Sums of Rational Functions

by

James Wells Ratcliffe

B.Sc. (Hons.), Bishop's University, 2008

Thesis Submitted In Partial Fulfillment of the
Requirements for the Degree of
Master of Science or Doctor of

in the
Department of Mathematics
Faculty of Science

© James Wells Ratcliffe 2012

SIMON FRASER UNIVERSITY

Summer 2012

All rights reserved.
However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for "Fair Dealing." Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

APPROVAL

Name:	James Wells Ratcliffe
Degree:	Master of Science
Title of Thesis:	Sums of Rational Functions
Examining Committee:	**Nilima Nigam** Associate Professor (Chair)

Jason Bell
Senior Supervisor
Associate Professor

Imin Chen
Supervisor
Associate Professor

Michael Monagan
Internal Examiner
Professor

Date Approved: April 12, 2012

Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website (www.lib.sfu.ca) at http://summit/sfu.ca and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, British Columbia, Canada

revised Fall 2011

Abstract

Given a rational function $\phi(X)$ with rational coefficients that is defined at every positive integer, we consider the sum $\sum_{n=0}^{\infty} \phi(n)$. It is believed that when this sum converges, it converges to either a rational or transcendental number. We prove an analogue of this conjecture over fields of rational functions:

Let K be a field and let $\phi(X)$ be a rational function with coefficients in K such that $\phi(0) = 0$. Given a positive integer $d \geq 2$, we define $F(X) = \sum_{n=0}^{\infty} \phi(X^{d^n})$. If d is not a power of $\mathrm{char}(K)$, then $F(X)$ is either a rational function or transcendental over $K(X)$.

Our demonstration uses results from the theory of automatic sequences and from commutative algebra.

To my father.

Acknowledgments

I would like to thank my supervisor Jason Bell for introducing me to this problem and giving me a great deal of help completing my thesis, and Michael Monagan and Imin chen for reading it over very carefully. I would also like to thank my girlfriend Val for her love, encouragement and support and my son Nigel for keeping me cheerful near the end of my degree. This work was partially supported by NSERC CGS-M.

Contents

List of Figures

Part I

Background

Chapter 1

Introduction

1.1 Telescoping Sums

Let $\phi(X)$ be a rational function with rational coefficients such that $\phi(n)$ is defined for any nonnegative integer n. We are interested in the sum

$$S \coloneqq \sum_{n=0}^{\infty} \phi(n).$$

Example 1.1.1. Let $\phi(X) = \frac{1}{(X+1)(X+2)}$. Note that

$$\frac{1}{(X+1)(X+2)} = \frac{1}{X+1} - \frac{1}{X+2},$$

so for $N \geq 0$

$$\begin{aligned}\sum_{n=0}^{N} \phi(n) &= \sum_{n=0}^{N} \left(\frac{1}{n+1} - \frac{1}{n+2}\right) \\ &= 1 - \frac{1}{2} + \frac{1}{2} - \cdots - \frac{1}{N+1} + \frac{1}{N+1} - \frac{1}{N+2} \\ &= 1 - \frac{1}{N+2}.\end{aligned}$$

Thus

$$S = \sum_{n=0}^{\infty} \phi(n) = 1.$$

We say that $\phi(X)$ has a *telescoping* behaviour (or that the sum S is *telescoping*) since the number of terms in every partial sum is bounded after some cancellation, in this case by 2.

Example 1.1.2. Let $\phi(X) = \frac{2}{16X^2+16X+3}$. Now

$$\phi(X) = \frac{1}{4X+1} - \frac{1}{4X+3},$$

so

$$S = \sum_{n=0}^{\infty} \phi(n) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}.$$

This series is known as the Madhava–Leibniz series since it was shown to converge to $\pi/4$ by the Indian mathematician Madhava and later by Leibniz. See [Roy90].

Conjecture 1.1.3. *Let $\phi(X)$ be a rational function over $\mathbb{Q}$ such that $\phi(n)$ is defined for every positive integer n and define*

$$S := \sum_{n=0}^{\infty} \phi(n).$$

If S converges, then $S \in \mathbb{Q}$ or S is a transcendental number.

Assuming S is convergent, it is believed that $S \in \mathbb{Q}$ precisely when the sum has a telescoping behaviour.

In [ASST01], Adhikari, Saradha, Shorey and Tijdeman proved the following:

Theorem 1.1.4. *Let $P(X)$ and $Q(X)$ be a polynomials over the rationals such that $Q(X)$ has only simple rational roots. If*

$$S := \sum_{n=0}^{\infty} \frac{P(n)}{Q(n)}$$

converges, then it converges to either a rational or a transcendental number. Furthermore, if S is rational and the roots of $Q(X)$ all lie in the interval $[-1, 0)$, then $S = 0$.

Example 1.1.5. Let $\phi(X) = \frac{1}{(2X+1)(2X+2)}$. Then

$$S := \sum_{n=0}^{\infty} \frac{1}{(2n+1)(2n+2)}$$

converges and is clearly positive. Since the roots of $Q(X)$ are both in the interval $[-1, 0)$, we see that S is transcendental by the previous result. In fact, it can be shown that $S = \log(2)$.

1.2 Integers and Polynomials

When studying the ring of integers, one can sometimes gain insight into its structure by looking at polynomial rings. A polynomial ring in one variable over a field K has many properties in common with the ring of integers $\mathbb{Z}$. Many of these common properties stem from the fact that both are *Euclidean* rings.

For any two integers a and b where b is nonzero, there exist integers q and r with $O \leq |r| < |b|$ such that $a = qb + r$. Analogously, if $A(X)$ and $B(X) \neq 0$ are two polynomials over a field K, then there exist polynomials $Q(X)$ and $R(X)$ with $R(X) = 0$ or $\deg(R(X)) < \deg(B(X))$ such that $A(X) = Q(X)B(X) + R(X)$. The only difference in the statements is the function used to measure the "size" of the elements. This function is called a *Euclidean norm*. A ring R is called a Euclidean ring if there is a norm function $N : R \setminus \{0\} \to \mathbb{N}$ such that:

(1) $N(a) \leq N(ab)$ for all nonzero $a, b \in R$.

(2) For all $a, b \in R$ with $b \neq 0$, there exist $q, r \in R$ with $r = 0$ or $N(r) < N(b)$ such that $a = bq + r$.

It follows from (1) that any Euclidean ring is an integral domain, that is, the product of any two nonzero elements is always nonzero. For this reason they are usually called *Euclidean domains*. The second statement is often referred to as the division algorithm since it is usually proved using an algorithm (similar to grade-school long division).

An important consequence of the division algorithm is that both $\mathbb{Z}$ and $K[X]$ are *principal ideal domains*, that is, every ideal is generated by one element. In a principal ideal domain, every irreducible element is prime. A nonzero nonunit element p of a commutative ring is said to be *prime* if whenever p divides a product ab with $a, b \in R$, then either p divides a or p divides b. An *irreducible* element is a nonzero nonunit that cannot be written as the product of two nonunits. That primality and irreducibility are equivalent is an important similarity between $\mathbb{Z}$ and $K[X]$ since prime numbers play a central role in number theory.

In any principal ideal domain we can write any nonunit (except zero) as a product of irreducible elements, and this factorization is unique up to reordering and multiplication by a unit. Thus a Euclidean ring is a *unique factorization domain*, that is, an integral domain

with the unique factorization described above. This is a generalization of the Fundamental Theorem of Arithmetic to a larger class of integral domains.

The division algorithm and the fundamental theorem of arithmetic form the basis for a great deal of number theory, and so many of their consequences have analogues in polynomial rings. For example, if K is a field and $P(X), Q(X) \in K[X]$, then there is always a polynomial $R(X) \in K[X]$ that divides both $P(X)$ and $Q(X)$ and has maximal degree with respect to this property, that is, $R(X)$ is a *greatest common divisor* of $P(X)$ and $Q(X)$.

For proofs of these facts and more details, see [Irv04].

A polynomial ring over any field has a similar algebraic structure to that of the integers, but a polynomial ring over a finite field has even more similarities. One interesting example is the distribution of primes. The equivalent of the prime integers in a polynomial ring is the irreducible polynomials. We usually restrict to positive primes; we ignore the negative primes since they are just unit multiples of the positive primes. In the same way we will consider only monic irreducible polynomials since every element of the base field is a unit.

Euclid proved that there are infinitely many primes in $\mathbb{Z}$. The analogue of this result in a polynomial ring would be that there are irreducible polynomials of arbitrarily high degree. However, the only irreducible polynomials over $\mathbb{R}$ are linear and quadratic, and over $\mathbb{C}$, only linear polynomials are irreducible. This makes the disribution of irreducibles in $\mathbb{R}[X]$ and $\mathbb{C}[X]$ different from that of the primes in $\mathbb{Z}$, and so we would not expect to gain insight into the distribution of primes in $\mathbb{Z}$ by studying either of these polynomial rings.

In $\mathbb{Q}$ there are irreducible polynomials of arbitrarily high degree. Hayes [Hay65] proved that any polynomial over $\mathbb{Q}$ can be written as the sum of two irreducibles, and the proof is very elementary. This result is an analogue of the famous Goldbach Conjecture that any integer greater than 2 can be written as the sum of 2 primes. The Goldbach Conjecture has remained unresolved since 1742—this seems to suggest that the distribution of irreducibles in $\mathbb{Q}[\mathbb{X}]$ is more dense than that of the primes in $\mathbb{Z}$.

Over the finite field of size q, denoted $\mathbb{F}_q$, the number of monic irreducible polynomials of a given degree t is

$$N_q(t) = \frac{1}{t} \sum_{d|t} \mu(d) q^{t/d},$$

where μ is the *Möbius function* on $\mathbb{N}$ (see [LN97]). Thus the number of monic irreducible polynomials of degree t is asymptotic to $q^t/t = \frac{q^t}{\log_q(q^t)}$ as $t \to \infty$.

If we define the size of a polynomial in $\mathbb{F}_q[X]$ of degree t to be q^t, then the distribution of irreducibles in $\mathbb{F}_q[X]$ is similar to that of the primes in $\mathbb{Z}$, since the number of primes less than a positive integer n (or the number of primes of size less than n) is asymptotically

$$\pi(n) \sim \frac{n}{\log(n)}$$

by the well-known Prime Number Theorem.

The polynomial analogue of the Goldbach Conjecture, that any polynomial of degree n is the sum of two irreducible polynomials of degree n, has not been proved for $\mathbb{F}_q[X]$, though Pollack has recently made some progress towards a solution in [Pol11].

Another famous conjecture on the distribution of primes in $\mathbb{Z}$ is the Twin Prime Conjecture, which states that there are infinitely many primes p such that $p + 2$ is also prime. A generalized version was given by de Polignac (see [dP51]): for every positive even integer n, there are infinitely many primes p such that $p + n$ is also prime. See [Pol08] for a summary of work on the analogue of this conjecture for $\mathbb{F}_q[X]$, including new results by that paper's author.

For a more in-depth summary of the similarities between $\mathbb{Z}$ and $\mathbb{F}_q[X]$, see [EHM05].

1.3 The Main Result

We will prove an analogue of the conjecture from the first section for function fields. That is, we will prove the following result:

Main Theorem. *Let K be a field, let $\phi(X)$ be a rational function in $K(X)$ whose power series expansion lies in $XK[[X]]$, and let $d \geq 2$. Let $F(X) = \sum_{n=0}^{\infty} \phi(X^{d^n}) \in K[[X]]$. If d is not a power of* $\mathrm{char}(K)$*, then either $F(X) \in K(X)$ or $F(X)$ is transcendental over $K(X)$.*

Note that when $\mathrm{char}(K) = 0$ the theorem holds for any $d \geq 2$.

We require that $\phi(X) \in XK[[X]]$ and $d \geq 2$ to avoid problems with convergence. If $\phi(X)$ has a nonzero constant term a_0, then we would have

$$F(X) = \sum_{n=0}^{\infty} \phi(X^{d^n}) = \sum_{n=0}^{\infty} a_0 + G(X)$$

for some $G(X) \in XK[[X]]$. Since $\sum_{n=0}^{\infty} a_0$ is not necessarily defined in K, we may not get a valid power series.

With $\phi(X) \in XK[[X]]$, each coefficient of $F(X)$ is a polynomial in the coefficients of $\phi(X)$, so $F(X)$ is a well-defined power series over K. To see this, note that because $d \geq 2$, the degree of the lowest degree term in $\phi(X^{d^n})$ is strictly, and very rapidly, increasing as $n \to \infty$.

We consider sums of the form $\sum_{n=0}^{\infty} \phi(X^{d^n})$ because if $\phi(X) = \psi(X) - \psi(X^d)$ for some rational function $\psi(X)$, then we get exactly the type of telescoping behaviour that allows the sums to be rational.

To illustrate the main result, we give a few examples:

Example 1.3.1. Let $K = \mathbb{Q}$ and let

$$\phi(X) = \frac{X - X^2}{1 + X + X^2 + X^3} = \frac{X}{1 + X} - \frac{X^2}{1 + X^2}.$$

If we choose $d = 2$, then we have

$$\begin{aligned} F(X) &= \sum_{n=0}^{\infty} \phi(X^{2^n}) \\ &= \sum_{n=0}^{\infty} \frac{X^{2^n}}{1 + X^{2^n}} - \sum_{n=0}^{\infty} \frac{X^{2^{n+1}}}{1 + X^{2^{n+1}}} \\ &= \frac{X}{1 + X}. \end{aligned}$$

Example 1.3.2. Let $K = \mathbb{Q}$, let $\phi(X) = X \in K(X)$ and let $d = 2$. Then $F(X) = \sum_{n=0}^{\infty} X^{2^n} = \sum_{n \geq 0} a_n X^n$ where $a_n = 1$ if n is a power of 2 and $a_n = 0$ otherwise.

To see that $F(X)$ is not rational, suppose towards a contradiction that $F(X) = P(X)/Q(X)$ for some $P(X), Q(X) \in \mathbb{F}_2(X)$. We can assume that $P(0) = 0$ and $Q(0) \neq 0$. If we write $Q(X) = q_0 + \cdots + q_s X^s$, then the n-th coefficient of $P(X)$ is $q_0 a_n + \cdots + q_s a_{n-s}$. Choose N such that $2^N > \deg(P(X))$ and $2^N - 2^{N-1} > s$. Then for all $n \geq N$, the 2^n-th coefficient of $P(X)$ is $q_0 \neq 0$, a contradiction.

It will be easy to show that $F(X)$ is transcendental over $\mathbb{Q}(X)$ after we have seen some results on automatic sequences. See Example 4.4.4.

The following example demonstrates why we require that d not be a power of $\mathrm{char}(K)$.

Example 1.3.3. Let $K = \mathbb{F}_2$ and let $\phi(X)$ and d be defined as in Example 1.3.2. Then $d = \mathrm{char}(K)$. By the same argument as in 1.3.2, $F(X)$ is not rational. However, we can show that $F(X)$ is algebraic over $\mathbb{F}_2(X)$. To see this, note that since we are working over $\mathbb{F}_2$ we have

$$\begin{aligned} F(X)^2 &= F(X^2) \\ &= \sum_{n=0}^{\infty} X^{2^{n+1}} \\ &= F(X) - X, \end{aligned}$$

and so $F(X)$ is algebraic over $\mathbb{F}_2(X)$

We will first prove the Main Theorem in the case where the base field is finite using automata theory and automatic sequences.

A *deterministic finite automaton with output* is a finite collection of states and rules for moving between these states depending on the input provided. The automaton starts in a fixed initial state and is given a string as input. The input is read one symbol at a time and the automaton moves into different states based on the symbol read and the current state. The state in which the automaton ends up after reading the output string determines the output symbol.

Let k be an integer greater than 1. A sequence $\mathbf{a} = (a_0, a_1, \ldots)$ is said to be *k-automatic* if there is a deterministic finite automaton taking input strings from alphabets of the form $\{0, 1, \ldots, k-1\}$ that, when given the base k expansion of a positive integer n, outputs a_n.

Example 1.3.4. Let $\mathbf{a} = (0, 1, 1, 1, \ldots)$. Given $k \geq 2$, we can construct a deterministic finite automaton with output M that generates $\mathbf{a}$ as follows. Let M have 2 states, q_0 and q_1 with associated outputs 0 and 1, respectively. When reading an input string over $\{0, 1, \ldots, k-1\}$, M will start in q_0 and move to q_1 when it reads a nonzero symbol. Once in q_1, M stays in q_1. Thus M gives 0 as output after reading a base-k representation of 0 and 1 otherwise, so we can conlude that $\mathbf{a}$ is k-automatic for any $k \geq 2$.

We are interested in automatic sequences because of a result of Cobham's [Cob69] that states that if a sequence is k- and l-automatic for two multiplicatively independent integers k and l, then it is eventually periodic. We will also show see that a formal power series over a finite field is rational (a quotient of polynomials) if and only if its sequence of coefficients is eventually periodic. Thus, we need to show that the coefficients of $F(X)$ are k- and l-automatic for some multiplicatively independent integers k and l.

First we will show that the coefficients of $F(X) = \sum_{n=0}^{\infty} \phi(X^{d^n})$ are d-automatic. We will see that a power series over a finite field is rational if and only if its sequence of coefficients is eventually periodic. Büchi [Büc60] proved that an eventually periodic sequence is k-automatic for all $k \geq 2$, so we can conclude that the coefficients of $\phi(X)$ are d-automatic since $\phi(X)$ is a rational function over a finite field. We then show that this implies that the coefficients of $F(X)$ are also d-automatic since they depend on the coefficients of $\phi(X)$.

Christol's Theorem [Chr79] states that if a power series over a finite field of characteristic p is algebraic, then its coefficients are p-automatic.

When d is not a power of p, then d and p are multiplicatively independent. Thus, if $F(X)$ is algebraic, it follows that the coefficients of $F(X)$ are both d- and p-automatic and thus eventually periodic by Cobham's Theorem, so we can conclude that $F(X)$ is rational. We will then show that the degrees of the numerator and denominator polynomials of $F(X)$ are bounded in terms of the degrees of the numerator and denominator polynomials of $\phi(X)$.

In the general case, we will use some results from commutative algebra to reduce to the special case. The main result we will use is the Nullstellensatz. The Nullstellensatz, which is German for "zero-locus theorem", was originally proved by Hilbert. His version demonstrated a bijection between the radical ideals of a polynomial ring over an algebraically closed field and subvarieties of the corresponding affine space. We will use a more general version of the Nullstellensatz that was proved by Bourkbaki and applies to the larger class of Jacobson rings.

A ring R is called a *Jacobson ring* if the intersection of all the maximal ideals of R/P is (0) for any prime ideal P of R. For example, $\mathbb{Z}$ is Jacobson.

The general form of the Nullstellensatz that we will use states that if R is a Jacobson ring and S is a finitely-generated R-algebra, then S is also Jacobson; also, if N is a maximal ideal of S, then $M \coloneqq N \cap R$ is a maximal ideal of R and S/N is a finite field extension of R/M.

To use the Nullstellensatz, we will construct a $\mathbb{Z}$-algebra $R \subset K$ generated by the coefficients of the numerator and denominator polynomials of $\phi(X)$ such that $\phi(X)$ is contained in $R[[X]]$. It follows that $F(X) \in R[[X]]$ as well. Given a maximal ideal M of R, R/M is a finite field by the Nullstellensatz. If $F(X)$ is algebraic over $K(X)$, we can show that the image of $F(X)$ modulo M is algebraic over $(R/M)(X)$. We then show that since the image of $F(X)$ modulo M is rational and the degrees of its numerator and denominator polynomials are uniformly bounded, then $F(X)$ must be rational, which will complete the proof of the main result.

Chapter 2

Commutative Algebra

In this chapter we will give the background in commutative algebra that we will need in order to reduce the main result to the finite field case. Specifically, we need to define the class of Jacobson rings and present a version of the Nullstellensatz generalized to this class. This result will be a crucial part of the proof of the main result in the general case. Sections 2.2, 2.5, and 2.6 follow the apporoach of [Eis95]. Section 2.1 was written using [Gal10] as reference, Section 2.3 using [Gal10] and [AS03] and Section 2.4 using [Eis95] and [Har77].

2.1 Rings and Ideals

Let R be a commutative ring with multiplicative identity. Throughout this chapter, we will assume that all rings are commutative and have a multiplicative identity. We recall that an *ideal* of R is a nonempty set $I \subset R$ that is closed under addition and under multiplication by any element in R. If I is a proper subset of R, then we call I a *proper* ideal of R. We note that I is a proper ideal of R if and only if $1 \notin I$.

If $S \subset R$, then the *ideal generated by* S, denoted by (S), is the smallest ideal of R containing S, or, equivalently, the set of all linear combinations of elements of S with coefficients in R. We will leave out braces for the sake of clarity: we will write (a, b) instead of $(\{a, b\})$, for example. Note that for $a \in R$, $(a) = aR \coloneqq \{ar : r \in R\}$.

We note that if I and J are ideals of R, then $I \cap J$ is also an ideal. In general $I \cup J$ is not an ideal, but it is easy to see that $(I \cup J) = I + J \coloneqq \{a + b : a \in I, b \in J\}$. We also define $IJ = (ab : a \in I, b \in J)$, since $\{ab : a \in I, b \in J\}$ is generally not an ideal. Note that

$IJ \subset I \cap J$.

If I is an ideal of R, then we can define a relation on R as follows: if $a, b \in R$, then we say a is *congruent modulo* I to b if and only if $b - a \in I$. It is easy to check that this is an equivalence relation and that for $a \in R$, the equivalence class of a is $a + I \coloneqq \{a + r : r \in I\}$. We will generally use the more compact notation $\overline{a}$ to denote the equivalence class of a. We let R/I denote the *quotient ring* of R modulo I, which is the set of equivalence classes of R with addition and multiplication defined as follows for $a, b \in R$:

$$\overline{a} + \overline{b} = \overline{a + b};$$
$$\overline{a}\overline{b} = \overline{ab}.$$

Since I is an ideal, it follows that R/I is a ring under these operations. We have a surjective homomorphism $R \to R/I$ given by $a \mapsto \overline{a}$, and so we sometimes call $\overline{a}$ the *image* of a in R/I or the image of a modulo I.

Example 2.1.1. A subset I of the ring of integers $\mathbb{Z}$ is an ideal if and only if $I = n\mathbb{Z}$ for some $n \in \mathbb{Z}$.

