404, 405, 406

[138] P. Mathys, A generalization of the discrete fourier transform in finite fields, in *Proc. IEEE Symp. Info. Th.* 1990, San Diego, pp. 14–19. 404

[139] H. Matsumura, *Commutative Algebra*, W. A. Benjamin, New York, 1970. 96

[140] M. Matumoto and T. Nishimura, Mersenne twister: A 623–dimensionally equidistributed uniform pseudo–random number generator. *ACM Trans. Model. Computer Simulation* **8** (1998), 3–30.

[141] B. MacDonald, *Finite Rings with Identity*, Marcel Dekker, New York, 1974. 95, 96, 97, 103, 104

[142] R. McEliece, *Finite Fields for Computer Scientists and Engineers*, Kluwer Academic Publishers, Norwell, MA, 1987.

[143] R. J. McEliece, Linear recurring sequences over finite fields, Ph.D. thesis, California Inst. of Technology, 1967. http://etd.caltech.edu/etd/available/etd–10012002–154611/ 173

[144] R. McFarland, A family of difference sets in non-cyclic groups, J. Comb. Theory Ser. A, **15** (1976), 1–10. 317

[145] W. Meier and O. Staffelbach, Correlation properties of combiners with memory in stream ciphers, in I. Damgård, ed., *Advances in Cryptology – EUROCRYPT '90*, Lecture Notes in Computer Science **473**, Springer–Verlag, Berlin, 1991, pp. 204–13 441

[146] W. Meier and O. Staffelbach, Correlation properties of combiners with memory in stream ciphers, *Journal of Cryptology* **5** (1992), 67–86. 441

[147] W. H. Mills, Continued fractions and linear recurrences, *Math. Comp.* **29** (1975), 173–180. 145, 395

[148] L. M. Milne–Thompson, *The Calculus of Finite Differences*, Macmillan and Co., London, 1933. 173

[149] G. Mrugalski, J. Rajski, and J. Tyszer, Ring generators - new devices for embedded test applications, *IEEE Trans. on CAD of Integrated Circuits and Systems* **23(9)** (2004), 1306–1320. 267

[150] H. Niederreiter, The probabilistic theory of linear complexity, in C.G. Gunther, ed., *Advances in Cryptology EUROCRYPT 88*, Lecture Notes in Computer Science, **330**, Springer–Verlag, Berlin, 1988, pp. 191209. 494

[151] H. Niederreiter: Random Number Generation and Quasi–Monte Carlo Methods, SIAM, Philadelphia PA, 1992.

[152] J.-S. No, G.-M. Gil and D.-J. Shin, Generalized construction of binary bent sequences with optimal correlation property. IEEE. Trans. Info. Theory **49** (2003), 1769–1780. 317

[153] C. D. Olds, *Continued Fractions* New Mathematical Library, Mathematical Association of America, New York, 1963 141, 145

[154] J. Olsen, R. Scholtz and L. Welch, Bent-function sequences, IEEE. Trans. Info. Theory **IT-28** (1982), 858–864. 317

[155] F. Pappalardi and I. Shparlinski, On Artin's conjecture over function fields, *Finite fields and their applications* **1** (1991), 399–404. 354

[156] D. Peng and P. Fan, Lower bounds on the Hamming auto- and cross correlations of frequency-hopping sequences. IEEE. Trans. Info. Theory **50** (2004), 2149–2154. 293

[157] O. Perron, Bemerkungen über die Verteilung der quadratischen Reste. Math. Zeit, **56** (1952), 122–130. 339

[158] W. W. Peterson, *Error Correcting Codes*, MIT Press, Cambridge MA, 1961. 14

[159] W. W. Peterson and E. J. Weldon, Jr., *Error–Correcting Codes* second edition, MIT Press, Cambridge MA, 1972. 15, 168, 201

[160] J. Pollard, The Fast Fourier Transform in a Finite Field, *Math. Comp.,* **25** (1971), 365–374. 440

[161] W. Qi and H. Xu, Partial period distribution of FCSR sequences Publication information, *IEEE Trans. Info. Theory* **49** (2003), 761–765. 365

[162] N. Quang, J. Massey and L. Györfi, Families of sequences with optimal generalized Hamming correlation properties, http://www.isiweb.ee.ethz.ch/archive/massey_pub/pdf/Bl625.pdf 292, 293

[163] T. R. N. Rao, *Error Coding For Arithmetic Processors*, Academic Press, New York N. Y., 1974

[164] A. Ralston, De Bruijn sequences – a model example of the interaction of discrete mathematics and computer science, *American Math. Monthly* **55**(1982), 131–143. 279

[165] I. S. Reed, R. A. Scholtz, T. K. Truong and L. R. Welch, The fast decoding of Reed–Solomon codes using Fermat theoretic transforms and continued fractions, *IEEE Trans. Info. Theory* **IT–25** (1978), 100–106.