Example 2.1.2. Let $R = \mathbb{C}[x, y]$. Then $I = (x - 1, y - 1) = (x - 1)R + (y - 1)R$ is an ideal of R, and $\mathbb{C}[x, y]/I \cong \mathbb{C}$, as this amounts to evaluating each polynomial at $x = 1$, $y = 1$.

Two special types of ideals feature prominently in commutative algebra, namely *prime* and *maximal* ideals.

Definition 2.1.3. Let R be a ring and I a proper ideal of R. We say that I is *prime* if whenever $ab \in I$, then either $a \in I$ or $b \in I$.

Example 2.1.4. An ideal $I = n\mathbb{Z}$ of $\mathbb{Z}$ is prime if and only if $n = \pm p$ for some prime number p.

Definition 2.1.5. Let R be a ring and I a proper ideal of R. We say that I is *maximal* if the only ideal of R properly containing I is R itself

We can obtain some important results concerning maximal ideals in a ring R by ordering R by inclusion and studying the resulting structure.

Definition 2.1.6. A *partially ordered set* is a set S along with a relation, or *order*, $\leq$ such that for any $a, b, c \in S$ we have

(1) $a \leq a$;

(2) if $a \leq b$ and $b \leq c$, then $a \leq c$;

(3) if $a \leq b$ and $b \leq a$, then $a = b$.

An element $a \in S$ is called *maximal* if there is no $b \in S$ such that $a \leq b$. A subset T of S is said to be *totally ordered* if for any $a, b \in T$ either $a \leq b$ or $b \leq a$.

If we order the set of proper ideals of a ring R by inclusion, then the maximal ideals of R are precisely those ideals that are maximal elements of this partially ordered set. We can then use Zorn's lemma to prove the existence of maximal ideals.

Lemma 2.1.7 (Zorn, [Zor35])**.** *Let S be a partially ordered set. Suppose that for every subset $T \subset S$, there exists $b \in S$ such that $a \leq b$ for all $a \in T$. Then S has at least one maximal element.*

For a proof of Zorn's lemma, see [HJ99, Theorem 1.3].

Proposition 2.1.8. *Let R be a ring and let I be a proper ideal of R. Then R has a maximal ideal containing I.*

Proof. Let $\mathcal{S}$ be the set of all proper ideals of R containing I, ordered by containment. We will use Zorn's lemma to show that $\mathcal{S}$ has a maximal element. Let A be an index set and let $\{I_\alpha : \alpha \in A\}$ be a totally ordered subset of $\mathcal{S}$. Let $J = \cup_{\alpha \in A} I_\alpha$; we claim that $J \in \mathcal{S}$.

Since J_α is nonempty for each $\alpha \in A$, then J is nonempty. If $a, b \in J$, then $a \in J_\alpha$ and $b \in J_\beta$ for some $\alpha, \beta \in A$. Without loss of generality, we have $J_\alpha \supset J_\beta$, so $a, b \in J_\alpha$, and thus $a + b \in J_\alpha \subset J$. Also, if $r \in R$, $ra \in J_\alpha \subset J$. Finally, since $J_\alpha \supset I$ for all $\alpha \in A$, we have $J \supset I$, and so $J \in \mathcal{S}$. Thus, by Zorn's lemma, $\mathcal{S}$ has a maximal element, M. To see that M is a maximal ideal of R, suppose there is an ideal K of R such that $M \subsetneq K \subset R$. Then $I \subset K$, and thus, since M is maximal in $\mathcal{S}$, $K = R$. □

Corollary 2.1.9. *Every nontrivial ring has a maximal ideal.*

Proof. Let R be a nontrivial ring, that is, $R \neq \{0\}$. Then (0) is a proper ideal of R, so, by Proposition 2.1.8, R has a maximal ideal. □

The argument using Zorn's lemma in the proof of Proposition 2.1.8 is fairly standard, and we will refer to it in this chapter.

Next we define two important classes of rings that give very useful characterizations of prime and maximal ideals.

Definition 2.1.10. A ring R is called an *integral domain* if whenever $ab = 0$ for some $a, b \in R$, we have either $a = 0$ or $b = 0$.

If there exist nonzero elements a, b of a ring R such that $ab = 0$, we call a and b *zero divisors*. Note that R is an integral domain if and only if R has no zero divisors. We also note that (0) is a prime ideal of a ring R if and only if R is an integral domain; this follows directly from the definitions.

Definition 2.1.11. Let R be a ring. An element $r \in R$ is called a *unit* if there exists another element $s \in R$, called the *inverse* of r, such that $rs = 1$. We denote by $U(R)$ the multiplicative *units group* of R. If every nonzero element of R is a unit, we say that R is a *field*.

Example 2.1.12. The ring of integers $\mathbb{Z}$ is an integral domain and $U(R) = \{-1, 1\}$, so $\mathbb{Z}$ is not a field, but $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are.

We note that if I is an ideal of R and I contains a unit of R, then $1 \in I$ and thus $I = R$.

Proposition 2.1.13. *Let I be an ideal of a ring R, then*

(1) I is prime if and only if R/I is an integral domain.

(2) I is maximal if and only if R/I is a field.

Proof. For the first statement, suppose that I is prime and let $a, b \in R$ such that $\overline{ab} = \overline{0}$. Then $ab \in I$, and so either $a \in I$ or $b \in I$. Thus either $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$.

Conversely, suppose that R/I is an integral domain and let $a, b \in I$. Then $\overline{ab} = \overline{0}$, so either $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$. Thus either $a \in I$ or $b \in I$.

For the second statement, suppose that I is maximal and let $a \in R$ such that $\overline{a} \neq \overline{0}$. Then $a \notin I$, so $(a) + I \supsetneq I$. Since I is maximal, $(a) + I = R$, and so $1 = ab + x$ for some $b \in R$ and $x \in I$. Thus $ab - 1 \in I$, which implies that $\overline{a}\overline{b} = \overline{1}$, and so $\overline{a}$ is a unit.

Conversely, suppose that R/I is a field and let J be an ideal properly containing I. Choose $a \in J \setminus I$. Then $\overline{a} \neq \overline{0}$, so there exists $b \in R$ such that $\overline{a}\overline{b} = \overline{1}$. Thus $1 - ab \in I \subset J$. Now $ab \in J$, and so $1 \in J$, which implies that $J = R$. Therefore I is maximal. □

An immediate consequence of Proposition 2.1.13 is that every maximal ideal is also prime. Moreover, we get a useful characterization of fields, namely that their only proper ideal is (0). To see this, note that $R \cong R/(0)$, which is a field if and only if (0) is maximal.

We end with a few more definitions specific to integral domains that will be needed later.

Definition 2.1.14. Let R be an integral domain and let a be a nonzero nonunit of R. We say that a is irreducible if whenever we write $a = bc$ with $b, c \in R$, either b or c is a unit.

Example 2.1.15. In $\mathbb{Z}$, the irreducible elements are all $\pm p$ where p is prime.

An integral domain in which every ideal is generated by one element is called a *principal ideal domain*. It is not hard to show that in a principal ideal domain, every prime ideal is maximal and that (a) is prime and maximal if and only if a is irreducible.

2.2 Localization

It can often be useful to make nonunits in a ring into units. We do this by constructing fractions of the elements in the ring, which work in the same way as the elements of $\mathbb{Q}$. Let R be a ring and let U be a subset of R that is *multiplicatively closed*, that is any product of elements of U is in U, including 1, the "empty product". We define an equivalence relation on $R \times U$ as follows: $(a, b) \sim (c, d)$ if there exists $x \in U$ such that $x(ad - bc) = 0$. We then define $R[U^{-1}]$, called the *localization of R at U*, to be the set of equivalence classes of $R \times U$, where we denote the equivalence class of (a, b) by $\frac{a}{b}$. We can show that $R[U^{-1}]$ is a ring under the following operations:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{cd};$$
$$\frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$

We note that if $a \in R$, then $\frac{a}{1} = 0$ in $R[U^{-1}]$ if and only if there exists an element $b \in U$ such that $ab = 0$. In particular, $R[U^{-1}] = \{0\}$ precisely when there exists $b \in U$ such that $bR = (0)$. We can avoid these situations by requiring that U contains no zerodivisors.

There is a natural map $R \to R[U^{-1}]$ that send a to $\frac{a}{1}$. If U contains no zerodivisors, then this map is an injection, so we can view R as a subring of $R[U^{-1}]$. When $a \in U$, we have $a\frac{1}{a} = 1$, and so a is a unit with inverse $\frac{1}{a}$ in $R[U^{-1}]$.

When U is the set of all nonzerodivisors of R, we call $R[U^{-1}]$ the *total ring of fractions* of R. We note that when R is an integral domain, its total ring of fractions is a field, and we call it the *field of fractions* of R.

Example 2.2.1. The field of fractions of $\mathbb{Z}$ is $\mathbb{Q}$.

Example 2.2.2. The ring of *Gaussian integers* is

$$\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\},$$

where $i = \sqrt{1}$. If $a, b, c, d \in \mathbb{Z}$, then

$$\begin{aligned}\frac{a+bi}{c+di} &= \left(\frac{a+bi}{c+di}\right)\left(\frac{c-di}{c-di}\right) \\ &= \frac{ac+bd+(bc-ad)i}{c^2+d^2}.\end{aligned}$$

Therefore the field of fractions of $\mathbb{Z}[i]$ is

$$\mathbb{Q}[i] = \{r + si : r, s \in \mathbb{Q}\}.$$

2.3 Polynomials and Power Series

In this section we will look at polynomial rings and their extensions. Let R be a ring. We denote by $R[X]$ the *ring of polynomials* in the variable X over R, that is,

$$R[X] = \left\{a_0 + a_1X + a_2X^2 + \cdots + a_dX^d : d \geq 0, a_0, \ldots, a_d \in R\right\}.$$

Polynomials are added and multiplied by the same operations of the ring of coefficients, R. In particular, if $\sum_{i=0}^{d} a_iX^i$ and $\sum_{i=0}^{e} b_iX^i$ are two polynomials in $R[X]$, then their product is $\sum_{k=0}^{d+e} c_kX^k$, where $c_i = \sum_{j=0}^{k} a_{k-j}b_j$.

It is generally more convenient to think of polynomials as having infinitely many terms but only finitely many with nonzero coefficients. The *degree* of a polynomial $P(X)$, written $\deg(P(X))$, is defined to be the highest power of X with a nonzero coefficient. For example, the polynomial $1 + X + X^2$ has degree 2. By convention, we define the degree of the 0 polynomial to be $-\infty$.

The ring R is embedded in $R[X]$ as each nonzero element in R corresponds to a polynomial of degree 0. If R is an integral domain, it is easy to see that $R[X]$ is also an integral domain, and in this case the units of $R[X]$ are precisely the units of R.

We will be primarily interested in polynomials over a field K. It can be shown using the division algorithm that $K[X]$ is a Euclidean domain, that is, given $A(X), B(X) \in K[X]$ with $B(X) \neq 0$, we can write

$$A(X) = B(X)Q(X) + R(X)$$

with $Q(X), R(X) \in K[X]$ and $\deg(R(X)) < \deg(B(X))$.

It is a direct consequence of this that $K[X]$ is a principal ideal domain and a *unique factorization domain*, that is, that any nonconstant element of $K[X]$ can be written uniquely (up to reordering and multiplication by a unit) as a product of one or more irreducible elements. It follows from this that any two elements of $K[X]$ have a *greatest common divisor*, that is, a common divisor with maximal degree. Proofs of these facts and more details about polynomials can be found in most undergraduate algebra texts; see, for example, [Gal10].

The field of fractions of $K[X]$ is denoted by $K(X)$ and is called the *function field*, or the field of *rational functions* over K.

We can extend the ring of polynomials over a field K by allowing infinitely many nonzero coefficients, obtaining the ring of *formal power series* over K,

$$K[[X]] = \left\{ \sum_{n \geq 0} a_n X^n : a_0, a_1, \ldots \in K \right\}.$$

These are called formal power series since we do not care whether or not the sum converges at a given value X in K. Thus, in general we cannot talk about the value of a power series for certain values of X as we can with polynomials or rational functions. However, we will still use function notation such as $F(X)$ to refer to power series for consistency.

Proposition 2.3.1. *Let R be a ring and let $F(X) = \sum_{n \geq 0} a_n X^n \in R[[X]]$. Then $F(X)$ is a unit in $R[[X]]$ if and only if a_0 is a unit in R.*

Proof. If there exists $G(X) = \sum_{n\geq 0} b_n X^n \in R[[X]]$ such that $F(X)G(X) = 1$, then we see that $a_0 b_0 = 1$, so a_0 is a unit in R.

Conversely, suppose that a_0 is a unit in R and let $b_0 = 1/a_0$. For $n \geq 1$, define

$$b_n = -\frac{1}{a_0} \sum_{k=1}^{n} a_k b_{n-k}.$$

Define $G(X) = \sum_{n\geq 0} b_n X^n$; it is straightforward to check that $F(X)G(X) = 1$. □

We can see $K[X]$ as a subring of $K[[X]]$. As a consequence of Proposition 2.3.1, all polynomials with nonzero constant coefficients have inverses in $K[[X]]$, or to put it another way, if $\phi(X) \in K(X)$, then $\phi(X) \in K[[X]]$ whenever $\phi(X) = P(X)/Q(X)$ where $P(X), Q(X) \in K[X]$ with $Q(0) \neq 0$.

Example 2.3.2. Let $F(X) = 1 - X$ and $G(X) = \sum_{n\geq 0} X^n$. Then

$$\begin{aligned} F(X)G(X) &= (1 - X)(1 + X + X^2 + \cdots) \\ &= 1 - X + X - X^2 + X^2 - \cdots \\ &= 1. \end{aligned}$$

Thus

$$\frac{1}{1 - X} = G(X) = \sum_{n\geq 0} X^n.$$

The field of fractions of $K[[X]]$ is called the field of *formal laurent series* over K. It is denoted by $K((X))$ and given by

$$K((X)) = \left\{ \sum_{n\geq N} a_n X^n : n \in \mathbb{Z}, a_0, a_1, \ldots \in K \right\}.$$

Example 2.3.3. The polynomial $X + X^2$ is not a unit in $K[[X]]$ for any field K. However, we can find its inverse in $K((X))$. We saw in Example 2.3.2 that

$$\frac{1}{1 - X} = \sum_{n\geq 0} X^n.$$

We replace X with $-X$ to obtain

$$\frac{1}{1 + X} = \sum_{n\geq 0} (-1)^n X^n,$$

and so

$$\frac{1}{X + X^2} = \sum_{n\geq -1} (-1)^{n+1} X^n.$$

2.4 The Zariski Topology

We let Spec(R) denote the set of all prime ideals of a ring R, called the *spectrum* of R.

Example 2.4.1. If K is a field, then

$$\mathrm{Spec}(K[X]) = \{0\} \cup \{\langle P(X)\rangle : P(X) \in K[X] \text{ and } P(X) \text{ is irreducible }\}.$$

If K is algebraically closed, then

$$\mathrm{Spec}(K[X]) = \{0\} \cup \{\langle X - \alpha\rangle : \alpha \in K\}.$$

In this case the spectrum is in bijection with the base field K plus an extra element corresponding to (0), called the *generic point*.

We can endow Spec(R) with a topology, called the Zariski topology, in a natural way. First, however, we will look at an ideal of R that relates to Spec(R).

Definition 2.4.2. Let R be a ring. We say that an element $a \in R$ is *nilpotent* if there exists a positive integer n such that $a^n = 0$. We say that an ideal I of R is *nil* if every element of I is nilpotent.

Proposition 2.4.3. *If R is a ring and $\{N_\alpha : \alpha \in A\}$ is a collection of nil ideals of R, then $\sum_{\alpha\in A} N_\alpha$ is also nil. In particular, R has a unique largest nil ideal, denoted* Nil(R)*, formed by taking the sum of all nil ideals of R.*

Proof. Let $\{N_\alpha : \alpha \in A\}$ be a collection of nil ideals of R. Let $N = \sum_{\alpha\in A} N_\alpha$. If $a \in N$, then $a = \sum_{i=1}^{t} a_i$ where $\alpha_i \in A$, and $a_i \in N_{\alpha_i}$, for each $i = 1, \ldots, t$. Since each N_{α_i} is nil, there exist natural numbers $n_1, \ldots, n_t$ such that $a_i^{n_i} = 0$ for $i = 1, \ldots, t$.

First, suppose that $t = 2$. Note that

$$(a_1 + a_2)^{n_1+n_2} = \sum_{i=0}^{n_1+n_2} \binom{n_1 + n_2}{i} a_1^i a_2^{n_1+n_2-i}.$$

For each $0 \le i \le n_1 + n_2$, either $i \ge n_1$ or $n_1 + n_2 - i \ge n_2$, so each term of the right-hand side is 0, and so $a_1 + a_2$ is nilpotent. By induction we see that any finite sum of nilpotent elements of R is nilpotent, and so N is nil.

If we let Nil(R) be the sum of all nil ideals of R, then it follows that Nil(R) is nil and that every nil ideal of R is contained in Nil(R). □

The ideal $\mathrm{Nil}(R)$ is called the *nilradical* of R. Note that if R is an integral domain, then the only nilpotent element of R is 0, and so $\mathrm{Nil}(R) = (0)$. We note that $\mathrm{Nil}(R)$ is the set of all nilpotent elements of R, since if $a \in R$ is nilpotent, (a) is nil, and so $(a) \subset \mathrm{Nil}(R)$. The next proposition characterizes $\mathrm{Nil}(R)$ is terms of the prime ideals of R.

Proposition 2.4.4. *Let R be a ring. Then*

$$\mathrm{Nil}(R) = \bigcap_{P \in \mathrm{Spec}(R)} P.$$

Proof. First we show that $\mathrm{Nil}(R) \subset P$ for all $P \in \mathrm{Spec}(R)$. Let $a \in \mathrm{Nil}(R)$. Then there is a positive integer n such that $a^n = 0$. It follows that $a^n \in P$ for all $P \in \mathrm{Spec}(R)$ since P is an ideal and so $0 \in P$. Therefore $a \in P$ for each $P \in \mathrm{Spec}(R)$ since every $P \in \mathrm{Spec}(R)$ is prime.

Now we show that $a \in \cap_{P \in \mathrm{Spec}(R)} P \subset \mathrm{Nil}(R)$. Let $a \in \cap_{P \in \mathrm{Spec}(R)} P$. To show that $a \in \mathrm{Nil}(R)$, it suffices to show that a is nilpotent. Suppose a is not nilpotent and let $S = \{1, a, a^2, \ldots\}$. Consider the set $\mathcal{S}$ of ideals I of R such that $I \cap S = \emptyset$, ordered by inclusion. By an argument similar the one using Zorn's lemma in the proof of Proposition 2.1.8, we can show that $\mathcal{S}$ has a maximal element Q.

We claim that Q is prime. To see this, let $x, y \in R$ with $xy \in Q$ but neither $x \in Q$ nor $y \in Q$. Define $Q_1 = Q + xR$ and $Q_2 = Q + yR$, then both Q_1 and Q_2 properly contain Q. Since Q is maximal in $\mathcal{S}$, we have $Q_1 \cap S \neq \emptyset$ and $Q_2 \cap S \neq \emptyset$, thus there exist positive integers n_1 and n_2 such that $x^{n_1} \in Q_1$ and $x^{n_2} \in Q_2$. Then

$$\begin{aligned} x^{n_1+n_2} \in Q_1 Q_2 &= (Q + xR)(Q + yR) \\ &= Q(Q + xR + (yR)) + xyR \\ &\subset Q + xyR \\ &\subset Q, \end{aligned}$$

since $xy \in Q$. This is a contradiction, so Q is prime. By construction, $a \notin Q$, but this contradicts our choice of $a \in \bigcap_{P \in \mathrm{Spec}(R)} P$. Therefore a is nilpotent. □

Proposition 2.4.4 is a very powerful tool in commutative algebra. The following result is one application.

Corollary 2.4.5. *Let R be a ring. Then the group of units of $R[X]$ is*

$$U(R[X]) = \{a_0 + a_1X + \cdots + a_nX^n : n \in \mathbb{N}, a_0 \in U(R), a_1, \ldots, a_n \text{ are nilpotent}\}$$

Proof. Let $V = \{a_0 + a_1X + \cdots + a_nX^n : n \in \mathbb{N}, a_0 \in U(R), a_1, \ldots, a_n$ are nilpotent$\}$. We will first show that any element of V is a unit in $R[X]$. Let $A(X) = a_0 + a_1X + \cdots a_nX^n \in V$, then $a_0 \in U(R)$, and $a_1, \ldots, a_n$ are nilpotent elements of R. Let $B(X) = a_1X + \cdots + a_nX^n$. Each term, a_iX^i, is nilpotent, so, as seen in the proof of Proposition 2.4.3, $B(X)$ is nilpotent, and so there exists $m \in \mathbb{N}$ such that $B(X)^m = 0$. Now

$$\begin{aligned}(a_0 + B(X))\left(a_0^{m-1} - a_0^{m-2}B(X) + a_0^{m-3}B(X)^2 - \cdots + (-B(X))^{m-1}\right) &= a_0^m + (-1)^{m-1}B(X)^m \\ &= a_0^m.\end{aligned}$$

Since $a_0 \in U(R)$, then $a_0^m \in U(R)$. Let

$$C(X) = \left(a_0^{m-1} - a_0^{m-2}B(X) + a_0^{m-3}B(X)^2 - \cdots + (-B(X))^{m-1}\right)(a_0^m)^{-1};$$

then $(a_0 + B(X))C(X) = 1$, so $A(X) \in U(R[X])$.

Now suppose $A(X) = a_0 + a_1X + \cdots + a_nX^n \in U(R[X])$ with inverse $B(X) = b_0 + b_1X + \cdots + b_mX^m$. We wish to show that $A(X) \in V$. Let $P \in \operatorname{Spec}(R)$. We denote by $\overline{A}(X)$ and $\overline{B}(X)$ the polynomials in $(R/P)[X]$ obtained by reducing each coefficient of $A(X)$ and $B(X)$, respectively, modulo P. Then $\overline{A}(X) \in U(R/P[X])$ with inverse $\overline{B}(X)$.

We claim that $\overline{a_1} = \cdots = \overline{a_n} = 0$. If not, then there exists $j \geq 1$ such that $\overline{a_j} \neq 0$ and $\overline{a_i} = 0$ for all $i > j$. Similarly, there exists $k \geq 0$ such that $\overline{b_k} \neq 0$ and $\overline{b_i} = 0$ for all $i > k$. Then

$$\begin{aligned}1 &= \overline{A}(X)\overline{B}(X) \\ &= \left(\overline{a_0} + \overline{a_1}X + \cdots + \overline{a_n}X^n\right)\left(\overline{b_0} + \overline{b_1}X + \cdots + \overline{b_n}X^n\right) \\ &= \overline{a_j}\overline{b_k}X^{i+k} + F(X),\end{aligned} \tag{2.1}$$

where $i+k > 0$ and $F(X) \in R[X]$ with degree less than $i+k$. However, since P is prime, R/P is an integral domain, and so $\overline{a_j}\overline{b_k} \neq 0$, which contradicts (2.1). Thus $\overline{a_1} = \cdots = \overline{a_n} = 0$, that is, $a_1, \ldots, a_n \in P$. This holds for any $P \in \operatorname{Spec}(R)$, so $a_1, \ldots, a_n \in \operatorname{Nil}(R)$ by Proposition 2.4.4. Finally, we note that $1 = A(X)B(X)$ implies that $1 = a_0b_0$, and thus $a_0 \in U(R)$. □

We will now introduce the Zariski topology.

Definition 2.4.6. Let R be a ring. We call a subset $C \subset \mathrm{Spec}(R)$ *closed* if there is an ideal I of R such that

$$C = \{P \in \mathrm{Spec}(R) : P \supset I\}.$$

We call a subset of $\mathrm{Spec}(R)$ *open* if it is the complement of a closed set.

The collection of open subsets defined above defines the *Zariski topology* on $\mathrm{Spec}(R)$.

Example 2.4.7. Let K be a field. If $C \subset \mathrm{Spec}(K[X])$ is a closed set under the Zariski topology, then

$$C = \{\langle Q(X)\rangle : Q(X) \in K[X], Q(X) \text{ is irreducible, and } Q(X) \mid P(X)\}$$

for some $P(X) \in K[X]$.

When K is algebraically closed, we see that each proper closed subset of $\mathrm{Spec}(K[X])$ is of the form

$$C = \{\langle X - \alpha\rangle : P(\alpha) = 0\}$$

for some $P(X) \in K[X] \setminus (0)$.

Proposition 2.4.8. *Let R be a ring. Then the collection of open sets from Definition 2.4.6 forms a topological space.*

Proof. We will check that $\emptyset$ and $\mathrm{Spec}(R)$ are closed and that arbitrary intersections and finite unions of closed sets are themselves closed.

We note that $\emptyset = \{P \in \mathrm{Spec}(R) : P \supset R\}$ and $\mathrm{Spec}(R) = \{P \in \mathrm{Spec}(R) : P \supset (0)\}$, so $\emptyset$ and R are closed.

Now suppose that for each α in some index set J we have and ideal I_α of R and a closed set $C_\alpha = \{P \in \mathrm{Spec}(R) : P \supset I_\alpha\}$. Let $I = \sum_{\alpha\in A} I_\alpha$. We claim that $\{P \in \mathrm{Spec}(R) : P \supset I\} = \bigcap_{\alpha\in A} C_\alpha$. To see this, observe that $P \in \bigcap_{\alpha\in A} C_\alpha$ if and only if $P \supset I_\alpha$ for all $\alpha \in A$, which is true precisely when $P \supset \sum_{\alpha\in A} I_\alpha = I$.

Finally, suppose that C_1 and C_2 are closed sets. Then there exist ideals I_1 and I_2 of R such that $C_i = \{P \in \mathrm{Spec}(R) : P \supset I_i\}$ for $i = 1, 2$. Let $I = I_1 I_2$; we claim that $\{P \in \mathrm{Spec}(R) : P \supset I\} = C_1 \cup C_2$. To see this, note that $P \in C_1 \cup C_2$ if and only if $P \supset I_i$ for some $i = 1, 2$. If $P \supset I_i$ for some $i = 1, 2$, then $P \supset I$ since $I = I_1 I_2 \subset I_i$. Now suppose that $P \supset I$ and $P \not\supset I_i$ for $i = 1, 2$. Then there exist $a \in I_1 \setminus P$ and $b \in I_2 \setminus P$, and so $ab \in I_1 I_2 \subset P$.

However, since P is a prime ideal, then we have either $a \in P$ or $b \in P$, a contradiction. Therefore $C_1 \cup C_2 = \{P \in \text{Spec}(R) : P \supset I\}$. A simple induction shows that any finite union of closed sets is closed. □

Many topological properties of Spec(R) can be characterized by algebraic properties of R. For example, we say that a topological space is *disconnected* if it is the union of two disjoint proper closed subsets; the following theorem gives two algebraic characterizations of rings R such that Spec(R) is disconnected.

Theorem 2.4.9. *Let R be a ring. Then the following are equivalent:*

(1) Spec(R) *is disconnected;*

(2) R has two nonzero idempotents e_1 and e_2 such that $e_1 + e_2 = 1$ and $e_1 e_2 = 0$;

(3) There exist nontrivial subrings R_1 and R_2 of R such that $R \cong R_1 \times R_2$.

Proof. Suppose that Spec(R) is disconnected. Then Spec(R) $= C_1 \cup C_2$ where C_1 and C_2 are closed, disjoint proper subsets of Spec(R). We can therefore choose two proper ideals I_1 and I_2 of R such that $C_i = \{P \in \text{Spec}(R) : P \supset I\}$ for $i = 1, 2$. Since $C_1 \cap C_2 = \emptyset$, we have $I_1 + I_2 = R$, and since $C_1 \cup C_2 = \text{Spec}(R)$, every $P \in \text{Spec}(R)$ must contain $I_1 I_2$.

Choose $a \in I_1$ and $b \in I_2$ such that $a + b = 1$. Then $ab \neq 0$ since $I_1, I_2 \subsetneq R$. Now $ab \in I_1 I_2 \subset \text{Nil}(R)$, so there exists a positive integer n such that $(ab)^n = 0$. Let

$$e_1 = \sum_{j=0}^{2n} \binom{2n}{j} a^j b^{2n-j}, e_2 = \sum_{j=n+1}^{2n} \binom{2n}{j} a^j b^{2n-j};$$

then

$$e_1 + e_2 = \sum_{j=0}^{n} \binom{2n}{j} a^j b^{2n-j} = (a+b)^{2n} = 1.$$

Notice that each term in e_1 is in Rb^n, so $e_1 \in Rb^n$. Similarly, $e_2 \in Ra^n$, and so $e_1 e_2 \in R(ab)^n = (0)$. Thus $e_1 e_2 = 0$, and so

$$e_1^2 = e_1^2 + e_1 e_2 = e_1(e_1 + e_2) = e_1.$$

Similarly, $e_2^2 = e_2$. Since e_1 and e_2 are in proper ideals of R, neither of them is 1, and so R has two nonzero idempotents whose sum is 1 and whose product is 0.

Now suppose that R has nontrivial idempotents e_1 and e_2 such that $e_1 + e_2 = 1$ and $e_1e_2 = 0$. For $i = 1, 2$, let $R_i = Re_i$ as a subring of R with identity e_i. Then we have a ring homomorphism $\phi : R \to R_1 \times R_2$ given by $\phi(r) = (re_1, re_2)$. To see that this is a homomorphism, we note that if $r, s \in R$, then

$$\phi(rs) = (rse_1, rse_2) = (r, r)(se_1, se_2) = (re_1, re_2)(se_1, se_2) = \phi(r)\phi(s).$$

It is easy to check that ϕ satisfies the other conditions of a ring homomorphism.

Notice that

$$\begin{aligned}\mathrm{Ker}(\phi) &= \{r \in R : re_1 = re_2 = 0\} \\ &\subset \{r \in R : r(e_1 + e_2) = 0\} \\ &= (0)\end{aligned}$$

since $e_1 + e_2 = 1$. Therefore ϕ is injective.

To see that ϕ is surjective, let $x, y \in R$ and consider the element $(xe_1, ye_2) \in R_1 \times R_2$. Then

$$\begin{aligned}\phi(xe_1 + ye_2) &= ((xe_1 + ye_2)e_1, (xe_1 + ye_2)e_2) \\ &= (xe_1^2 + ye_1e_2, xe_1e_2 + ye_2^2) \\ &= (xe_1, ye_2)\end{aligned}$$

since $e_1e_2 = 0$ and e_1 and e_2 are idempotent. Therefore ϕ is surjective, and so $R \cong R_1 \times R_2$.

Finally, suppose that $R \cong R_1 \times R_2$ for two nontrivial subrings R_1 and R_2. Without loss of generality, we can assume that $R = R_1 \times R_2$. Let $I_1 = R_1 \times \{0\}$ and $I_2 = \{0\} \times R_2$ and let $C_i = \{P \in \mathrm{Spec}(R) : P \supset I_i\}$ for $i = 1, 2$.

Since $I_1I_2 = \{0\}\times\{0\}$, it follows that every $P \in \mathrm{Spec}(R)$ contains I_1 or I_2. Thus $C_1 \cup C_2 = \mathrm{Spec}(R)$. Also, since $I_1 + I_2 = R$, we have $C_1 \cap C_2 = \emptyset$.

In order to show that C_1 and C_2 are nontrivial, it suffices to show that $\mathrm{Spec}(R_i)$ is nonempty for $i = 1, 2$. Let $P \in \mathrm{Spec}(R)$, and define $P_1 = \{a \in R_1 : (a, b) \in P \text{ for some } b \in R_2\}$. It is easy to check that P_1 is an ideal of R_1. Suppose that $aa' \in P_1$. Then $(aa', b) \in P$ for some $b \in R$. Since P is prime, either $(a, b) \in P$ or $(a', b) \in P$, so we have either $a \in P_1$ or $a' \in P_1$, and so P_1 is prime. We can construct a prime ideal of R_2 in a similar fashion, so we see that $\mathrm{Spec}(R_1), \mathrm{Spec}(R_2) \neq \emptyset$, and so $C_1, C_2 \neq \emptyset$. We thus conclude that $\mathrm{Spec}(R)$ is disconnected. □

The correspondence theorem, one of the basic results of ring theory, states that for any ideal I of a ring R, there is a bijection between the ideals of R/I and the ideals of R that contain I (see [Gal10]). The next result shows that if we restrict to prime ideals, we get a homeomorphism between $\mathrm{Spec}(R/I)$ and a closed set of $\mathrm{Spec}(R)$.

Proposition 2.4.10. *Let R be a ring, let I be an ideal of R, and let $C = \{P \in \mathrm{Spec}(R) : P \supset I\}$. Then there is a bijection $\phi : \mathrm{Spec}(R/I) \to C$. Furthermore, ϕ is a homeomorphism.*

Proof. Let $Q \in \mathrm{Spec}(R/I)$; we define $\phi(Q) = \{a \in R : \overline{a} \in Q\}$. We must first show that $\phi(Q) \in C$. If $a \in I$, then $\overline{a} = \overline{0} \in Q$, so $a \in \phi(Q)$, and thus $I = \phi(Q) \supset I$. Since Q is a prime ideal of R/I, it follows that $\phi(Q)$ is a prime ideal of R, and so $\phi(Q) \in C$.

Let $Q, Q' \in \mathrm{Spec}(R/I)$ such that $\phi(Q) = \phi(Q')$. Then

$$\begin{aligned} \overline{a} \in Q &\Leftrightarrow a \in \phi(Q) \\ &\Leftrightarrow a \in \phi(Q') \\ &\Leftrightarrow \overline{a} \in Q', \end{aligned}$$

so ϕ is injective.

To see that ϕ is surjective, let $P \in C$ and let $Q = \{\overline{a} : a \in P\}$. Clearly, if $a \in P$, then $\overline{a} \in Q$. Now if $\overline{a} \in Q$, then $\overline{a} = \overline{b}$ for some $b \in P$, and so $a - b \in I$. Since $P \supset I$, we have $a - b \in P$, and so $a \in P$. It then follows that Q is a prime ideal of R/I, since P is a prime ideal of R, and that $\phi(Q) = P$.

It remains to be proven that ϕ is a homeomorphism. Let C be a closed set of C; then $C = C' \cap C$ for some closed set C' of $\mathrm{Spec}(R)$, so $C = \{P \in \mathrm{Spec}(R) : P \supset I, J\}$ for some ideal J of R. To show that ϕ is continuous, we must show that $\phi^{-1}(C) = \{Q \in \mathrm{Spec}(R/I) : \phi(Q) \in C\}$ is closed in R/I. Let $J' = \langle \overline{a} \in R/I : a \in J \rangle \subset R/I$ and let $Q \in \mathrm{Spec}(R/I)$.

Suppose that $Q \supset J'$. Then if $a \in J$, we have $\overline{a} \in J' \subset Q$, so $a \in \phi(Q)$, and we see that $\phi(Q) \supset J$. Conversely, suppose that $\phi(Q) \supset J$. If $\overline{a} \in J'$, then $\overline{a} = \overline{b}$ for some $b \in J \subset Q$, so $\overline{a} = \overline{b} \in Q$, and we see that $Q \supset J'$. Since $\phi(Q) \supset I$ for all $Q \in \mathrm{Spec}(R/I)$, it follows that $\phi^{-1}(C) = \{Q \in \mathrm{Spec}(R/I) : Q \supset J'\}$, a closed set, so ϕ is continuous.

Now let $\psi : C \to \mathrm{Spec}(R/I)$ be the inverse of ϕ. We saw above that $\psi(P) = \{\overline{a} \in R/I : a \in P\}$ for $P \in C$. Let D be a closed set in R/I. Then $D = \{Q \in \mathrm{Spec}(R/I) : Q \supset K\}$ for some ideal K of R/I. Let $K' = \langle a \in R : \overline{a} \in K \rangle$ and let $P \in C$.

Suppose that $P \supset K'$. Then if $\overline{a} \in K$, we have $a \in K' \subset P$, so $\overline{a} \in \psi(P)$, and thus $\psi(P) \supset K$. Conversely, suppose that $\psi(P) \supset K$. Then if $a \in K'$, we have $\overline{a} \in K \subset \psi(P)$, so $a \in P$, as seen above. Thus $P \subset K'$, and we have shown that $\psi^{-1}(D) = C \cap \{P \in \mathrm{Spec}(R) : P \subset J\}$, which is a closed set of C. Therefore ψ is continuous, and so ϕ is a homeomorphism. □

Corollary 2.4.11. *Let R be a ring, let I be an ideal of R and let $S = R/I$. Then for every $Q \in \mathrm{Spec}(S)$, there exists $P \in \mathrm{Spec}(R)$ with $P \supset I$ such that $R/P \cong S/Q$.*

Proof. Let $Q \in \mathrm{Spec}(S)$ and let $P = \{a \in R : \overline{a} \in Q\}$ were $\overline{a}$ denotes $a + I$. We saw in the proof of Proposition 2.4.10 that $P \in \mathrm{Spec}(R)$ and that $P \supset I$. We have a surjection $R \to S = R/I \to S/Q$ sending a to $\overline{a} + Q$. The kernel of this surjection is P, so by the isomorphism theorem, we have $R/P \cong S/Q$. □

Corollary 2.4.12. *Let R be a ring, let I be an ideal of R, let $S = R/I$, and define the map ϕ with inverse ψ as in the proof of Proposition 2.4.10. Then ϕ gives a bijection between the set of maximal ideals of R/I and the set of maximal ideals of R containing I.*

Proof. Let N be a maximal ideal of R/I and let $M = \phi(N)$. By Corollary 2.4.12, $R/M \cong S/N$, so M is maximal.

Conversely, let M be a maximal ideal of R containing I and let $N = \psi(M) = \{\overline{a} \in S : a \in M\}$. We claim that S/N is a field. To see this, let $\overline{a} + N \in S/N$ with $\overline{a} \notin N$. Then $a \notin M$, so there exists $b \in R$ such that $ab \equiv 1 \pmod{M}$. Thus $ab - 1 \in M$, which implies that $\overline{ab - 1} \in N$, so $\overline{a}\overline{b} \equiv 1 \pmod{N}$. Therefore S/N is a field, and so N is maximal. □

2.5 The Jacobson Radical

In this section we introduce the *Jacobson Radical* of a ring R, a special ideal with many useful properties.

Definition 2.5.1. Let R be a ring. We define the *Jacobson Radical* of R, denoted by $J(R)$, to be the intersection of all maximal ideals of R.

Example 2.5.2. Let K be a field. The maximal ideals of $K[X]$ are of the form $\langle P(X) \rangle$ where $P(X)$ is irreducible. Thus if $Q(X) \in J(K[X])$, then every irreducible $P(X) \in K[X]$ divides $Q(X)$, and so $J(K[X]) = (0)$.

Example 2.5.3. Let $R = \mathbb{Z}/12\mathbb{Z}$. The maximal ideals of R are in bijection with the maximal ideals of $\mathbb{Z}$ containing $12\mathbb{Z}$, namely $2\mathbb{Z}$ and $3\mathbb{Z}$ by Corollary 2.4.12. So we see that R has two maximal ideals, $2R$ and $3R$. Therefore $J(R) = 2R \cap 3R = 6R$.

Example 2.5.4. Let $R = \left\{\frac{a}{2b+1} : a, b \in \mathbb{Z}\right\} \subset \mathbb{Q}$, that is the set of all rational numbers with odd denominators. Then R is a subring of $\mathbb{Q}$. There is a unique maximal ideal of R, namely $M = 2R$. Note that we have a surjective homomorphism,

$$\begin{aligned} \phi : R &\to \mathbb{Z}/2\mathbb{Z} \\ \frac{a}{2b+1} &\mapsto a + 2\mathbb{Z}, \end{aligned}$$

with kernel M. Therefore, by the isomorphism theorem, $R/M \cong \mathbb{Z}/2\mathbb{Z}$, which is a field, and thus M is maximal by Proposition 2.1.13. To see that M is the unique maximal ideal, observe that if I is a proper ideal of R, I cannot contain any elements of the form $\frac{2a+1}{2b+1}$ for $a, b \in \mathbb{Z}$, since these are units. Thus $I \subset M$. Since this is the only maximal ideal, $J(R) = 2R$.

Example 2.5.4 hints at a criterion for deciding whether or not an element of R is in $J(R)$.

Proposition 2.5.5. *Let R be a ring. Then $x \in J(R)$ if and only if $1 + ax \in U(R)$ for all $a \in R$.*

Proof. Suppose that $x \in J(R)$ and $a \in R$. Since $J(R)$ is an ideal, then $ax \in J(R)$. If $1 + ax \notin U(R)$, then $R(1 + ax)$ must be a proper ideal of R, and hence there is a maximal ideal M of R containing $R(1 + ax)$ by Proposition 2.1.8. But since $ax \in J(R)$ and $J(R)$ is the intersection of all maximal ideals, then $ax \in M$, which implies that $1 \in M$, and so $M = R$, a contradiction. Thus $1 + ax \in U(R)$.

Conversely, suppose that $x \in R$ and $1 + ax \in U(R)$ for all $a \in R$. We claim that $x \in J(R)$. If not, then there exists a maximal ideal M of R such that $x \notin M$. Since M is maximal, we have $Rx + M = R$, so there exist $a \in R$ and $b \in M$ such that $ax + b = 1$. Thus $b = 1 + (-a)x$, but b cannot be a unit since $b \in M$, contradicting our choice of x. Therefore we must have $x \in J(R)$. □

We can use Proposition 2.5.5 to show that the Jacobson radical is preserved by surjective ring homomorphisms.

Proposition 2.5.6. *Let R and S be rings and let $\phi : R \to S$ be a surjective ring homomorphism. Then $\phi(J(R)) \subset J(S)$.*

Proof. Let $y \in \phi(J(S))$ and let $b \in S$. Then $y = \phi(x)$ for some $x \in J(R)$ and $b = \phi(a)$ for some $a \in R$. Since $x \in J(R)$, then by Proposition 2.5.5 there exists $r \in R$ such that $(1 + ax)r = 1$. Thus

$$\phi((1 + ax)r) = \phi(1),$$

so

$$(1 + by)\phi(r) = 1,$$

since ϕ is a homomorphism. Thus $1 + by \in U(S)$ for all $b \in S$, so $y \in J(S)$. □

One of the most useful results concerning the Jacobson radical is Nakayama's lemma. If R is a ring, an additive abelian group M is called an *R-module* if M is closed under multiplication by elements of R (which can be defined as any binary operator $R \times M \to M$) and multiplication by elements of R distributes over addition in M.

Theorem 2.5.7 (Nakayama's Lemma, [Nak51])**.** *Let R be a ring and let M be a finitely generated R-module. If $J(R)M = M$, then $M = (0)$.*

Proof. Suppose that $J(R)M = M$, but $M \neq (0)$. Let $\{m_1, \ldots, m_k\}$ be a minimal generating set of M, that is, $M = Rm_1 + \cdots + Rm_k$ and no proper subset of $\{m_1, \ldots, m_k\}$ generates M. Since $M \neq (0)$, $k \geq 1$.

Since $J(R)M = M$, there exist $a_1, \ldots, a_k \in J(R)$ such that

$$m_1 = a_1m_1 + \cdots + a_km_k,$$

so

$$(1 - a_1)m_1 = a_2m_2 + \cdots + a_km_k.$$

By Proposition 2.5.5, $(1 - a_1) \in U(R)$, so there exists $s \in R$ such that $s(1 - a_1) = 1$. This gives

$$m_1 = sa_2m_2 + \cdots + sa_km_k.$$

However, this means that $m_1 \in Rm_2 + \cdots + Rm_k$, so $\{m_2, \ldots, m_k\}$ generates M, which is a contradiction. Thus $M = (0)$. □

The condition that M be finitely generated is needed, as the next example shows.

Example 2.5.8. Let $R = \left\{\frac{a}{2b+1} : a, b \in \mathbb{Z}\right\}$ as in Example 2.5.4 and let $M = \mathbb{Q}$. Then $\mathbb{Q}$ is an R-module by ordinary multiplication. In Example 2.5.4 we showed that $J(R) = 2R$, so $J(R)\mathbb{Q} = 2R\mathbb{Q} = \mathbb{Q}$, but of course $\mathbb{Q} \neq (0)$.