[166] O. S. Rothaus, On "bent" functions, J. Comb. Theory A **20** (1976), 300–305. 317

[167] R. Rueppel, New approaches to stream ciphers, Ph.D. thesis, Swiss Federal Institute of Technology, Zürich, 1984. 16, 411

[168] R. Rueppel, Correlation immunity and the summation generator, in H. C. Williams, ed., *Advances in Cryptology, Crypto '85*, Lecture Notes in Computer Science **218**, Springer–Verlag, Berlin, 1986, pp. 260-272. 413

[169] R. Rueppel, *Analysis and Design of Stream Ciphers.* Springer–Verlag, Berlin, 1986. 408, 409, 411, 413, 424, 441

[170] R. Rueppel and O. Staffelbach, Products of sequences with maximum linear complexity, *IEEE Trans. Info. Theory* **IT–33** (1987), 124–131. 413

[171] D. Sarwate and M. Pursley, Crosscorrelation properties of pseudorandom and related sequences. Proc. IEEE **68** (1980), 593–619.

[172] M. R. Schroeder, *Number Theory in Science and Communication*, fourth ed., Springer–Verlag, Berlin, 2006.

[173] W. Schmidt, *Equations Over Finite Fields, An Elementary Approach*, Springer–Verlag, Berlin, 1976. 75

[174] B. Schneier, *Applied Cryptography.* John Wiley & Sons, New York, 1996.

[175] R. Scholtz and L. Welch, GMW sequences, *IEEE Trans. Info. Theory* **IT–30** (1984), 548–553. 336

[176] E. Selmer, *Linear Recurrence Relations Over Finite Fields* Lecture notes, Dept. of Mathematics, University of Bergen, Norway, 1966 14

[177] A. Schönhage and V. Strassen, Schnelle Multiplikation Grosser Zahlen, *Computing,* **7** (1971), 281–292. 440

[178] C. Seo, S. Lee, Y. Sung, K. Han, and S. Kim, A Lower Bound on the Linear Span of an FCSR, *IEEE Trans. Info. Theory* **46** (2000), 691–693. 417

[179] M. Simon, J. Omura, R. Scholtz and B. Levitt, *Spread Spectrum Communications Handbook*, second edition, McGraw–Hill, Inc. New York, 1994. 289, 292

[180] T. Tian and W. Qi, Period and complementarity properties of FCSR memory sequences, *IEEE Trans. Info. Theory* **53** (2007), 2966–2970. 360

[181] J. Tate, Classes d'isogénies de variétés abéliennes sur un corps finis (d'àpres T. Honda), *Sém. Bourbaki* **21** (1968/69), Exp. 352. 326

[182] O. Teichmüller, Differentialrechnung bei charackteristik p, *J. Reine. u. Ang. Math.* (Crelle's J.) **175** (1936), 89–99. 403

[183] S. Tezuka, P. L'Ecuyer, and R. Couture, On the lattice structure of add–with–carry and subtract–with–borrow random number generators, *ACM Trans. Model. Comp. Sim.* **3** (1993), 315–331. 204

[184] V. Tspyschev, Full periodicity of Galois polynomials over nontrivial Galois rings of odd characteristic, J. Math. Sciences, **131** (2005). 341, 344

[185] P. Udaya and M. U. Siddiqi, Optimal large linear complexity frequency hopping patterns derived from polynomial residue class rings. IEEE. Trans. Info. Theory **44**(1988), 1492–1503. 341, 344

[186] R. Walters, *Spread Spectrum: Hedy Lamarr and the Mobile Phone*, BookSurge Publishing (Amazon.com), (2006). 289

[187] M. Ward, The arithmetical theory of linear recurring series, Trans. Amer. Math. Soc. **35** (1933), 600–628. 341, 344

[188] B. M. M. de Weger, Approximation lattices of p–adic numbers, *J. Num. Th.* **24** (1986), 70–88. 431, 436, 437

[189] A. Weil, On some exponential sums, *Proc. Nat. Acad. Sci.* **34** (1948), 204–207. 75

[190] L. R. Welch, Lower bounds on the maximum correlation of signals, *IEEE Trans. Info. Theory* **IT–20** (1974), 397–399.

[191] L. R. Welch and R. A. Scholtz, Continued fractions and Berlekamp's algorithm. *IEEE Trans. Info. Theory* **25** (1979), 19–27. 145, 395

[192] H. Xu and W. Qi, Autocorrelations of maximum period FCSR sequences, *SIAM Journal on Disc. Math.* **20** (2006), 568–577. 365

[193] H. Xu and W. Qi, Further results on the distinctness of decimations of ℓ-sequences, *IEEE Trans. Info. Theory* **52** (2006), 3831–3836. 368

[194] J. Xu and A. Klapper, Feedback with carry shift registers over $\mathbf{Z}/(n)$, in C. Ding, T. Helleseth, and H. Niederreiter, eds., *Proceedings of International Conference on Sequences and their Application (SETA), Singapore, December 1998*, Springer-Verlag, 1999.

[195] N. Zierler, Several binary sequence generators, Technical report 95, MIT Lincoln Laboratories, 1955. 14

[196] N. Zierler, Linear recurring sequences, *J. Soc. Indust. Appl. Math.* **7** (1959), 31–48. 14, 175, 178, 295, 298

Index