Nakayama's lemma does not apply because $\mathbb{Q}$ is not finitely generated as an R-module. To see this, suppose that $\frac{c_1}{d_1}, \ldots, \frac{c_k}{d_k} \in \mathbb{Q}$, where $\gcd(c_i, d_i) = 1$ for each $1 \geq i \geq k$, generates $\mathbb{Q}$ as an R-module. Let $l = \operatorname{lcm}(d_1, \ldots, d_k)$. Then there exists N such that $2^N \mid l$, but $2^{N+1} \nmid l$. Now let $\frac{a_1}{2b_1+1}, \ldots, \frac{a_k}{2b_k+1} \in R$ and let $l' = \operatorname{lcm}((2b_1 + 1)d_1, \ldots, (2b_k + 1)d_k)$. Then $2^{N+1} \nmid l'$ since each $2b_i + 1$ is odd. Therefore there are no $\frac{a_1}{2b_1+1}, \ldots, \frac{a_k}{2b_k+1} \in R$ such that

$$\frac{a_1}{2b_1 + 1}\frac{c_1}{d_1} + \cdots + \frac{a_k}{2b_k + 1}\frac{c_k}{d_k} = \frac{1}{2^{N+1}}.$$

We use the Jacobson radical to define a class of rings that we will be concerned with in the next section.

Definition 2.5.9. We say that a ring R is a *Jacobson ring* if $J(S) = (0)$ whenever S is a prime homomorphic image of R, that is, $S \cong R/P$ for some prime ideal of P of R.

Example 2.5.10. If R is a field, then R is a Jacobson ring, since (0) is the only proper ideal of R and $J(R) = (0)$.

Example 2.5.11. Let $R = \mathbb{Z}$. Then R is Jacobson. To see this, note that if S is a prime homomorphic image of R, then either $S \cong \mathbb{Z}$ or $S \cong \mathbb{Z}/p\mathbb{Z}$ for some prime p. If p is prime, then $\mathbb{Z}/p\mathbb{Z}$ is a field, and so $J(\mathbb{Z}/p\mathbb{Z}) = (0)$. On the other hand, $J(\mathbb{Z}) = \bigcap_{p \text{ prime}} p\mathbb{Z} = (0)$, so R is Jacobson.

Example 2.5.12. The ring $R = \left\{\frac{a}{2b+1} : a, b \in \mathbb{Z}\right\}$ from Example 2.5.4 is not a Jacobson ring. Note that R is an integral domain, so R is a prime homomorphic image of itself. Recall from Example 2.5.4 that $J(R) = 2R \neq (0)$, so R is not Jacobson.

The next two results give basic closure properties of Jacobson rings.

Proposition 2.5.13. *Let R be a Jacobson ring and I be an ideal of R. Then R/I is Jacobson.*

Proof. Let $S = R/I$ and let $Q \in \operatorname{Spec}(S)$. By Proposition 2.4.10, Q corresponds uniquely to a prime ideal of R containing I, namely $P = \{a \in R : a + I \in Q\}$. This gives us a natural isomorphism $S/Q \to R/P$ defined by $(a + I) + Q \mapsto a + R$. If R is Jacobson, then $J(S/Q) = J(R/P) = (0)$. Thus S is Jacobson. □

In the proof of Proposition 2.5.13, it is in fact sufficient to see that the map given is surjective and then apply Proposition 2.5.6.

Example 2.5.14. Let $R = \mathbb{Z}/12\mathbb{Z}$. Recall from Example 2.5.3 that $J(R) = 6R$. This does not mean that R is not Jacobson, as in Example 2.5.12, since R is not an integral domain. The prime ideals of R are $2R$ and $3R$. These are both maximal, so every prime homomorphic image of R is a field and so has Jacobson radical (0). Thus R is Jacobson.

Proposition 2.5.15. *If R is a Jacobson ring, then $R[X]$ is also Jacobson.*

Proposition 2.5.15 and its converse were proved in [Gol51]. We omit the proof, but note that the result is a special case of Theorem 2.6.5, the Nullstellensatz.

Example 2.5.16. We saw in Example 2.5.2 that $J(K[X]) = (0)$. We have seen that every prime ideal of $K[X]$ is of the form $\langle P(X)\rangle$ for some irreducible $P(X) \in K[X]$. It follows that every prime ideal of $K[X]$ is maximal. Thus if P is a prime ideal of $K[X]$, then $K[X]/P$ is a field, and so $J(K[X]/P) = (0)$. Therefore $K[X]$ is Jacobson.

Let R and S be rings with $R \subset S$ and let $y \in S$. We define the *R-algebra generated by y* as

$$R[y] \coloneqq \{a_0 + a_1 y + \cdots + a_n y^n : n \geq 0, a_0, \ldots, a_n \in R\}.$$

It is not hard to see that $R[y]$ is the smallest subring of S containing R and y. In particular, if S is the total ring of fractions of R and $x \in R$ is not a zero divisor, then $R[x^{-1}]$ is the smallest ring containing R in which x is a unit.

We can use this construction to give a useful characterization of Jacobson rings. First, however, we will prove a lemma concerning $\operatorname{Spec}\left(R[x^{-1}]\right)$.

Lemma 2.5.17. *Let R be a ring and let $x \in R$ such that x is not a zero divisor. Then there is a homeomorphism*

$$\phi : \operatorname{Spec}\left(R[x^{-1}]\right) \to \mathcal{U} \coloneqq \{P \in \operatorname{Spec}(R) : x \notin P\}.$$

Proof. We note that if $b \in R[x^{-1}]$, then $x^d b \in R$ for some $d \geq 0$. Let $Q \in \operatorname{Spec}\left(R[x^{-1}]\right)$ and define

$$\phi(Q) = \left\{a \in R : x^d a = x^d b \text{ for some } b \in Q, d \geq 0\right\}.$$

We claim that $\phi(Q) = Q \cap R$. To see this, let $x \in \phi(Q)$. Then $x^d a = x^d b$ for some $b \in Q$ and $d \geq 0$, so $x^d a \in Q$. Since Q is prime, we have either $a \in Q$ or $x \in Q$, but $x \notin Q$ since x is a unit in $R[x^{-1}]$. Therefore $a \in Q \cap R$. We see that $\phi(Q) \subset Q \cap R$ by construction, so $\phi(Q) = Q \cap R$. It follows that $\phi(R) \in \text{Spec}(R)$ and that $a \in \phi(Q)$, so ϕ is well-defined.

To see that ϕ is injective, let $Q, Q' \in \text{Spec}\left(R[x^{-1}]\right)$ such that $\phi(Q) = \phi(Q')$. Then $Q \cap R = Q' \cap R$. Let $\frac{a}{x^d} \in Q$ where $a \in R$. Now Q is an ideal of $R[x^{-1}]$, so $a \in Q \cap R$, and thus $a \in Q' \cap R$, which implies that $\frac{a}{x^d} \in Q'$. Similarly, we can show that $Q' \subset Q$, so we have $Q = Q'$.

Now we will show that ϕ is surjective. Let $P \in \mathcal{U}$, and let $Q = P\left(R[x^{-1}]\right)$. We claim that $Q \cap R = P$. We have $P \subset Q \cap R$ by construction. Let $b \in Q \cap R$. We have $0 \in P$, so assume that $b \neq 0$. Then $b = \frac{r_1}{x^{d_1}} a_1 + \cdots \frac{r_k}{x^{d_k}} a_k$ where $k \geq 1$ and $r_i \in R, d_i \geq 0$ for each $i = 1, \ldots, k$. Adding, we get $b = \frac{a}{x^d}$ for some $a \in P$ (since P is an ideal) and $d \geq 0$. Therefore $x^d b \in P$, so $b \in P$ since $b \in R$, P is prime and $x \notin P$. Therefore $\phi(Q) = Q \cap R = P$.

We have shown that ϕ is a bijection, so it remains to show that ϕ and its inverse are continuous. Let C be a closed subset of $\mathcal{U}$. Then

$$\begin{aligned} C &= \mathcal{U} \cap \{P \in \text{Spec}(R) : P \supset I\} \\ &= \{P \in \text{Spec}(R) : P \not\supset (a), P \supset I\} \end{aligned}$$

for some ideal I of R. Thus

$$\begin{aligned} \phi^{-1}(C) &= \left\{Q \in \text{Spec}\left(R[x^{-1}]\right) : a \notin \phi(Q), \phi(Q) \supset I\right\} \\ &= \left\{Q \in \text{Spec}\left(R[x^{-1}]\right) : \phi(Q) \supset I\right\}. \end{aligned}$$

Let $J = I\left(R[x^{-1}]\right)$. We note that $J \cap R = I$ by an earlier argument. Now by the same argument used to prove that ϕ is injective, we see that $Q \supset J$ if and only if $\phi(Q) \supset I$. Therefore $\phi^{-1}(C) = \left\{Q \in \text{Spec}\left(R[x^{-1}]\right) : Q \supset J\right\}$, which is closed in $\text{Spec}\left(R[x^{-1}]\right)$, so ϕ is continuous.

Now let ψ be the inverse of ϕ and let D be a closed subset of $\text{Spec}\left(R[x^{-1}]\right)$. Then $D = \left\{Q \in \text{Spec}\left(R[x^{-1}]\right) : Q \supset J\right\}$ for some ideal J of $R[x^{-1}]$. Let $I = J \cap R$. Then it follows from earlier arguments that

$$\psi^{-1}(D) = \mathcal{U} \cap \{P \in \text{Spec}(R) : P \supset I\},$$

which is closed. Therefore ϕ is a homeomorphism. □

Proposition 2.5.18. *Let R be a ring. Then the following are equivalent:*

(1) R is Jacobson;

(2) if $P \in \mathrm{Spec}(R)$ and $S \cong R/P$ has the property that there exists $x \in S \setminus \{0\}$ such that $S[x^{-1}]$ is a field, then S is a field.

Proof. Suppose that R is a Jacobson ring and suppose that S is a prime homomorphic image of R and that $S[x^{-1}]$ is a field for some $x \in S \setminus \{0\}$. Suppose that S is not a field and let M be a maximal ideal of S.

Since S is not a field, then $M \neq (0)$, so there exists $a \in M \setminus \{0\}$. Now $S[x^{-1}]$ is a field, so there exist $s \in S$ and $d \geq 0$ such that $a\frac{s}{x^d} = 1$, which implies that $x^d = as \in M$. Now M is prime, since it is maximal, so $x \in M$. This argument holds for any maximal ideal of S, so $x \in J(S)$, a contradiction.

Conversely, suppose that (2) holds and that R is not Jacobson. Since R is not Jacobson, there exists $P \in \mathrm{Spec}(R)$ such that $J(R/P) \neq (0)$. Let $T = R/P$. Then there exists $y \in J(T) \setminus \{0\}$. Let $\mathcal{S} = \{1, y, y^2, \ldots\}$. We saw in the proof of Proposition 2.4.4 that the nilradical $N := \mathrm{Nil}(T)$ is maximal with respect to the property that $N \cap \mathcal{S} = \emptyset$. Moreover N is prime.

Let $S = T/N$ and let $Q = \{a \in R : a + P \in N\}$. Then Q is a prime ideal of R and $S \cong P/Q$ by arguments used in the proof of Proposition 2.5.13.

Let x denote the image of y in S. Then $x \neq 0$ since $y \notin Q$. We claim that $S[x^{-1}]$ is a field. To see this, note that there is a bijection i between $\mathrm{Spec}(S)$ and prime ideals in T containing Q. Thus if $P' \in \mathrm{Spec}(S) \setminus \{(0)\}$, then there exists $P \in \mathrm{Spec}(T)$ such that $i(P) = P'$, and so $P \supsetneq Q$. This implies that $y^d \in P$ for some $d > 0$, so $y \in P$ since P is prime. It follows from Lemma 2.5.17 that $\mathrm{Spec}(S[x^{-1}]) = \{(0)\}$, so $S[x^{-1}]$ is a field, and therefore S is a field by our hypothesis.

Now $y \in J(T)$, so $x \in J(S)$ by Lemma 2.5.6. But S is a field, so $J(S) = (0)$, which means that $x = 0$, a contradiction. Therefore $J(T) = (0)$, so R is Jacobson. □

2.6 The Nullstellensatz

In this section we will prove the generalized form of Hilbert's Nullstellensatz. The Nullstellensatz was originally proved by Hilbert [Hil93] in the case where R is a polynomial

ring over an algebraically closed field. We will prove a more general version that applies to the class of all Jacobson rings.

Definition 2.6.1. Let R be a ring and let S be an R-algebra. We say that S is *integral* over R if for every $s \in S$, there exist $d \geq 0$ and $r_0, \ldots, r_d$ such that

$$s^{d+1} = r_0 + r_1 s + \cdots + r_d s^d.$$

Example 2.6.2. Let $R = \mathbb{Z}$ and let $S = \mathbb{Z}[i]$, where $i^2 = -1$. Then S is integral over R. To see this, we note that if $s \in S$, then $s = a + ib$ for some $a, b \in R$. Then

$$\begin{aligned} s^2 &= (a + ib)^2 \\ &= a^2 + 2abi - b^2 \\ &= 2a^2 + 2abi - a^2 - b^2 \\ &= 2as - (a^2 + b^2). \end{aligned}$$

Example 2.6.3. Let $R = \mathbb{Z}$ and $S = \mathbb{Q}$. Then S is not integral over R. If it were, then there would exist $r_0, \ldots, r_d \in \mathbb{Z}$ such that

$$\frac{1}{2^{d+1}} = r_0 + \frac{r_1}{2} + \cdots + \frac{r_d}{2^d}.$$

which implies

$$\begin{aligned} 1 &= 2^{d+1} r_0 + 2^d r_1 + \cdots + 2r_d \\ &= 2(2^d r_0 + 2^{d-1} r_1 + \cdots + r_d) \end{aligned}$$

a contradiction since $2^d r_0 + 2^{d-1} r_1 + \cdots + r_d \in \mathbb{Z}$.

Lemma 2.6.4. *Let R and S be integral domains. If R is a field and S is integral over R, then S is a field.*

Proof. Suppose R is a field and that S is integral over R. Let $s \in S \setminus \{0\}$. Since S is integral over R, there exists $d \geq 0$ and $r_0, \ldots, r_d \in R$ such that

$$s^{d+1} = r_0 + r_1 s + \cdots + r_d s^d.$$

Since s is nonzero and S is an integral domain, we can assume without loss of generality that $r_0 \neq 0$.

Now R is a field, so there exists $x \in R$ such that

$$\begin{aligned} 1 = xr_0 &= x\left(s^{d+1} - r_d s^d - \cdots - sr_1\right) \\ &= s\left(xs^d - xr_d s^{d-1} - \cdots - xr_1\right). \end{aligned}$$

Since S is an R-algebra, $x \in S$. Therefore $x \in U(S)$ for all $s \in S \setminus \{0\}$, so S is a field. □

We can now prove the generalized Nullstellensatz.

Theorem 2.6.5 (The Nullstellensatz)**.** *Let R be a Jacobson ring and let S be a finitely generated R-algebra. Then S is a Jacobson ring. Furthermore, if N is a maximal ideal of S, then $M \coloneqq N \cap R$ is a maximal ideal of R, and S/N is a finite extension of R/M.*

Proof. We will first prove this in the special case where R is a field and $S = R[X]$. In this case, S is a principal ideal domain, and so every prime ideal is maximal. Thus S/Q is a field for all $Q \in \operatorname{Spec}(S)$.

Let $P(X) \in J(S)$. If $B(X) \in S$ is irreducible, then $(B(X))$ is a maximal ideal of S, and so $P(X) \in (B(X))$. Thus every irreducible polynomial in S divides $P(X)$. We claim that there are infinitely many irreducible polynomials in S. To see this, suppose towards a contradiction that there are finitely many irreducible polynomials in $R[X]$, $B_1(X), \ldots, B_t(X)$. However, none of these divides $\left(\prod_{i=1}^{t} B_i(X)\right) + 1$, a contradiction. Thus $P(X)$ has infinitely many irreducible factors, so $P(X) = 0$. Therefore S is Jacobson.

Now suppose that N is a maximal ideal of S. Then $N = (A(X))$ for some irreducible polynomial $A(X)$. Now if $M \coloneqq N \cap R = R$, we have $1 \in N$, which implies that $N = S$, a contradiction. Thus $M = (0)$ since R is a field, and so M is maximal, and $[S/N : R]$ is equal to the degree of $A(X)$, so we have proved the theorem in this case.

Now suppose that R is any Jacobson ring and suppose that S is generated as an R-algebra by one element, t. To show that S is Jacobson, we will use Proposition 2.5.18. Let $P \in \operatorname{Spec}(S)$. We can replace S by S/P and R by $R/(P \cap R)$ and thus assume that R is an integral domain since $P \cap R$ is prime. To prove the first statement of the theorem, we must show that if there exists $b \in S$ such that $S[b^{-1}]$ is a field, then S is a field. We can in fact show that R itself is a field and that S is a finite extension of R.

Suppose there exists $b \in S$ such that $S[b^{-1}]$ is a field. Since S is generated by t over R, there is some $Q \in \operatorname{Spec}(R[X])$ such that $S \cong R[X]/Q$. Indeed, Q is the kernel of the surjective homomorphism $S \to R[X]$ sending t to X.

We claim that $Q \neq (0)$. To see this, suppose that $Q = (0)$. Then $R[X] \cong S$, so we have an element $a \in R[X]$ such that $R[X][a^{-1}]$ is a field. If we let K be the field of fractions of R, then $K[X][a^{-1}]$ is a field as well. But $K[X]$ is Jacobson, so $K[X]$ is a field by Proposition 2.5.18, a contradiction. Thus $Q \neq (0)$, and so $S[b^{-1}] \cong K[X]/QK[X]$. Thus $S[b^{-1}]$ is finite-dimensional over K by an argument similar to that used in the first special case above.

Now let $P(X) = p_m X^m + \cdots + p_0 \in Q$. Then

$$P(t) = p_m t^m + \cdots + p_0 = 0$$

in S. We can invert p_m and get

$$p_m^{-1} P(t) = t^m + p_m^{-1} p_{m-1} t^{m-1} \cdots + p_m^{-1} p_0 = 0,$$

so t is integral over $R[p_m^{-1}]$, which implies that S is integral over $R[p_m^{-1}]$. Now there exists $q_0, \ldots, q_n \in R$ such that

$$Q(b) := q_n b^n + \cdots + q_0 = 0.$$

We may divide by a power of b if necessary, since S is an integral domain, and assume that $q_0 \neq 0$. We multiply $Q(b)$ by $(q_0 b^n)^{-1}$ to obtain

$$q_n q_0^{-1} + \cdots + q_1 q_0^{-1} (b^{-1})^{m-1} + (b^{-1})^m = 0.$$

Thus b^{-1} is integral over $R[q_0^{-1}]$, and so $S[b^{-1}]$ is integral over $R[(p_m q_0)^{-1}]$. Since $S[b^{-1}]$ is a field by assumption, Lemma 2.6.4 implies that $R[(p_m q_0)^{-1}]$ is a field, and so R is a field by Proposition 2.5.18. This means that $R = K$, so S is a finite extension of R.

If we use the same reduction but assume that S is a field, the argument above is sufficient to prove the second statement. Note that any nonzero element of S can be chosen as b and that $S[b^{-1}] = S$ when S is a field.

We can now prove the result in the general case where S is generated as an R-algebra by r elements. Suppose that $r > 1$ and that the result has been proved for all R-algebras with at most $r - 1$ generators. Choose $r - 1$ of the generators of S and let S' be the R-algebra generated by them. Then S is an S'-algebra generated by one element. Now S' is Jacobson by induction, and so S is Jacobson by the case proved above.

If N is a maximal ideal of S, then $S' \cap N$ is maximal in S' by the $r = 1$ case. By induction, $R \cap N$ is maximal in R and $S'/(S' \cap N)$ is a finite extension of $R/(R \cap N)$. By the $r = 1$ case, S/N is a finite extension of $S'/(S' \cap N)$, so it follows that S/N is a finite extension of $R/(R \cap N)$. □

Chapter 3

Automatic Sequences

In this chapter we discuss sequences that can be computed by an automatic process, specifically by deterministic finite automata (DFA, for short), a simple computational model. These automata take input in the form of strings over a given set of symbols, and in the case that we are interested in, return output from another set of symbols. When we say that a sequence $(a_n)_{n\geq 0}$ is computed by an automaton, we mean that when the automaton is given a string representation of n as input, it gives a_n as output.

McCulloch and Pitts [MP90] developed a formal computational model to use in the study of neural activity that was basically the same as modern finite automata. Later, Huffman [Huf54], Mealy [Mea55] and Moore [Moo56] wrote about computational models equivalent to the finite automata we discuss. We follow the approach of Allouche and Shallit [AS03].

Sequences that can be computed by automata are called automatic sequences.

3.1 Strings

Computations by DFAs are performed on strings, so we will give a brief explanation of the basic terms. An *alphabet* is a set of *symbols*; we will consider only finite alphabets. Symbols can really be anything, but we will be looking mainly at alphabets made up of integers, though letters are commonly used as well. The Greek letter Σ is commonly used to denote an alphabet. Symbols are usually displayed in a different typeface, like this: `0`,`1`,`2`. Often, however, we will identify symbols with the integers that they represent, for example,

we might identify 0 with 0.

Let k be a positive integer. We define

$$\Sigma_k = \{0, 1, \ldots, k-1\}.$$

We will mainly be concerned with alphabets of this form since for any given $k \geq 2$ we can represent positive integer as strings over Σ_k. We will discuss this in Section 3.4.

A *string* or *word* is a sequence of symbols from some alphabet. More formally, a finite string w over an alphabet Σ is a function $w : \{0, \ldots, n-1\} \to \Sigma$. We write w as $w(0)w(1)\cdots w(n-1)$, or, more commonly, as $w_0 w_1 \cdots w_{n-1}$, where $w_i = w(i)$.

Example 3.1.1. The word $w = 012$ over Σ_3 is the function

$$\begin{aligned} w : \{0,1,2\} &\to \Sigma_3 \\ 0 &\mapsto 0 \\ 1 &\mapsto 1 \\ 2 &\mapsto 2. \end{aligned}$$

So $w = w_0 w_1 w_2$ where $w_0 = 0$, $w_1 = 1$ and $w_2 = 2$.

Infinite strings are possible as well, though we will generally call them *infinite sequences*, or often just sequences. An infinite sequence or string over an alphabet Σ is, formally, a map $\mathbf{u} : \{0, 1, 2, \ldots\} \to \Sigma$. As above, we often write $u_n = u(n)$, and we write $\mathbf{u} = (u_n)_{n \geq 0}$ or sometimes $\mathbf{u} = u_0 u_1 \cdots$. We will assume that all strings are finite unless stated otherwise.

A set of words over a given alphabet is called a *language*. We denote by Σ^* the language made up of all strings over Σ, including the empty string with no characters, which we denote by ϵ.

Two strings can be *concatenated* by juxtaposing their symbols, and we write wx for the concatenation of strings w and x. For an integer $t \geq 0$, we define

$$w^t = \underbrace{ww\cdots w}_{t \text{ times}}.$$

In the case where $t = 0$, we have $w^0 = \epsilon$. We define w^ω to be the infinite string (or sequence) $www\cdots$.

We say that x is a *substring* of w if $w = yxz$ for some strings y and z. The number of characters in a string is called its *length* and is denoted $|w|$.

If $w = w_0 w_1 \cdots w_t$, then we define $w^R = w_t w_{t-1} \cdots w_0$, the string made up of the symbols in w in reverse order.

Example 3.1.2. Let $w = \texttt{dog}$ and $x = \texttt{cow}$, then $wx = \texttt{dogcow}$, $x^3 = \texttt{cowcowcow}$, $\texttt{do}$ is a substring of w, $|wx| = 6$, and $w^R = \texttt{god}$.

3.2 Deterministic Finite Automata

A deterministic finite automaton, or DFA for short, is a computational model that takes a string as input that is either accepted or rejected.

A DFA consists of a finite number of *states* and can be in exactly one state at a time. The DFA is given an input string w, which is read one symbol at a time from left to right. One state is chosen as an initial state, and the DFA moves from state to state based on the symbols read from the input. Some of the states are designated as *accepting states*. If the DFA is in one of these accepting states after reading the entire input, we say that w is accepted by the DFA, otherwise we say it is rejected by the DFA. We now give a more formal definition.

Definition 3.2.1. A *deterministic finite automaton* is a 5-tuple

$$M = (Q, \Sigma, \delta, q_0, F)$$

where Q is finite set of states, Σ is a finite *input alphabet*, $\delta : Q \times \Sigma \to Q$ is the *transition function*, $q_0 \in Q$ is the *initial state*, and $F \subset Q$ is a set of *accepting states*.

We can extend the transition function δ to $Q \times \Sigma^*$ by setting $\delta(q, \epsilon) = q$, for all $q \in Q$, and setting $\delta(q, xw) = \delta(\delta(q, x), w)$ for all $q \in Q$, $x \in \Sigma$, and $w \in \Sigma^*$. Given a string $w \in \Sigma^*$, we say that w is *accepted* by a finite automaton $M = (Q, \Sigma, \delta, q_0, F)$ if $\delta(q_0, w) \in F$. If $\delta(q_0, w) \notin F$, then we say that w is *rejected* by M. The set of all strings over Σ that are accepted by a DFA M is called the language accepted or generated by M, written $L(M)$.

We often represent a DFA visually with a *transition diagram*. This is a directed graph in which the vertices correspond to the states of the DFA. Edges are drawn to represent transition between states, with labels indicating the symbol read. The initial state is indicated by an unlabelled arrow entering the state and the accepting states by double outlines.

Example 3.2.2. For example, Figure 3.1 depicts a DFA that accepts all binary strings containing no consecutive 1s. As the DFA reads a string, it will move to state q_1 upon reading a 1, but move back to the initial state q_0 if the next bit is a 0. Two consecutive 1s will cause the DFA to move to q_2, and since there are no transitions out of q_2, it will stay there. Thus the accepting states are q_0 and q_1. If we call this DFA M, we have

$$L(M) = \{w \in \{0, 1\}^* : 11 \text{ is not a substring of } w\}.$$

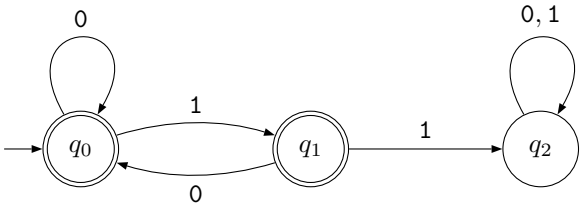

Figure 3.1: A DFA that accepts strings with no consecutive 1s.

Example 3.2.3. Figure 3.2 shows a DFA that accepts binary strings whose second to last bit is 1. The states are labelled to indicate the last two bits read from the input string (this can be verified by the reader). After reading an input string, the DFA will be in the state corresponding to its last two bits, so the accepting states are 10 and 11.

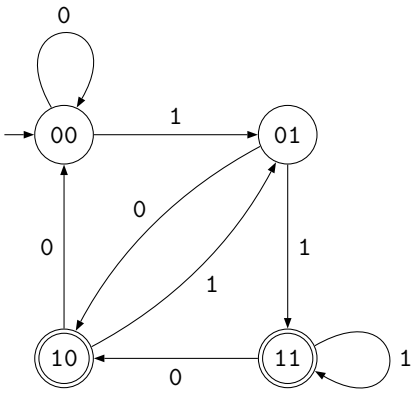

Figure 3.2: A DFA that accepts strings ending in 10 or 11.

3.3 Deterministic Finite Automata with Output

Let M be a DFA with input alphabet Σ. Given any string in Σ^*, M will either accept it or reject it, so we can think of M as computing a function $f_M : \Sigma^* \to \{0, 1\}$, where

$$f(w) = \begin{cases} 1, & \text{if } w \text{ is accepted by } M; \\ 0, & \text{if } w \text{ is not accepted by } M. \end{cases}$$

for all $w \in \Sigma^*$. Expanding on this idea, we can build automata with output from any set of symbols. Instead of designating certain states as accepting states, we can associate a specific output to each state. In this way we can, if we wish, have as many outputs as there are states in the automaton.

Definition 3.3.1. A *deterministic finite automaton with output*, or DFAO, is a 6-tuple

$$M = (Q, \Sigma, \delta, q_0, \Delta, \tau)$$

where Q is finite set of states, Σ is a finite input alphabet, $\delta : Q \times \Sigma \to Q$ is the *transition function*, $q_0 \in Q$ is the *initial state*, Δ is the output alphabet, and $\tau : Q \to \Delta$ is the *output function*.

As before, a DFAO M starts in the initial state q_0 and moves from state to state based on the symbols read left to right from an input string $w \in \Sigma^*$. Once M has read all of w, it will be in some state q. Instead of accepting or rejecting w, M gives an output symbol from Δ, namely $\tau(q)$. Thus we can say that M computes a function from Σ^* to Δ. If we extend δ to $Q \times \Sigma^*$ in the same way as for DFAs, we can formally say that M computes the function $f_M : \Sigma^* \to \Delta$ defined by

$$f_M(w) = \tau(\delta(q_0, w)).$$

Any function that can be computed by a DFAO is called a *finite-state function*.

We draw transition diagrams for DFAOs in the same way as for DFAs, indicating the output $a \in \Delta$ associated to each state q inside the state as q/a.

Example 3.3.2. Figure 3.3 shows a DFAO that takes a binary string and computes the sum of the digits mod 2. If we call the DFAO M, then we can write

$$f_M(x_1 x_2 \cdots x_t) = (x_1 + x_2 + \cdots + x_t) \bmod 2$$

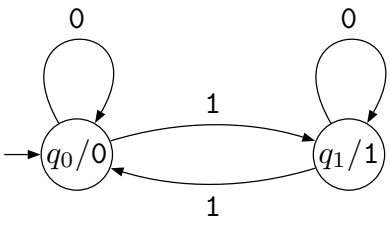

Figure 3.3: A DFAO that computes the sum of the digits in the input mod 2.

Example 3.3.3. In Figure 3.4 we see a DFAO that takes as input a binary string and computes the number of trailing 0s mod 3. Each state q_0, q_1, q_2 corresponds to the output 0,1,2, respectively, and the DFAO cycles through them as it reads 0s from the input, thus counting the 0s mod 3. When a 1 is encountered, the DFAO moves to q_0, resetting the count of 0s.

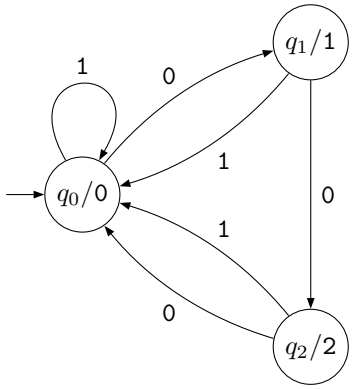

Figure 3.4: A DFAO that computes the number of trailing 0s in the input mod 3.

3.4 Representation of Integers

We cannot give an integer directly as input to a DFAO, but we can represent any integer as a string, and then give the string representation to a DFAO with an appropriate input alphabet.

Let $k \geq 2$ be an integer. Recall that we defined $\Sigma_k = \{0, 1, \ldots, k-1\}$. From here on, we will be fairly loose about the distinction between a symbol and the integer it represents.

Definition 3.4.1. Let $w = a_0 a_1 \cdots a_t$ be a string over Σ_k. We define

$$[w]_k \coloneqq \sum_{i=0}^{t} a_i k^{t-i}.$$

(Here we are actually taking the integer represented by a_i). If $[w]_k = n$, then we say that w is a *base-k representation* of n.

For example, $[111]_2 = 7$ and $[111]_3 = 10$. Note that we also have $7 = [0111]_2 = [00111]_2 = [000111]_2 = \cdots$.

Definition 3.4.2. Let n and k be non-negative integers with $k \geq 2$. Then we can write $n = \sum_{i=0}^{s} a_i k^i$ where $a_i \in \Sigma_k$ for each $0 \leq i \leq s$ and $a_s \neq 0$. Moreover, s and $a_0, \ldots, a_s$ are unique. The *canonical base-k representation* of n is $(n)_k \coloneqq a_s \ldots a_0$.

We note that $[(n)_k]_k = n$, but we do not always have $([w]_k)_k = w$. For example, $([00101]_2)_k = 101$. Given a string $w \in \Sigma_k$, it is easy to see that $[w]_k = n$ if and only if $w = 0^t(n)_k$ for some $t \geq 0$. We should also point out that $(0)_k = \epsilon$, even though 0 is a more familiar representation of 0.

It is possible to extend Definition 3.4.2 to include $k = 1$ by taking

$$(n)_1 = 1^n = \underbrace{11\cdots1}_{n \text{ times}}.$$

This is interesting in that it gives the simplest numeration system possible. It will not, however, be useful here since automata with input alphabet $\{1\}$ have more limited behaviour than those with larger input alphabets (such as Σ_k with $k \geq 0$).

3.5 Automatic Sequences

We are now ready to look at sequences that can be generated by automata. For convenience, a DFAO whose input alphabet is Σ_k for some integer $k \geq 2$ is called a k-DFAO.

Definition 3.5.1. Let $(a_n)_{n\geq 0}$ be a sequence over a finite alphabet Δ. We say that $(a_n)_{n\geq 0}$ is *k-automatic* if there exists a k-DFAO $M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ such that $a_n = f_M(w)$ for all $n \geq 0$ and all $w = x_1 x_2 \ldots x_t$ with $[w]_k = n$. We then say that M *generates* the sequence $(a_n)_{n\geq 0}$.

While the above definition requires that M output a_n no matter what base-k representation of n is given as input, it is sufficient that M gives the correct output for only canonical representation of n.

Proposition 3.5.2. *[AS03] The sequence* $\mathbf{a} = (a_n)_{n\geq 0}$ *is k-automatic if and only if there exists a k-DFAO M such that $f_M((n)_k) = a_n$ for all $n \geq 0$. Furthermore, we can choose $M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ such that $\delta(q_0, 0) = 0$.*

Proof. If $\mathbf{a}$ is k-automatic, then from the definition there exists a DFAO that will compute a_n given $(n)_k$ as input. Conversely, suppose there exists a DFAO $M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ such that $f_M((n)_k) = \tau(\delta((n)_k, q_0)) = a_n$ for all $n \geq 0$. Define a new DFAO $M' = (Q', \Sigma_k, \delta', q_0', \Delta, \tau')$ with

$$\begin{aligned} Q' &= Q \cup \{q_0'\}, \\ \delta'(q, a) &= \delta(q, a) \qquad \text{for all } q \in Q, a \in \Sigma_k, \\ \delta'(q_0', a) &= \begin{cases} \delta(q_o, a) & \text{if } a \neq 0, \\ q_0' & \text{if } a = 0, \end{cases} \\ \tau'(q_0') &= \tau(q_0), \\ \tau'(q) &= \tau(q) \qquad \text{for all } q \in Q. \end{aligned}$$

Let $w \in \Sigma_k$ such that $[w]_k = n$, then $w = \mathbb{0}^i(n)_k$ for some $i \geq 0$. By our construction of M',

$$\delta'(q_0', w) = \delta'(q_0', \mathbb{0}^i(n)_k) = \delta'(q_0', (n)_k).$$

If $n = 0$, then $(n)_k = \epsilon$, so $\delta'(q_0', (n)_k) = q_0'$ and $\delta(q_0, (n)_k) = q_0$. Since $\tau'(q_0') = \tau(q_0)$, we have $\tau'(\delta'(q_0', (n)_k)) = \tau(\delta(q_0, (n)_k))$.

If $n \neq 0$, then $(n)_k = ax$ for some $a \in \Sigma_k \setminus \{0\}, x \in \Sigma_k^*$. Thus

$$\delta'(q_0', (n)_k) = \delta'(q_0', ax) = \delta(\delta(q_0, a), x) = \delta(q_0, ax) = \delta(q_0, (n)_k),$$

so

$$\tau'(\delta'(q_0', (n)_k)) = \tau'(\delta(q_0, (n)_k)) = \tau(\delta(q_0, (n)_k))$$

since M will never move back to q_0' once it moves away from it.

Thus, in all cases,

$$\begin{aligned} f_{M'}(w) &= \tau'(\delta'(q_0', w)) \\ &= \tau'(\delta'(q_0', (n)_k)) \\ &= \tau(\delta(q_0, (n)_k)) \\ &= f_M((n)_k) \\ &= a_n. \end{aligned}$$

□

We can also allow the automaton to generate a sequence by reading the representation of each n starting with the least significant digit, which is sometimes more convenient. We will use this in Section 3.7. Recall that for a string $w = w_0 w_1 \cdots w_t$, we define $w^R = w_t w_{t-1} \cdots w_0$.

Theorem 3.5.3. *[AS03] The following are equivalent:*

(1) $(a_n)_{n\geq 0}$ *is a k-automatic sequence;*

(2) there exists a k-DFAO M *such that* $a_n = f_M(w^R)$ *for all* $n \geq 0$ *and* $w \in \Sigma_k$ *such that* $[w]_k = n$*;*

(3) there exists a k-DFAO M *such that* $a_n = f_M((n)_k^R)$ *for all* $n \geq 0$*.*

Proof. See [AS03, Theorem 5.2.3]. □

Example 3.5.4 (The Thue-Morse Sequence)**.** The sequence $\mathbf{t} = (t_n)_{n\geq 0}$, called the Thue-Morse sequence, counts the number of 1s modulo 2 in the base-2 representation of n. The first twenty terms are

$$\mathbf{t} = (\mathtt{0}, 1, 1, \mathtt{0}, 1, \mathtt{0}, \mathtt{0}, 1, 1, \mathtt{0}, \mathtt{0}, 1, \mathtt{0}, 1, 1, \mathtt{0}, 1, \mathtt{0}, \mathtt{0}, 1, \mathtt{0}, \ldots).$$

This sequence was first described by Thue in [Thu12], and then independently by Morse in [Mor21], although it appears implicitly in [Pro51].

We recall that the automaton from Example 3.3.2 (Figure 3.5) computes the sum of bits in its input string modulo 2, which is equivalent to counting the number of 1s modulo 2. Thus, when given the binary expansion of n, the automaton computes t_n, so the Thue-Morse sequence is 2-automatic.

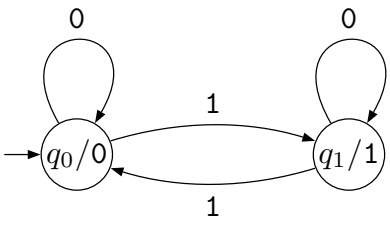

Figure 3.5: A DFAO that computes the Thue-Morse sequence.

Example 3.5.5 (The Rudin-Shapiro Sequence)**.** The Rudin-Shapiro sequence is given by $\mathbf{r} = (r_n)_{n\geq 0}$, where $r_n = 1$ if the number of (possibly overlapping) occurrences of 11 in the binary expansion of n is even, and $r_n = -1$ if it is odd. The first few terms are

$$\mathbf{r} = (1, 1, 1, -1, 1, 1, -1, 1, 1, 1, 1, -1, -1, -1, 1, -1, \ldots)$$

The automaton in Figure 3.6 computes $\mathbf{r}$. To make this easier to see, we have labelled the states to indicate the number of occurrences of 11 read so far modulo 2 and the last bit read by the automaton.

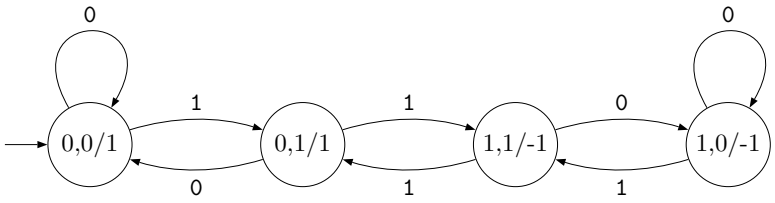

Figure 3.6: A DFAO that computes the Rudin-Shapiro sequence.

3.6 Uniform Morphisms

We can give a useful characterization, due to Cobham [Cob72], of automatic sequences using *morphisms*, maps on languages that behave nicely with respect to concatenation of words.

Definition 3.6.1. Let Σ and Δ be alphabets. A *homomorphism*, or just *morphism*, is a map $\phi : \Sigma^* \to \Delta^*$ that satisfies $\phi(wx) = \phi(w)\phi(x)$ for all $w, x \in \Sigma^*$.

For any morphism ϕ, we must have $\phi(\epsilon) = \epsilon$. Also, from the definition, we see that it is sufficient to define the action of ϕ on Σ, since ϕ can then be uniquely extended to a morphism on Σ^*.

If $|\phi(a)| = k$ for all $a \in \Sigma$, we say that ϕ is a *k-uniform morphism*. A 1-uniform morphism is called *coding*.

Example 3.6.2. Let $\Sigma = \{\mathtt{a}, \mathtt{c}, \mathtt{t}\}$, $\Delta = \{\mathtt{e}, \mathtt{m}, \mathtt{o}, \mathtt{s}, \mathtt{u}\}$ and define

$$\phi(\mathtt{a}) = \mathtt{u},$$
$$\phi(\mathtt{c}) = \mathtt{mo},$$
$$\phi(\mathtt{t}) = \mathtt{se}.$$

Then $\phi(\mathtt{cat}) = \mathtt{mouse}$. Note that ϕ is not uniform.

The next result shows that an automatic sequence remains automatic if we apply a coding.

Proposition 3.6.3. *If* $\mathbf{a} = (a_n)_{n\geq 0}$ *is a k-automatic sequence over* Δ *and* ρ *is a coding on* Δ*, then* $\rho(\mathbf{a})$ *is k-automatic.*

Proof. Since $\mathbf{a}$ is k-automatic, $\mathbf{a}$ is generated by a k-DFAO $M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$. Let $M' = (Q, \Sigma_k, \delta, q_0, \Delta, \rho \circ \tau)$, then clearly $\rho(\mathbf{a})$ is generated by M'. □

Proposition 3.6.3 along with the next result gives a useful closure property of automatic sequences.

Proposition 3.6.4. *Let* $\mathbf{a} = (a_n)_{n\geq 0}$ *and* $\mathbf{b} = (b_n)_{n\geq 0}$ *be k-automatic sequences over* Δ *and* Δ'*, respectively. Then* $\mathbf{a} \times \mathbf{b} = ((a_n, b_n))_{n\geq 0}$ *is a k-automatic sequence over* $\Delta \times \Delta'$.

Proof. Let $M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ and $M' = (Q', \Sigma_k, \delta', q_0', \Delta', \tau')$ be the k-DFAOs generating $\mathbf{a}$ and $\mathbf{b}$, respectively. Define $M'' = (Q \times Q', \Sigma_k, \delta'', (q_0, q_0'), \Delta \times \Delta', \tau'')$, where

$$\delta''((q, q'), c) = (\delta(q, c), \delta'(q', c)) \quad \text{and}$$
$$\tau''((q, q')) = (\tau(q), \tau'(q'))$$

for all $q \in Q, q' \in Q', c \in \Sigma_k$. It is easy to check that M'' generates $\mathbf{a} \times \mathbf{b}$. □

Corollary 3.6.5. *Let* $\mathbf{a} = (a_n)_{n\geq 0}$ *and* $\mathbf{b} = (b_n)_{n\geq 0}$ *be k-automatic sequences over* Δ *and* Δ'*, respectively, and let* $f : \Delta \times \Delta' \to \Delta''$*, where* Δ'' *is a finite alphabet. Then the sequence* $f(\mathbf{a}, \mathbf{b}) = (f(a_n, b_n))_{n\geq 0}$ *is k-automatic.*

Proof. Apply Proposition 3.6.4, then Proposition 3.6.3 using $\rho((a_n, b_n)) = f(a_n, b_n)$. □

For an application of Corollary 3.6.5, see Example 4.4.2.

We will be concerned with morphisms where $\Sigma = \Delta$. In this case, the morphism $\phi : \Sigma^* \to \Sigma^*$ can be iterated. For $a \in \Sigma$, we write $\phi^0(a) = a$ and $\phi^i(a) = \phi(\phi^{i-1}(a))$ for $i \geq 0$. A word $w \in \Sigma^*$ is called a *fixed-point* of ϕ if $\phi(w) = w$.

Example 3.6.6. The Thue-Morse morphism $\mu : \Sigma_2^* \to \Sigma_2^*$ is defined by $\mu(\mathtt{0}) = \mathtt{01}$ and $\mu(1) = \mathtt{10}$, so μ is 2-uniform. We have

$$\begin{aligned}
\mu^2(\mathtt{0}) &= \mathtt{0110},\\
\mu^3(\mathtt{0}) &= \mathtt{01101001},\\
\mu^4(\mathtt{0}) &= \mathtt{0110100110010110}\\
&\vdots
\end{aligned}$$

In fact, for each k, we have $\mu^k(\mathtt{0}) = \mathtt{0}1\mu(1)\mu^2(1)\cdots\mu^{k-1}(1)$, and $\mu^k(\mathtt{0}) = t_0t_1\cdots t_{2^k}$, where $(t_n)_{n\geq 0} = \mathbf{t}$ is the Thue-Morse sequence from Example 3.5.4. So, as k increases, $\mu^k(\mathtt{0})$ matches more and more of the Thue-Morse infinite string, that is, the Thue-Morse sequence written as an infinite string of bits, which is formally the same thing.

Let ϕ be a k-uniform morphism on some alphabet Σ. If there exists $a \in \Sigma$ such that $\phi(a) = aw$ for some string $w \in \Sigma^*$ with $|w| = k - 1$, we say that ϕ is *prolongable* on a. Then $\phi^i(a) = aw\phi(w)\phi^2(w)\cdots\phi^{i-1}(w)$ for $i \geq 2$. As i approaches infinity, $\phi^i(a)$ is "approaching" an infinite word, so we define

$$\phi^\omega(a) = aw\phi(w)\phi^2(w)\cdots .$$

It's not hard to see that $\phi(\phi^\omega(a)) = \phi^\omega(a)$. In fact, in this case, $\phi^\omega(a)$ is the unique fixed point of ϕ starting with a.

In Example 3.6.6, μ is prolongable in $\mathtt{0}$ and $\phi^\omega(\mathtt{0}) = \mathbf{t}$, so $\mathbf{t}$ is the fixed point of μ starting with $\mathtt{0}$.

We have that **t** is a fixed point of a 2-uniform morphism. In fact, every k-automatic sequence is a fixed point of a k-uniform morphism or the image under a coding (a 1-uniform morphism) of a fixed point of a k-uniform morphism.

Before we prove this, we need the following lemma.

Lemma 3.6.7. *Let* $\mathbf{u} = u_0u_1\cdots$ *be an infinite string such that* $\phi(\mathbf{u}) = \mathbf{u}$ *for some* k*-uniform morphism* ϕ*. Then* $\phi(u_i) = u_{ki}u_{ki+1}\cdots u_{ki+k-l}$ *for all* $i \geq 0$*.*

Proof. Since $\mathbf{u} = \phi(\mathbf{u})$ and ϕ is k-uniform, we have

$$\phi(u_0u_1\cdots u_i) = u_0u_1\cdots u_{ki+k-1}.$$

Now

$$\begin{aligned}\phi(u_0u_1\cdots u_i) &= \phi(u_0u_1\cdots u_{i-1})\phi(u_i)\\ &= u_0u_1\cdots u_{ki-1}\phi(u_i),\end{aligned}$$

so

$$\phi(a_i) = u_0u_1\cdots u_{ki+k-1}.$$ □

Theorem 3.6.8. *[Cob69] Let* $k \geq 2$*. Then a sequence* $\mathbf{u} = (u_n)_{n\geq 0}$ *is* k*-automatic if and only if it is the image under a coding of a fixed point of a* k*-uniform morphism.*

Proof. Suppose **u** is the image of a fixed point of a k-uniform morphism, that is, suppose that $\mathbf{u} = \tau(\mathbf{v})$ for some coding $\tau : \Delta \to \Delta$ and $\phi(\mathbf{v}) = \mathbf{v}$ for some k-uniform morphism $\phi : \Delta^* \to \Delta^*$. Write $\mathbf{v} = v_0v_1\cdots$ where $v_i \in \Delta$ for each i.

Let $M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$, where $\delta(a, b)$ is defined to be the bth letter of $\phi(a)$ for $a \in \Delta$ and $b \in \Sigma_k$. We claim that M generates **u**, and thus **u** is k-automatic.

We first show, by induction on n, that $v_n = \delta(w_0, (n)_k)$ for all $n \geq 0$. This holds trivially when $n = 0$. Assume $v_i = \delta(w_0, (i)_k)$ for all $i < n$. Write $(n)_k = a_1a_2\cdots a_t$, then $n = kn' + a_t$

where $(n')_k = n_1 \cdots n_{t-1}$. Then

$$\begin{aligned}
\delta(v_0, (n)_k) &= \delta(v_0, n_1 \cdots n_t) \\
&= \delta(\delta(v_0, n_1 \cdots n_{t-1}), n_t) \\
&= \delta(\delta(v_0, n'), n_t) \\
&= \delta(v_{n'}, n_t) && \text{(by induction)} \\
&= \text{the } n_t\text{th symbol of } \phi(v_{n'}) \\
&= v_{kn'+n_t} && \text{(by Lemma 3.6.7)} \\
&= v_n.
\end{aligned}$$

Therefore $u_n = \tau(v_m) = \tau(\delta(v_0), (n)_k)$, and so $\mathbf{u}$ is generated by M.

Conversely, suppose that $\mathbf{u}$ is k-automatic. Then $\mathbf{u}$ is generated by a k-DFAO $M = (Q, \Sigma_k, \delta, q_o, \Delta, \tau)$. By Theorem 3.5.2, we can assume that $\delta(q_0, 0) = q_0$. We define a k-morphism on Q by

$$\phi(q) = \delta(q, 0)\delta(q, 1) \cdots \delta(q, k-1).$$

Now $\phi(q_0) = q_0\delta(q, 1) \cdots \delta(q, k-1)$, so ϕ is prolongable on q_0. Let $\mathbf{v} = v_0 v_1 \cdots$ be the fixed point of ϕ starting with q_0. We claim that $\delta(q_0, w) = [v]_k$ for all $w \in \Sigma_k^*$.

We will prove our claim by induction on $|w|$. If $|w| = 0$, then we have $\delta(q_0, \epsilon) = q_0 = v_0$ by definition of $\mathbf{v}$. Now assume that the claim is true for all $x \in \Sigma_k^*$ with $|x| < i$. Suppose $|w| = i$, and write $w = xa$, where $a \in \Sigma_k$. Then

$$\begin{aligned}
\delta(q_0, w) &= \delta(\delta(q_0, x), a) \\
&= \delta(v_{[x]_k}, a) && \text{(by induction)} \\
&= \text{the } a\text{-th symbol of } \phi(v_{[x]_k}) && \text{(by definition of } \phi) \\
&= v_{k[x]_k+a} && \text{(by Lemma 3.6.7)} \\
&= v_{[xa]_k} \\
&= v_{[w]_k},
\end{aligned}$$

which proves the claim. Thus we have $\tau(w_n) = \tau(\delta(q_0, (n)_k)) = u_n$, and so $\mathbf{u}$ is the image under τ of a fixed point of ϕ. □

Using Theorem 3.6.8, we can prove one direction of the following very useful theorem on automatic sequences.

Theorem 3.6.9. *Let $m \geq 1$. A sequence $\mathbf{u} = (u_n)_{n\geq 0}$ is k-automatic if and only if it is k^m-automatic.*

Proof. Suppose $\mathbf{u}$ is k-automatic. By Theorem 3.6.8, $\mathbf{u}$ is the image under a coding τ of a fixed point of some k-uniform morphism ϕ. That is,

$$u_0u_1\cdots = \tau(\phi^\omega(a))$$

for some symbol a. Define $\gamma = \tau^m$, then γ is k^m-uniform. Furthermore, $\gamma^\omega(a) = \phi^\omega(a)$, so we have

$$u_0u_1\cdots = \tau(\gamma^\omega(a)).$$

Thus, $\mathbf{u}$ is k^m-automatic.

The proof of the converse is not difficult, but it requires some theory on regular languages (languages accepted by deterministic finite automata) and so is omitted. See [Eil74]. □

3.7 The Kernel of a Sequence

In this section we introduce the *k-kernel* of a sequence, and show how it gives another characterization of k-automatic sequences, as Theorem 3.7.5 shows. This characterization is perhaps not quite as interesting as Theorem 3.6.8, but it is very helpful in proving Christol's Theorem in Chapter 4.

Definition 3.7.1. Let $\mathbf{u} = (u_n)_{n\geq 0}$ be a sequence and let $k \geq 2$ be an integer. The *k-kernel* of $\mathbf{u}$, denoted by $K_k(\mathbf{u})$, is the set of subsequences of $\mathbf{u}$ of the form $(u_{k^en+r})_{n\geq 0}$ where e is any nonnegative integer and r is any integer at least 0 and less that k^e. That is,

$$K_k(\mathbf{u}) = \{(u_{k^en+r})_{n\geq 0} : e \geq 0 \text{ and } 0 \leq r < k^e\}.$$

We illustrate Definition 3.7.1 with a generic example:

Example 3.7.2. Let $\mathbf{u} = (u_n)_{n\geq 0}$ be a sequence. Then the 2-kernel of $\mathbf{u}$ is

$$\begin{aligned} K_2(\mathbf{u}) = \{&\mathbf{u} = (u_n)_{n\geq 0},\\ &(u_{2n})_{n\geq 0}, (u_{2n+1})_{n\geq 0},\\ &(u_{4n})_{n\geq 0}, (u_{4n+1})_{n\geq 0}, (u_{4n+2})_{n\geq 0}, (u_{4n+3})_{n\geq 0},\\ &(u_{8n})_{n\geq 0}, (u_{8n+1})_{n\geq 0}, (u_{8n+2})_{n\geq 0}, (u_{8n+3})_{n\geq 0}, (u_{8n+4})_{n\geq 0}\ldots\}. \end{aligned}$$

Example 3.7.3. We will determine the 2-kernel of the Thue-Morse sequence $\mathbf{t} = (t_n)_{n\geq 0}$ from Example 3.5.4. Let $\mathbf{t}' = (t_{2^e n+r})_{n\geq 0}$ be in the 2-kernel of $\mathbf{t}$. Now $(2^e n + r)_2 = (n)_2 \mathtt{0}^i (r)_2$ for some $i \geq 0$ since $0 \leq r < 2^e$. Then $t_{2^e n+r} = t_n + t_r \mod 2$, so either $\mathbf{t}' = \mathbf{t}$ or $\mathbf{t}' = \bar{\mathbf{t}} = (\overline{t_n})_{n\geq 0}$, where $\overline{t_n} = \mathtt{0}$ if $t_n = \mathtt{1}$ and vice-versa. Therefore $K_2(\mathbf{t}) = \{\mathbf{t}, \bar{\mathbf{t}}\}$.

Example 3.7.4. Let $\mathbf{r} = (r_n)_{n\geq 0}$ be the Rudin-Shapiro sequence from Example 3.5.5. Recall that r_n is 1 or -1 when the number of times $\mathtt{11}$ occurs in $(n)_k$ is even or odd, respectively, so the first few terms are

$$\mathbf{r} = (1, 1, 1, -1, 1, 1, -1, 1, 1, 1, 1, -1, -1, -1, 1, -1, \ldots).$$

Let $n \geq 0$ and let $w = (n)_2$. Then $(2n)_2 = w\mathtt{0}$ and $(4n + 1)_2 = w\mathtt{01}$, so $r_{2n} = r_n$. Also, since $(2n + 1)_2 = w\mathtt{1}$ and $(8n + 7)_2 = w\mathtt{111}$, there are two more occurrences of $\mathtt{11}$ in $(8n + 7)_2$ than in $(2n + 1)_2$, so $r_{8n+7} = r_{2n+1}$. Using arguments like this we can find the following identities (repeating those found above):

$$\begin{aligned} r_{2n} &= r_n, \\ r_{4n+1} &= r_n, \\ r_{8n+7} &= r_{2n+1}, \\ r_{16n+3} &= r_{8n+3}, \\ r_{16n+11} &= r_{4n+3}. \end{aligned}$$

These cover every case, so we conclude that

$$K_2(\mathbf{r}) = \{\mathbf{r}, (r_{2n+1})_{n\geq 0}, (r_{4n+3})_{n\geq 0}, (r_{8n+3})_{n\geq 0}\}.$$

Theorem 3.7.5. *Let $k \geq 2$ and let $\mathbf{u} = (u_n)_{n\geq 0}$ be a sequence. Then $\mathbf{u}$ is k-automatic if and only if $K_k(\mathbf{u})$ is finite.*

Proof. Suppose $\mathbf{u}$ is k-automatic. It follows from Theorem 3.5.3 that there exists a k-DFAO $M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ such that

$$u_n = \tau(\delta(q_0, (n)_k^R \mathtt{0}^t)) \qquad \text{for all } t \geq 0, n \geq 0.$$

Let $\mathbf{u}' = (u_{k^e n+r})_{n\geq 0} \in K_k(\mathbf{u})$, and let $w \in \Sigma_k^*$ with $|w| = e$ and $[w]_k = r$. Also, let $q = \delta(q_0, w^R)$. If $n > 0$, we have

$$(k^e n + r)_k = (n)_k w,$$

so

$$\begin{aligned}\delta(q_0, (k^e n + r)_k^R) &= \delta(\delta(q_0, w^R), (n)_k^R) \\ &= \delta(q, (n)_k^R).\end{aligned}$$

If $n = 0$, then

$$(k^e n + r)_k = (r)_k,$$

so $w = \mathbb{0}^i$ for some $i \geq 0$. Then

$$\begin{aligned}\delta(q_0, (k^e n + r)_k^R) &= \delta(q_0, (r)_k^R) \\ &= \delta(q_0, (r)_k^R \mathbb{0}^i) \\ &= \delta(q_0, w^R) \\ &= q \\ &= \delta(q, (0)_k^R).\end{aligned}$$

Therefore for any $n \geq 0$, we have

$$\delta(q_0, (k^e n + r)_k^R) = \delta(q, (n)_k^R),$$

so $\mathbf{u}'$ is generated by $M' = (Q, \Sigma_k, \delta, q, \Delta, \tau)$. Since Q is finite, there are only finitely many such automata, so $K_k(\mathbf{a})$ is finite.

Conversely, suppose that $K_k(\mathbf{u})$ is finite. Define an equivalence relation on Σ_k^* by

$$w \equiv x \text{ if and only if } u_{k^{|w|}n+[w]_k} = u_{k^{|x|}n+[x]_k} \text{ for all } n \geq 0.$$

Since $K_k(\mathbf{u})$ is finite, this equivalence relation partitions Σ_k^* into finitely many equivalence classes, so we can construct a k-DFAO with

$$\begin{aligned}Q &= \{[x] : x \in \Sigma_k^*\}, \\ \delta([x], a) &= [ax], \\ \tau([w]) &= u_{[w]_k}, \\ q_0 &= [\epsilon],\end{aligned}$$

where $[x]$ denotes the equivalence class containing [x]. We must of course make sure that M is well defined, that is, we must check that δ and τ are well-defined functions.

Suppose that $[x] = [w]$. Then

$$u_{k^{|w|}n+[w]_k} = u_{k^{|x|}n+[x]_k} \text{ for all } n \geq 0. \tag{3.1}$$

Substituting $n = km + r$ into the index on the left hand side, where $0 \leq a < k$, we get

$$\begin{aligned} k^{|w|}n + [w]_k &= k^{|w|+1}m + k^{|w|}a + [w]_k \\ &= k^{|aw|}m + [a\mathbb{0}^{|w|}]_k + [w]_k \\ &= k^{|aw|}m + [aw]_k. \end{aligned}$$

The same holds for any string in Σ_k^*, so we have

$$u_{k^{|aw|}m+[aw]_k} = u_{k^{|ax|}m+[ax]_k} \qquad \text{for all } m \geq 0.$$

Therefore $\delta([x], a) = [ax] = [aw] = \delta([w], a)$, so δ is well-defined.

For τ, we again assume that $[x] = [w]$, that is, that (3.1) holds. Setting $n = 0$, we have

$$\tau([x]) = u_{[x]_k} = u_{[w]_k} = \tau([w]),$$

so τ is well defined.

It can now be shown by a straightforward induction on $|w|$ that $\delta(q_0, w^R) = [w]$, and thus $\tau(\delta(q_0, w^R)) = u_{[w]_k}$, so M generates $\mathbf{u}$. □

Note that Examples 3.7.3 and 3.7.4 agree with 3.7.5 since both the Thuue-Morse sequence and the Rudin-Shapiro sequence are 2-automatic and have finite 2-kernels.

3.8 Fibres and Syndetic Sets

Let $\mathbf{u} = (u_n)_{n\geq 0}$ be a sequence over Δ and let $a \in \Delta$. The *fibre* of f at a is the set

$$\mathbf{u}^{-1}(a) = \{n : u_n = a\}.$$

We note that if $\mathbf{u}$ is generated by a k-DFAO $M = (Q, \Sigma_k, \delta, q_0, \tau, \Delta)$, then for each $a \in \Delta$, $\mathbf{u}^{-1}(a)$ is accepted by the DFA $M' = (Q, \Sigma_k, \delta, q_0, F)$ where $F = \{q : \tau(q) = a\}$

Example 3.8.1. Let $\mathbf{t}$ be the Thue-Morse sequence from Example 3.5.4, then

$$\mathbf{t} = (\mathtt{0}, 1, 1, \mathtt{0}, 1, \mathtt{0}, \mathtt{0}, 1, 1, \mathtt{0}, \mathtt{0}, 1, \mathtt{0}, 1, 1, \mathtt{0}, 1, \mathtt{0}, \mathtt{0}, 1, \mathtt{0}, \ldots),$$

so

$$\mathbf{t}^{-1}(\mathtt{0}) = \{0, 3, 5, 6, 9, 10, 12, 15, 17, 18, 20, \ldots\}$$

and

$$\mathbf{t}^{-1}(1) = \{1, 2, 4, 7, 8, 11, 13, 14, 16, 19, \ldots\}.$$

In this section we give a property of automatic sequences involving the gaps between elements of their fibres. A set of nonnegative integers S is called *syndetic* if the gaps between elements is bounded, or, more formally, if there exists d such that for all sufficiently large n, there is an element of S in the interval $[n, n + d]$.

Proposition 3.8.2. *Let $\mathbf{u} = (u_n)_{n\geq 0}$ be a k-automatic sequence over Δ and let $a \in \Delta$. Then either $\mathbf{u}^{-1}(a)$ is syndetic, or there exists $c > 1$ such that for infinitely many $m \geq 1$ there are no elements of $\mathbf{u}^{-1}(a)$ in the interval $[m, cm]$, that is, $\mathbf{u}^{-1}(a) \cap [m, cm] = \emptyset$ for infinitely many m.*

Proof. See Eilenberg [Eil74, Theorem 5.4]. □

Example 3.8.3. Let $\mathbf{s} = (s)_{n\geq 0}$ be the characteristic sequence of squares, that is,

$$s_n = \begin{cases} 1, & \text{if } n \text{ is a square;} \\ \mathtt{0}, & \text{otherwise.} \end{cases}$$

We can use Proposition 3.8.2 to show that $\mathbf{s}$ is not automatic.

The fibre $\mathbf{s}^{-1}(1)$ is the set of all squares (hence the name "characteristic sequence of squares"), $\{1, 4, 9, 16, \ldots\}$, which is clearly not syndetic since $(n + 1)^2 - n^2 = 2n + 1$ is not bounded. Therefore, if $\mathbf{s}$ is automatic, then there must be some $c \geq 1$ such that $[m, cm] \cap \mathbf{s}^{-1}(a) = \emptyset$ for infinitely many m.

However, we see that if m is sufficiently large, $\sqrt{c}\sqrt{m} - \sqrt{m} = \sqrt{m}(\sqrt{c} - 1) > 1$, since $\sqrt{c} > 1$. Then there exists some integer d such that $\sqrt{m} \leq d \leq \sqrt{c}\sqrt{m}$, and so $d^2 \in [m, cm]$. Thus $[m, cm] \cap \mathbf{s}^{-1}(a) \neq \emptyset$ for all but finitely many m, so it follows from Proposition 3.8.2 that $\mathbf{s}$ is not k-automatic for any $k \geq 2$.

3.9 Cobham's Theorem

We are interested in automatic sequences primarily because of a result from Cobham that we can use to prove that certain sequences are eventually periodic. We will see in a later chapter that when the sequence of coefficients of a formal power series is eventually periodic, then the power series must be rational.

When we say that a sequence $\mathbf{a} = (a_n)_{n\geq 0}$ is *eventually periodic*, we mean that there exist positive integers N and P such that $a_{n+P} = a_n$ for all $n \geq N$. If $N = 0$, we say that $(a_n)_{n\geq 0}$ is *totally periodic* or just *periodic*.

Two positive integers k, l are said to be *multiplicatively independent* if there exist no positive integers r, s such that $k^r = l^s$. Equivalently, k, l are multiplicatively independent if $\log k$ and $\log l$ are linearly independent over $\mathbb{Q}$. For example, 2 and 3 (or indeed any two distinct primes) are multiplicatively independent.

Theorem 3.9.1 (Cobham's Theorem)**.** *Let $k, l \geq 2$ be multiplicatively independent integers, and let $\mathbf{a} = (a_n)_{n\geq 0}$ be a sequence that is both k- and l-automatic. Then $\mathbf{a}$ is eventually periodic.*

The proof given by Cobham in [Cob69] is very difficult (although a footnote states that the proof has been substantially simplified since it was first published in an IBM research note). A simpler proof was given in [Han82] and then again in [Per90] and [AS03]. The last two contained an error which was corrected in [RW06].

Chapter 4

Christol's Theorem

Christol's Theorem is a major result linking automatic sequences to algebraic power series over finite fields. It will be signifigant in the proof of the main result in the case where the base field is finite.

Theorem 4.0.1 (Christol's Theorem)**.** *Let* $\mathbf{a} = (a_n)_{n\geq 0}$ *be a sequence over an alphabet* Δ*, and let* p *be a prime number. Then* $\mathbf{a}$ *is* p*-automatic if and only if there exists an injection* $h : \Delta \hookrightarrow \mathbb{F}_{p^s}$ *for some positive integer* s *such that* $F(X) = \sum_{n\geq 0} h(a_n)X^n$ *is algebraic over* $\mathbb{F}_{p^s}(X)$*.*

This result was originally proved by Christol in [Chr79] in the case where $\Delta = \Sigma_2 = \{\mathtt{0}, \mathtt{1}\}$, and then generalized by Christol, Kamae, Mendès France and Rauzy in [CKMFR80]. The proof presented in this chapter (including the preliminary results) is adapted from [AS03].

4.1 Some Examples

First we give a few examples illustrating Christol's Theorem.

Example 4.1.1. Let $F(X) = X + X^2 + X^4 + \cdots = \sum_{i\geq 0} X^{2^i}$. We saw in Example 1.3.3 that $F(X)$ is algebraic over $\mathbb{F}_2(X)$.

We also saw in that example that $F(X) = \sum_{n\geq 0} a_n X^n$ where $\mathbf{a} = (a_n)_{n\geq 0}$ is the *characteristic sequence of powers of two*, that is, a_n is $\mathtt{1}$ if n is a power of 2 and $\mathtt{0}$ otherwise. Then,

by Christol's Theorem, **a** must be 2-automatic. To see this, we note that if n is a power of 2, the $(n)_2 = 10^k$ for some $k \geq 0$. Therefore the DFAO in figure 4.1 generates **a**, so **a** is 2-automatic.

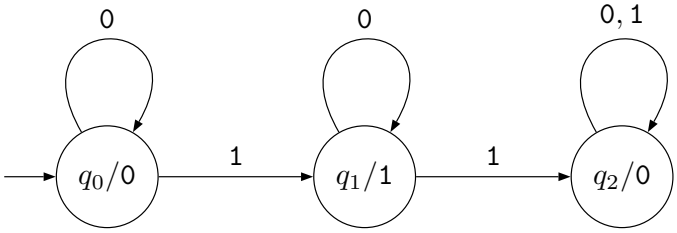

Figure 4.1: A 2-DFAO generating the characteristic sequence of powers of two.

Example 4.1.2. We go back to the Thue-Morse sequence $\mathbf{t} = (t)_{n\geq 0}$ from Example 3.5.4. We saw that **t** was 2-automatic, so by Christol's Theorem, the power series $T(X) = \sum_{n\geq 0} t_n X^n$ must be algebraic over $\mathbb{F}_2(X)$.

To see this, we first note that $(2n)_2 = (n)_2 0$ and $(2n+1)_2 = (n)_2 1$, and so we have $t_{2n} = t_n$ and $t_{2n+1} = (t_n + 1) \mod 2$. Then

$$\begin{aligned} T(X) &= \sum_{n\geq 0} t_{2n} X^{2n} + \sum_{n\geq 0} t_{2n+1} X^{2n+1} \\ &= \sum_{n\geq 0} t_n X^{2n} + X \sum_{n\geq 0} (t_n + 1) X^{2n} \\ &= \sum_{n\geq 0} t_n X^{2n} + X \sum_{n\geq 0} (t_n) X^{2n} + X \sum_{n\geq 0} X^{2n} \\ &= T(X^2) + XT(X^2) + \frac{X}{1 - X^2}. \end{aligned}$$

Therefore, over $\mathbb{F}_2$, we have

$$(1 + X)^3 T(X)^2 + (1 + X)^2 T(X) + X = 0,$$

so $T(X)$ is algebraic over $\mathbb{F}_2(X)$.

Example 4.1.3. Let $\mathbf{s} = (s)_{n\geq 0}$ denote the Rudin-Shapiro sequence on $\{0, 1\}$, so s_n is the number of possibly overlapping occurrences modulo 2 of 11 in $(n)_2$. If **r** is the Rudin-Shapiro sequence over $\{1, -1\}$ from Example 3.5.5 and $\mu : \{0, 1\} \to \{1, -1\}$ is a morphism defined by $\mu(0) = 1, \mu(1) = -1$, then $\mathbf{s} = \mu(\mathbf{r})$.

Since $\mathbf{s}$ is 2-automatic, then $S(X) = \sum_{n\geq 0} s_n X^n$ must be algebraic over $\mathbb{F}_2(X)$. We recall from Example 3.5.5 that $s_{2n} = s_{4n+1} = s_n$, and we also can show that $s_{4n+3} = s_{2n+1} + 1 \mod 2$. Let $T(X) = \sum_{n\geq 0} s_{2n-1} X^n$, then, in $\mathbb{F}_2[[X]]$, we have

$$\begin{aligned} S(X) &= \sum_{n\geq 0} s_{2n} X^{2n} + \sum_{n\geq 0} s_{2n+1} X^{2n+1} \\ &= \sum_{n\geq 0} s_n X^{2n} + X \sum_{n\geq 0} s_{2n+1} X^{2n} \\ &= S(X^2) + XT(X^2) \\ &= S(X)^2 + XT(X)^2. \end{aligned} \tag{4.1}$$

and

$$\begin{aligned} T(X) &= \sum_{n\geq 0} s_{4n+1} X^{2n} + \sum_{n\geq 0} s_{4n+3} X^{2n+1} \\ &= \sum_{n\geq 0} s_n X^{2n} + \sum_{n\geq 0} (s_{2n+1} + 1) X^{2n+1} \\ &= \sum_{n\geq 0} s_n X^{2n} + X\left(\sum_{n\geq 0} s_{2n+1} X^{2n} + \sum_{n\geq 0} X^{2n} \right) \\ &= S(X^2) + XT(X^2) + \frac{X}{1+X^2} \\ &= S(X)^2 + XT(X)^2 + \frac{X}{(1+X)^2}. \end{aligned} \tag{4.2}$$

If we add (4.1) and (4.2), we have

$$\begin{aligned} S(T) + T(X) &= 2S(X^2) + 2XT(X^2) + \frac{X}{(1+X)^2} \\ S(T) + T(X) &= \frac{X}{(1+X)^2} \\ S(T)^2 + T(X)^2 &= \frac{X^2}{(1+X)^4}. \end{aligned}$$

If we solve for $T(X)^2$ and substitute back into (4.2), we have

$$S(X) = S(X)^2 + X\left(S(X)^2 + \frac{X^2}{(1+X)^4} \right),$$

and so

$$(1+X)^5 S(X)^2 + (1+X)^4 S(X) + X^3 = 0$$

in $\mathbb{F}_2[[X]]$. Therefore $S(X)$ is algebraic over $\mathbb{F}_2(X)$.

4.2 Preliminaries

In order to prove Christol's Theorem, we need a few preliminary results about power series in $\mathbb{F}_q[[X]]$ and q-kernels. We define a class of linear transformations to help with the upcoming results. These are called *Cartier operators* since they were introduced by Cartier in [Car58].

Definition 4.2.1. Let $q > 1$ be a prime power. For $0 \leq r < q$, define

$$\Lambda_r : \mathbb{F}_q[[X]] \to \mathbb{F}_q[[X]]$$

by

$$\Lambda_r\left(\sum_{n\geq 0} a_n X^n\right) = \sum_{n\geq 0} a_{qn+r} X^n.$$

This map Λ_r is useful because of the following properties:

Lemma 4.2.2. *Let $F(X)$ and $G(X)$ be formal power series in $\mathbb{F}_q[[X]]$ and let $0 \leq r < q$. Then the following are true:*

(1) Λ_r is a linear transformation;

(2) if $P(X) \in \mathbb{F}_q[X]$, then

$$\deg(\Lambda_r(P(X))) \leq \frac{\deg(P(X))}{q};$$

(3) for $F(X) = \sum_{n\geq 0} a_n X^n \in \mathbb{F}_q[[X]]$,

$$F(X) = \sum_{0\leq r<q} X^r \Lambda_r(F(X))^q; \tag{4.3}$$

(4) for $G(X), H(X) \in \mathbb{F}_q[[X]]$,

$$\Lambda_r(G(X)^q H(X)) = G(X)\Lambda_r(H(X)). \tag{4.4}$$

Proof. See [AS03, Lemma 12.2.2]. □

The previous lemma will be useful right away to prove a stronger condition on algebraic power series over $\mathbb{F}_q$ than given by the definition.

Lemma 4.2.3. *Let $F(X) = \sum_{n\geq 0} a_n X^n \in \mathbb{F}_q[[X]]$, where q is a power of some prime p. Then $F(X)$ is algebraic over $\mathbb{F}_q(X)$ if and only if there exist polynomials $P_0(X), \ldots, P_t(X)$, with $P_0(X) \neq 0$, such that*

$$P_0(X)F(X) + P_1(X)F(X)^q + \cdots + P_t(X)F(X)^{q^t} = 0.$$

Proof. Suppose that $F(X)$ is algebraic, then the set $\{F(X), F(X)^q, F(X)^{q^2}, \ldots\}$ must be linearly dependent over $\mathbb{F}_q(X)$ (since the extension $\mathbb{F}_q(X)(F(X))$ of $\mathbb{F}_q(X)$ has finite degree). Therefore, clearing denominators if necessary, we can write

$$P_0(X)F(X) + P_1(X)F(X)^q + \cdots + P_t(X)F(X)^{q^t} = 0 \tag{4.5}$$

with $P_0(X), \ldots, P_t(X)$ in $\mathbb{F}_q[X]$, with t minimal. Take the smallest integer $j \geq 0$ such that $P_j(X) \neq 0$. We claim that $j = 0$. By (4.3),

$$P_j(X) = \sum_{0\leq r<q} X^r \Lambda_r(P_j(X))^q,$$

so we must have $\Lambda_r(P_j(X)) \neq 0$ for some r. Since Λ_r is linear, from (4.5) we get

$$\Lambda_r(P_j(X)F(X)^{q^j}) + \cdots + \Lambda_r(P_t(X)F(X)^{q^t}) = 0.$$

If $j \neq 0$, we can apply (4.4) to see that

$$\Lambda_r(P_j(X))F(X)^{q^{j-1}} + \cdots + \Lambda_r(P_t(X))F(X)^{q^{t-1}} = 0,$$

which contradicts the minimality of t. Therefore $j = 0$, and so $P_0 \neq 0$.

Conversely, if we have such a relation for $F(X)$, then $F(X)$ is algebraic by definition. □

We will prove Christol's Theorem using Theorem 3.7.5, so it will be useful to prove a special property of the k-kernel involving power series.

Lemma 4.2.4. *Let $\mathbf{a} = (a_n)_{n\geq 0}$ be a sequence over $\mathbb{F}_q$. Then $\mathbf{a}$ is q-automatic if and only if there exists a finite collection $\mathcal{F}$ of power series in $\mathbb{F}_q[[X]]$ such that:*

(1) we have $\sum_{n\geq 0} a_n X^n \in \mathcal{F}$;

(2) for all $A(X) \in \mathcal{F}$ and $0 \leq r < q$, we have $\Lambda_r(A(X)) \in \mathcal{F}$.

Proof. Suppose $\mathbf{a}$ is q-automatic. Then we can write $K_q(\mathbf{a}) = \{\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(d)}\}$, where $\mathbf{a}^{(1)} = \mathbf{a}$ and $\mathbf{a}^{(i)} = (a_n^{(i)})_{n\geq 0}$. Let

$$\mathcal{F} = \left\{ \sum_{n\geq 0} a_n^{(i)} X^n : 1 \leq i \leq d \right\}.$$

Since $\mathbf{a} = \mathbf{a}^{(1)}$, we have $\sum_{n\geq 0} a_n X^n \in \mathcal{F}$. Let $F(X) \in \mathcal{F}$, then $F(X) = \sum_{n\geq 0} a_n^{(i)} X^n$ for some $0 \leq i \leq d$, so

$$\Lambda_r(F(X)) = \sum_{n\geq 0} a_{qn+r}^{(i)} X^n.$$

Now $a_n^{(i)} = a_{q^e n+s}$ for some $e \geq 0$ and $0 \leq s < q^e$, so $a_{qn+r}^{(i)} = a_{q^e(qn+r)+s} = a_{g^{e+1}n+q^e r+s}$. Now

$$\begin{aligned} q^e r + s &< q^e r + q^e \\ &= q^e(r+1) \\ &\leq q^{e+1}, \end{aligned}$$

so $(a_{qn+r}^{(i)})_{n\geq 0}$ is in $K_q(\mathbf{a})$, and therefore $\Lambda_r(F(X)) \in \mathcal{F}$.

Conversely, suppose that $\mathcal{F}$ exists and satisfies the conditions of the lemma and let $(a_{q^e n+r})_{n\geq 0} \in K_q(\mathbf{a})$. If we apply Λ_0 to $\sum_{n\geq 0} a_n X^n$ a total of $e-1$ times, then apply Λ_r, we get $\sum_{n\geq 0} a_{q^e n+r} X^n$, and so $\sum_{n\geq 0} a_{q^e n+r} X^n \in \mathcal{F}$. Therefore $|K_q(\mathbf{a})| \leq |\mathcal{F}|$, so $K_q(\mathbf{a})$ is finite, and thus $\mathbf{a}$ is q-automatic. □

4.3 Proof of Christol's Theorem

We are now ready to prove Christol's Theorem.

Proof of Christol's Theorem. Choose an integer s large enough so that $|\Delta| \leq p^s$, and let $q = p^s$. Then there exists an injection $h : \Delta \hookrightarrow \mathbb{F}_q$, and we may assume that $\mathbf{a}$ is a sequence over $\mathbb{F}_q$.

Suppose that $\mathbf{a}$ is p-automatic. Then, by Theorem 3.6.9, $\mathbf{a}$ is q-automatic. Therefore $K_q(\mathbf{a})$ is finite by Theorem 3.7.5, so we can write $K_q(\mathbf{a}) = \{\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(d)}\}$ as in the proof of Lemma 4.2.4. For each $1 \leq j \leq d$, let $F_j(X) \coloneqq \sum_{n\geq 0} a_n^{(j)} X^n$. We can write

$$F_j(X) = \sum_{r=0}^{q-1} \sum_{m\geq 0} a_{qm+r}^{(j)} X^{qm+r} = \sum_{r=0}^{q-1} X^r \sum_{m\geq 0} a_{qm+r}^{(j)} X^{qm}.$$

As in the proof of Lemma 4.2.4, $(a^{(j)}_{qm+r})_{m\geq 0} \in K_q(\mathbf{a})$, so it is equal to $\mathbf{a}^{(i)}$ for some $1 \leq i \leq d$, and thus for each $0 \leq r < q$, $\sum_{m\geq 0} a^{(j)}_{qm+r} X^{qm} = F_i(X^q)$ for some $0 \leq i \leq d$. Therefore $F_j(X)$ is an $\mathbb{F}_q[X]$-linear combination of the power series $F_i(X^q)$, that is,

$$F_j(X) \in \mathrm{span}_{\mathbb{F}_q(X)}\{F_1(X^q), \ldots, F_d(X^q)\},$$

By substituting X^q for X, we get

$$F_j(X^q) \in \mathrm{span}_{\mathbb{F}_q(X)}\{F_1(X^{q^2}), \ldots, F_d(X^{q^2})\},$$

and so

$$\mathrm{span}_{\mathbb{F}_q(X)}\{F_1(X^q), \ldots, F_d(X^q)\} \subset \mathrm{span}_{\mathbb{F}_q(X)}\{F_1(X^{q^2}), \ldots, F_d(X^{q^2})\}.$$

Thus for any $1 \leq j \leq d$ we have

$$F_j(X), F_j(X^q) \in \mathrm{span}_{\mathbb{F}_q(X)}\{F_1(X^{q^2}), \ldots, F_d(X^{q^2})\}.$$

Again, if we substitute X^q for X, we get

$$F_j(X^q), F_j(X^{q^2}) \in \mathrm{span}_{\mathbb{F}_q(X)}\{F_1(X^{q^3}), \ldots, F_d(X^{q^3})\},$$

and so, as above,

$$F_j(X), F_j(X^q), F_j(X^{q^2}) \in \mathrm{span}_{\mathbb{F}_q(X)}\{F_1(X^{q^3}), \ldots, F_d(X^{q^3})\}.$$

Continuing in this way, we arrive at

$$F_j(X^{q^k}) \in \mathrm{span}_{\mathbb{F}_q(X)}\{F_1(X^{q^{d+1}}), \ldots, F_d(X^{q^{d+1}})\}$$

for all $0 \leq k \leq d$.

Now the dimension of the $\mathbb{F}_q(X)$-vector space spanned by $F_1(X^{q^{d+1}}), \ldots, F_d(X^{q^{d+1}})$ is at most d, so then for each j, $\{F_j(X), F_j(X^q), \ldots, F_j(X^{q^d})\}$ is linearly dependent over $\mathbb{F}_q(X)$. In particular, since $F = F_1$, there are $R_0(X), \ldots, R_d(X) \in \mathbb{F}_q(X)$, not all zero, such that

$$0 = R_0(X)F(X) + R_1(X)F(X^q) + \cdots + R_d(X)F(X^{q^d}).$$

Clearing denominators, we have

$$\begin{aligned} 0 &= P_0(X)F(X) + P_1(X)F(X^q) + \cdots + P_d(X)F(X^{q^d}) \\ &= P_0(X)F(X) + P_1(X)F(X)^q + \cdots + P_d(X)F(X)^{q^d}, \end{aligned}$$

Therefore $F(X)$ is algebraic over $\mathbb{F}_q(X)$.

Conversely, suppose that $F(X)$ is algebraic over $\mathbb{F}_q(X)$. By Lemma 4.2.3, there exist polynomials $P_0(X), \ldots, P_t(X)$ with $P_0 \neq 0$ such that

$$P_0(X)F(X) + P_1(X)F(X)^q + \cdots + P_t(X)F(X)^{q^t} = 0.$$

We divide by $P_0(X)^2$ and set $G = F(X)/P_0(X)$ to get

$$G(X) = \sum_{i=1}^{t} Q_i(X)G(X)^{q^i}, \quad \text{where } Q_i(X) = -P_i(X)P_0(X)^{q^i-2}. \tag{4.6}$$

We wish to use Lemma 4.2.4, so we let

$$\mathcal{H} = \left\{ H \in \mathbb{F}_q[[X]] : H = \sum_{i=0}^{t} R_i(X)G(X)^{q^i} \text{ with } R_i(X) \in \mathbb{F}_q[X] \text{ and } \deg(R_i(X)) \leq N \forall i \right\},$$

where

$$N = \max(\deg(P_0(X)), \deg(Q_1(X)), \ldots, \deg(Q_t(X))).$$

We can see that $\mathcal{H}$ is finite and that $F(X) = P_0(X)G(X) \in \mathcal{H}$, but we must still show that $\mathcal{H}$ is closed under Λ_r. Let $H(X) = \sum_{i=0}^{t} R_i(X)G(X)^{q^i} \in \mathcal{H}$. Then,

$$\begin{aligned} H(X) &= R_0(X)G(X) + \sum_{i=1}^{t} R_i(X)G(X)^{q^i} \\ &= \sum_{i=1}^{t} R_0(X)Q_i(X)G(X)^{q^i} + \sum_{i=1}^{t} R_i(X)G(X)^{q^i} \qquad \text{by (4.6)} \\ &= \sum_{i=1}^{t} (R_0(X)Q_i(X) + R_i(X))G^{q^i} \end{aligned}$$

By Lemma 4.2.2, we have

$$\Lambda_r(H(X)) = \sum_{i=1}^{t} \Lambda_r(R_0(X)Q_i(X) + R_i(X))G^{q^{i-1}}.$$

We have chosen N and defined $\mathcal{H}$ such that $\deg\left(R_0(X)Q_i(X) + R_i(X)\right) \leq 2N$, so we have

$$\deg\left(\Lambda_r(R_0(X)Q_i(X) + R_i(X))\right) \leq \frac{2N}{q} \leq N$$

by Lemma 4.2.2(b), and so $\Lambda_r(H(X)) \in \mathcal{H}$. Thus, by Lemma 4.2.4, $K_q(\mathbf{a})$ is finite, and thus $\mathbf{a}$ is q-automatic, and therefore p-automatic by Theorem 3.6.9. □

4.4 Applications of Christol's Theorem

We can use Christol's Theorem and a closure property of automatic sequences to quickly prove that the algebraic power series in $\mathbb{F}_q[[X]]$ are closed under the Hadamard product.

Definition 4.4.1. Let $F(X) = \sum_{n\geq 0} a_n X^n$ and $G(X) = \sum_{n\geq 0} b_n X^n$ be two formal power series over a field K. The *Hadamard product* of $F(X)$ and $G(X)$ is defined by

$$F(X) \odot G(X) = \sum_{n\geq 0} a_n b_n X^n.$$

Proposition 4.4.2. *Let $F(X), G(X) \in \mathbb{F}_q[[q]]$, where q is a power of some prime p. If $F(X)$ and $G(X)$ are algebraic, the $F(X) \odot G(X)$ is also algebraic.*

Proof. Let $F(X) = \sum_{n\geq 0} a_n X^n$ and $G(X) = \sum_{n\geq 0} b_n X^n$. By Christol's Theorem, $\mathbf{a} = (a_n)_{n\geq 0}$ and $\mathbf{b} = (b_n)_{n\geq 0}$ are both p-automatic. Define $f : \mathbb{F}_q \times \mathbb{F}_q \to \mathbb{F}_q$ by $f(c, d) = cd$, then $F(X) \odot G(X) = \sum_{n\geq 0} f(a_n, b_n)X^n$. By Corollary 3.6.5, $(f(a_n, b_n))_{n\geq 0}$ is p-automatic, and thus, using Christol's Theorem again, we see that $F(X) \odot G(X)$ is algebraic. □

Christol's Theorem, unsurprisingly, is useful for proving that certain power series are transcendental. For instance, we can use it show that power series are transcendental over $\mathbb{Q}(X)$. To do that, we need the following result.

Proposition 4.4.3. *Let $F(X) = \sum_{n\geq 0} a_n X^n \in \mathbb{Z}[[X]]$ and let p be prime. Define $F_p(X) = \sum_{n\geq 0}(a_n \bmod p)X$ to be the reduction of $F(X)$ modulo p. If $F(X)$ is algebraic over $\mathbb{Q}(X)$, then $F_p(X)$ is algebraic over $\mathbb{F}_p(X)$.*

Proof. Suppose $F(X)$ is algebraic over $\mathbb{Q}(X)$. Then there exist polynomials $P_0, P_1, \ldots, P_d \in \mathbb{Q}[X]$, not all zero, such that

$$P_0 + P_1 F(X) + \cdots + P_d F(X)^d = 0. \tag{4.7}$$

Clearing denominators if necessary, we can assume that $P_0, \ldots, P_d \in \mathbb{Z}[X]$. We can also assume that all the coefficients of $P_0, \ldots, P_d$ are relatively prime by dividing through by any common divisor. This means at least one coefficient of one of the P_i's is not a multiple of p. Thus, if we consider (4.7) modulo p, we get a nontrivial relation for $F_p(X)$, and we see that $F_p(X)$ is algebraic over $\mathbb{F}_p$. □

The following examples shows how we can use Christol's Theorem to prove that a power series is transcendental over $\mathbb{Q}(X)$.

Example 4.4.4. Let $F(X) = X + X^2 + X^4 + \cdots = \sum_{i\geq 0} X^{2^i}$ and let $\mathbf{a} = (a_n)_{n\geq 0}$ be the sequence of coeffients of $F(X)$. Suppose that $F(X)$ is algebraic over $\mathbb{Q}(X)$. It follows from Christol's Theorem and Proposition 4.4.3 that $\mathbf{a}$ is both 2- and 3-automatic (and indeed p-automatic for any prime p). Since 2 and 3 are multiplicatively independent, $\mathbf{a}$ is eventually periodic by Cobham's Theorem. This is a contradiction since $\mathbf{a}$ is the characteristic sequence of powers of 2, which is not eventually periodic since the gaps between 1s in the sequence increase without bound as $n \to \infty$.

Example 4.4.5. The *classical theta series* is defined as

$$\theta_3(X) = \sum_{k\in\mathbb{Z}} X^{k^2}.$$

We will show that $\theta_3(X)$ is transcendental over $\mathbb{Q}(X)$.

First we note that $\theta_3(X) = 1 + 2\sum_{k\geq 1} X^{k^2}$. Let $G(X) = \sum_{k\geq 1} X^{k^2}$. If $\theta_3(X)$ is algebraic over $\mathbb{Q}(X)$, then so is $G(X)$, and thus the power series $F_p(X)$ is algebraic over $\mathbb{F}_p(X)$ by Proposition 4.4.3. Therefore, by Christol's Theorem, the sequence of coefficients of $F_p(X)$ is p-automatic. However, the sequence of coefficients of $F_p(X)$ is the characteristic sequence of squares, $\mathbf{s}$, seen in Example 3.8.3, and we saw that $\mathbf{s}$ was not k-automatic for any $k \geq 2$. Therefore we conclude that $\theta_3(X)$ is transcendental over $\mathbb{Q}(X)$.

Part II

The Main Result

Chapter 5

The Finite Field Case

Before we prove the main result in general, we will prove the special case where the base field K is finite.

Theorem 5.0.1. *Let q be a power of some prime p, let $\phi(X)$ be a rational function in $\mathbb{F}_q(X)$ whose power series expansion lies in $X\mathbb{F}_q[[X]]$, and let $d \geq 2$. Let $F(X) = \sum_{n=0}^{\infty} \phi(X^{d^n}) \in \mathbb{F}_q[[X]]$. If d is not a power of p, then either $F(X) \in \mathbb{F}_q(X)$ or $F(X)$ is transcendental over $\mathbb{F}_q(X)$.*

To prove Theorem 5.0.1, we will use results presented in Chapters 3 and 4.

5.1 Periodicity of Coefficients

If the coefficients of a power series are eventually periodic, that is, they begin to repeat after a while, then we can show that the power series is rational. In fact, the converse holds over finite fields. It will then remain to show that the coefficients of the power series $F(X)$ from Theorem 5.0.1 are eventually periodic.

Definition 5.1.1. Let $(a_n)_{n\geq 0}$ be a sequence. We say that $(a_n)_{n\geq 0}$ is *eventually periodic* if there exist natural numbers N and M such that $a_{n+M} = a_n$ for all $n \geq N$. If $N = 0$, we say that $(a_n)_{n\geq 0}$ is *totally periodic* or just *periodic*. If we write $\mathbf{a} = a_0a_1a_2\cdots$, then the strings $a_0a_1a_{N-1}$ and $a_Na_{N+1}\cdots a_{N+M-1}$ are called the *preperiod* and *period* of $\mathbf{a}$, respectively. Using the notation of Section 3.1, we can write $\mathbf{a} = a_0a_1a_{N-1}(a_Na_{N+1}\cdots a_{N+M-1})^{\omega}$.

Proposition 5.1.2. *Let K be a (possibly infinite) field and let $\phi(X) \in K[[X]]$. If the sequence of coefficients $\mathbf{a} = (a_n)_{n\geq 0}$ of $\phi(X)$ is eventually periodic, then $\phi(X)$ is rational.*

Proof. Suppose that there exist $N, M \in \mathbb{N}$ such that $a_{n+M} = a_n$ for all $n \geq N$. Let $C(X) = a_0 + a_1X + \cdots + a_{N-1}X^{N-1}$. Then

$$\begin{aligned}\phi(X) &= C(X) + \sum_{n=N}^{\infty} a_n X^n \\ &= C(X) + \sum_{k=0}^{\infty}\sum_{j=0}^{M-1} a_{N+kM+j}X^{N+kM+j} \\ &= C(X) + \sum_{k=0}^{\infty} X^{kM}\sum_{j=0}^{M-1} a_{N+j}X^{N+j} \\ &= C(X) + \frac{R(X)}{1-X^M},\end{aligned}$$

where $R(X) = a_N + a_{N+1}X + \cdots + a_{N+M-1}X^{M-1}$. Therefore

$$\phi(X) = \frac{(1-X^M)C(X) + R(X)}{1-X^M},$$

and so $\phi(X)$ is rational. Note that coefficients of $C(X)$ and $R(X)$ make up the preperiod and period of $\mathbf{a}$, respectively. □

We will now prove the converse in the case that K is finite. The proof relies on the following lemma:

Lemma 5.1.3. *Let $q = p^r$ for some prime p, and let $A(X) \in \mathbb{F}_q[X]$ such that $A(0) = 1$. Then $A(X)\Big|(1 - X^{(p^s-1)p^t})$ for some integers s, t.*

Proof. Suppose that $A(X)$ is irreducible. Then $\mathbb{F}_q[X]/(A(X))$ is a field. This field is finite, so we have $\mathbb{F}_q[X]/(A(X)) \cong \mathbb{F}_{q'}$ for some q'. Now $\mathbb{F}_q[X]/(A(X))$ has characteristic p and we have $\mathbb{F}_q \hookrightarrow \mathbb{F}_q[X]/(A(X))$. Therefore $q' = p^s$ for some $s \geq r$.

For any nonzero $a \in \mathbb{F}_{q'}$, we have $a^{q'-1} = 1$, so then $1 - \overline{X}^{q'-1} = 0$ in $\mathbb{F}_q[X]/(A(X))$, where $\overline{X}$ is the equivalence class of X. Therefore $A(X)\Big|1 - X^{p^s-1}$.

If $A(X)$ is not irreducible, we can write

$$A(X) = A_1(X)\cdots A_k(X)$$

where $A_1(X), \ldots, A_k(X)$ are irreducible and not necessarily distinct. For each $1 \leq i \leq k$, we have $A_i(X) \Big| 1 - X^{p^{s_i}-1}$ for some s_i. Let $s = s_1 \cdots s_k$. Then, since

$$p^{ab} - 1 = (p^a)^b - 1 = (p^a - 1)(p^{a(b-1)} + \cdots + p^a + 1),$$

we have $p^{s_i} - 1 \Big| p^s - 1$ and, using a similar factorization, $1 - X^{p^{s_i}-1} \Big| 1 - X^{p^s-1}$ for each i. Thus $A_i(X) \Big| 1 - X^{p^s-1}$ for each i, and so

$$A_1(X) \cdots A_k(X) \Big| (1 - X^{p^s-1})^k.$$

If we choose t such that $p^t > k$, then we have

$$A(X) \Big| (1 - X^{p^s-1})^{p^t} = 1 - X^{(p^s-1)p^t} \qquad \square$$

We can now prove the converse of Theorem 5.1.2 in the finite field case.

Proposition 5.1.4. *Let $\phi = \sum_{n \geq 0} a_n X^n$ be a power series in $\mathbb{F}_q[[X]]$. If $\phi(X)$ is rational, then its sequence of coefficients $(a_n)_{n \geq 0}$ is eventually periodic.*

Proof. Write $\phi(X) = A(X)/B(X)$ where $A(X), B(X) \in \mathbb{F}_q[X]$, $B(0) = 1$ and $\gcd(A, B) = 1$. By Lemma 5.1.3, $B(X) \Big| (1 - X^{(p^s-1)p^t})$ for some s and t, so we can write

$$\phi(X) = \frac{\widetilde{A}(X)}{1 - X^M} \tag{5.1}$$

with $\widetilde{A}(X) \in \mathbb{F}_q[X]$ and setting $M = (p^s - 1)p^t$. Using the division algorithm, we can write

$$\widetilde{A} = C(X)(1 - X^M) + R(X),$$

where $C(X), R(X) \in \mathbb{F}_q[X]$ and $\deg(R(X)) < M$. We can then rewrite (5.1) as

$$\phi(X) = C(X) + \frac{R(X)}{1 - X^M}.$$

Since $\deg(R(X)) < M$, we can write $R(X) = r_0 + \cdots + r_{M-1}X^{M-1}$, and so

$$\begin{aligned}\frac{R(X)}{1-X^M} &= \sum_{j=0}^{M-1} r_j X^j \sum_{k=0}^{\infty} X^{kM} \\ &= \sum_{j=0}^{M-1}\sum_{k=0}^{\infty} r_j X^{kM+j} \\ &= \sum_{k=0}^{\infty}\sum_{j=0}^{M-1} r_j X^{kM+j} \\ &= \sum_{k=0}^{\infty} X^{kM} \sum_{j=0}^{M-1} r_j X^j. \end{aligned} \tag{5.2}$$

If $C(X) \neq 0$, we let $N = \deg(C(X)) + 1$; if $C(X) = 0$, we let $N = 0$. Given $n \geq N$, we let m denote the least residue of $n - N$ modulo M; then from (5.2) we can see that $a_{n+M} = a_n = r_m$. □

Example 5.1.5. Let $\phi = (1 - X - X^3)/(1 + X + X^2) \in \mathbb{F}_3[[X]]$. We can write

$$\begin{aligned}\phi(X) &= \frac{(1-X)(1-X-X^3)}{1-X^3} \\ &= \frac{1+X+X^2-X^3+X^4}{1-X^3} \\ &= 1 - X + \frac{-X+X^2}{1-X^3} \\ &= 1 - X + (-X+X^2)(1+X^3+X^6+\cdots) \\ &= 1 + 2X + (2X+X^2)(1+X^3+X^6+\cdots) \\ &= 1 + X + X^2 + 2X^4 + X^5 + 2X^7 + X^8 + 2X^9 + \cdots. \end{aligned} \tag{5.3}$$

From (5.3) we see that if $n \geq 2$, then then n-th coefficient of $\phi(X)$ is

$$a_n = \begin{cases} 1, \text{ if } n-2 \equiv 0 \pmod 3; \\ 0, \text{ if } n-2 \equiv 1 \pmod 3; \\ 2, \text{ if } n-2 \equiv 2 \pmod 3. \end{cases}$$

Therefore the coefficients of $\phi(X)$ are eventually periodic with $N = 2$ and $M = 3$ as defined in Definition 5.1.1. Then we can write the infinite string of coefficients of $\phi(X)$ as $11(102)^\omega$.

Proposition 5.1.2 is in fact true over any ring. It can be seen from the proof that any power series with eventually periodic coefficients is rational over any ring, but sometimes rational power series can have non-periodic coefficients. For example, consider

$$\frac{1}{1-2X} = 1 + 2X + 4X^2 + 8X^3 + \cdots .$$

The sequence $(2^n)_{n\geq 0}$ is not periodic over $\mathbb{Z}$, but it *is* eventually periodic over $\mathbb{Z}/N\mathbb{Z}$ for any N. In fact, this is the only kind of problem that can arise, so Proposition 5.1.4 holds over any ring where $(a^n)_{n\geq 0}$ is eventually periodic for all $a \in R$. See [HLPPA09].

5.2 The Coefficients of Our Series Are Automatic

In this section, we will show that the coefficients of $F(X)$ are d-automatic if d is not a power of p. This will rely on the fact that the coefficients of $\phi(X)$ are d-automatic, which follows from Proposition 5.1.4 and the result of Büchi [Büc60] that the coefficients of an eventually periodic sequence are k-automatic for any $k \geq 2$.

Proposition 5.2.1. *Let $\phi(X)$ and $F(X)$ be defined as in Theorem 5.0.1. Write $F(X) = \sum_{n\geq 0} b_n X^n$ and let $\mathbf{b} = (b_n)_{n\geq 0}$. If d is not a power of p, then $\mathbf{b}$ is d-automatic.*

Proof. Write $\phi(X) = \sum_{j\geq 0} a_j X^j$ and let $\mathbf{a} = (a_n)_{n\geq 0}$. Then

$$\begin{aligned} F(X) &= \sum_{k=0}^{\infty} \phi(X^{d^k}) \\ &= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} a_j X^{d^k j}. \end{aligned}$$

We wish to determine the coefficients of $\mathbf{b}$ in terms of the coefficients of $\mathbf{a}$. If $d^k j = m$, then $d^k | m$, and $j = \frac{m}{d^k}$.

$$b_m = \sum_{k=0}^{\infty} a_{\frac{m}{d^k}} \chi(d^k | m),$$

where $\chi(P)$ is 1 if the statement P is true and 0 if P is false.

Now let V be the vector space over $\mathbb{F}_q$ spanned by $K_d(\mathbf{a})$. Since $\mathbf{a}$ is d-automatic, V is finite. We claim that $K_d(\mathbf{b}) \subset V$. To see this, let $\mathbf{b}' \in K_d(\mathbf{b})$, then $\mathbf{b}' = (b_{d^e n+c})_{n\geq 0}$ for some e

and c. If $d \nmid c$, then $d \nmid d^e n + c$, so $b_{d^e n+c} = a_{d^e n+c}$ for all n, so $\mathbf{b}' \in K_d(\mathbf{a})$. If $c = d^r c'$ where $d \nmid c'$, then

$$\begin{aligned} b_{d^e n+c} &= b_{d^e n+d^r c'} \\ &= a_{d^e n+d^r c'} + a_{d^{e-1} n+d^{r-1} c'} + \cdots + a_{d^{e-r} n+c'}. \end{aligned}$$

This is true for all $n \geq 0$, so

$$\mathbf{b}' = \sum_{j=0}^{r} (a_{d^{e-j} n+d^{r-j} c'})_{n\geq 0} \in V,$$

and therefore $K_d(\mathbf{b}) \subset V$. Thus $K_d(\mathbf{b})$ is finite, and so $\mathbf{b}$ is d-automatic. □

5.3 Proof of the Finite Field Case

To tie the results of the previous two sections together, we use Cobham's Theorem (Theorem 3.9.1).

Proof of Theorem 5.0.1. Assume that d is not a power of p, then d and p are multiplicatively independent. By Proposition 5.2.1, we know that the coefficients of $F(X)$ are d-automatic. If $F(X)$ is algebraic over $\mathbb{F}_q(X)$, then, by Christol's Theorem (4.0.1), its coefficients are also p-automatic. Using Theorem 3.9.1, we see that the coefficients are eventually periodic, and therefore $F(X)$ is rational by Proposition 5.1.2. □

As mentioned above, we will use this special case to prove the main theorem in the general case, but we will need a slightly stronger version. We can strengthen the second conclusion by bounding the degree of the numerator and denominator polynomials of $F(X)$.

Theorem 5.3.1. *Let q be a power of some prime p and let $\phi(X)$ be a rational function in $\mathbb{F}_q(X)$ whose power series expansion lies in $X\mathbb{F}_q[[X]]$. Write $\phi(X) = A(X)/B(X)$ with $A(X), B(X) \in \mathbb{F}_q[X]$ having no common divisors. Let $d \in \mathbb{N}$ and let $F(X) = \sum_{n=0}^{\infty} \phi(X^{d^n}) \in \mathbb{F}_q[[X]]$. If d is not a power of p, then either $F(X) \in \mathbb{F}_q(X)$ or $F(X)$ is transcendental over $\mathbb{F}_q(X)$. In the case where $F(X)$ is rational, $F(X) = C(X)/D(X)$ with $C(X), D(X) \in \mathbb{F}_q[X]$ such that* $\deg(C(X)) \leq \max(\deg(A(X)), \deg(B(X)))$ *and* $\deg(D(X)) \leq \deg(B(X))$

Proof. Suppose that d is not a power of p and that $F(X)$ is rational. Then we can write $F(X) = C(X)/D(X)$ with $C(X), D(X) \in \mathbb{F}_q[X]$ having no common factors. Thus there exist polynomials $Q_1(X), Q_2(X) \in \mathbb{F}_q$ such that

$$Q_1(X)C(X) + Q_2(X)D(X) = 1.$$

Substituting X^d, we obtain

$$Q_1(X^d)C(X^d) + Q_2(X^d)D(X^d) = 1,$$

and so $C(X^d)$ and $D(X^d)$ have no common factors.

Now

$$F(X^d) = \sum_{n=0}^{\infty} \phi(X^{d^{n+1}}) = F(X) - \phi(X),$$

which gives

$$\frac{C(X^d)}{D(X^d)} = \frac{C(X)}{D(X)} - \frac{A(X)}{B(X)}.$$

Clearing denominators, we obtain

$$B(X)C(X^d)D(X) = C(X)D(X^d)B(X) - A(X)D(X)D(X^d). \tag{5.4}$$

Suppose that $P(X) \in \mathbb{F}_q[X]$ such that $P(X) \mid D(X^d)$, then $P(X)$ divides the right-hand side of (5.4), and so $P(X) \mid B(X)D(X)$ since $C(X^d)$ and $D(X^d)$ have no common factors. Thus

$$\deg(D(X^d)) \leq \deg(D(X)) + \deg(B(X)),$$

which gives

$$(d-1)\deg(D(X)) \leq \deg(B(X)).$$

Since $d \geq 2$, we see that $\deg(D(X)) \leq \deg(B(X))$.

Now from (5.4) we obtain

$$\deg(B(X)C(X^d)D(X)) = \deg(C(X)D(X^d)B(X) - A(X)D(X)D(X^d)).$$

If we denote by K, L, M and N the degrees of $A(X), B(X), C(X)$ and $D(X)$, respectively, we then obtain

$$\begin{aligned} L + dM + N &\leq \max(M + dN + L, K + N + dN) \\ &= \max(L + M, K + N) + dN, \end{aligned}$$

so

$$dM \leq \max(L + M, K + N) + (d - 1)N - L.$$

Now $N \leq L$, so we have

$$\begin{aligned} dM &\leq \max(L + M, K + L) + (d - 2)L \\ &= \max(M, K) + (d - 1)L. \end{aligned}$$

Now if $M \leq K$, then $\deg(C(X)) \leq \deg(A(X))$. If $M > K$, then we have $dM \leq M + (d - 1)L$, so $M \leq L$ and thus in this case $\deg(C(X)) \leq \deg(B(X))$. □

Recall that Theorem 5.0.1 holds only when d is not a power of p. Example 1.3.3 showed how this can happen over $\mathbb{F}_2$; the same idea can be generalized to any finite field using the following result:

Lemma 5.3.2. *Let q be a prime power and let $G(X) \in \mathbb{F}_q[[X]]$. Then $G(X)^q = G(X^q)$.*

Proof. See [[AS03]]. □

Theorem 5.3.3. *Let q and d be powers of some prime p and let $\phi(X)$ be a rational function in $\mathbb{F}_q(X)$ whose power series expansion lies in $X\mathbb{F}_q[[X]]$. Then $F(X) \coloneqq \sum_{n=0}^{\infty} \phi(X^{d^n})$ is algebraic over $\mathbb{F}_q(X)$.*

Proof. Write $q = p^s$ and $d = p^t$ with $s, t \geq 1$. By Lemma 5.3.2, we have

$$F(X)^{q^t} = F\left(X^{q^t}\right) = \sum_{n=0}^{\infty} \phi\left(X^{q^t d^n}\right).$$

Now $q^t d^n = p^{st} p^{nt} = p^{(n+s)t} = d^{n+s}$, so

$$\begin{aligned} F(X)^{q^t} &= \sum_{n=0}^{\infty} \phi\left(X^{d^{n+s}}\right) \\ &= F(X) - \phi(X) - \phi\left(X^d\right) - \cdots - \phi\left(X^{d^{s-1}}\right). \end{aligned}$$

Setting $\psi(X) = \phi(X) + \phi\left(X^d\right) + \cdots + \phi\left(X^{d^{s-1}}\right) \in \mathbb{F}_q(X)$, we see that $F(X)^{q^t} - F(X) + \psi(X) = 0$, and so $F(X)$ is algebraic over $\mathbb{F}_q$. □

Example 5.3.4. Let $\phi(X) = X \in \mathbb{F}_4(X)$ and let $d = 8$. Then $F(X) = \sum_{n=0}^{\infty} X^{8^n}$. We see that

$$\begin{aligned} F(X)^{4^3} &= F(X^{4^3}) && \text{(by Lemma 5.3.2)} \\ &= \sum_{n=0}^{\infty} \phi\left(X^{4^3 8^n}\right) \\ &= \sum_{n=0}^{\infty} \phi\left(X^{2^{3(n+2)}}\right) \\ &= \sum_{n=0}^{\infty} \phi\left(X^{8^{n+2}}\right). \end{aligned}$$

Therefore

$$F(X)^{64} - F(X) + X + X^8 = 0,$$

and so $F(X)$ is algebraic over $\mathbb{F}_4(X)$. We could show that $F(X)$ is not rational by an argument similar the one used in Example 1.3.2, but it is easier to use Proposition 5.1.4 and note that the sequence of coefficients of $F(X)$ is not eventually periodic since the size of the gaps between 1s increases without bound as $n \to \infty$.

Chapter 6

The Main Result

In this section we will prove the main result in the general case, which we restate:

Main Theorem. *Let K be a field, let $\phi(X)$ be a rational function in $K(X)$ whose power series expansion lies in $XK[[X]]$, and let $d \geq 2$. Let $F(X) = \sum_{n=0}^{\infty} \phi(X^{d^n}) \in K[[X]]$. If d is not a power of* $\mathrm{char}(K)$*, then either $F(X) \in \mathbb{F}_q(X)$ or $F(X)$ is transcendental over $\mathbb{F}_q(X)$.*

6.1 Preliminaries

We will use a lemma in the proof of the main result, which in turn relies on the following result from linear algebra.

Lemma 6.1.1. *Let A be an $n \times \aleph_0$ matrix with coefficients in an integral domain R. Then the rows of A are linearly dependent over R if and only if every $n \times n$ submatrix of A has determinant 0.*

Proof. Suppose that the rows of A are linearly dependent over R. Then there exists a vector

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \in R^n \setminus \{\mathbf{0}\}$$

such that $\mathbf{u}^T A = \begin{pmatrix} 0 & 0 & \cdots \end{pmatrix}$. If we let $\mathbf{a}^{(i)}$ denote the i-th column of A, then this is equivalent to saying that $\mathbf{u}^T \mathbf{a}^{(i)} = 0$ for all $i \geq 0$. Let B be a $n \times n$ submatrix of A. Since B is made up of columns of A, we see that $\mathbf{u}^T B = \mathbf{0}$, and so $\det(B) = 0$.

Conversely, suppose that the rows of A are linearly independent over R. Let K denote the field of fractions of R and suppose there exists a vector $\mathbf{u} \in K^n$ such that $\mathbf{u}^T A = \begin{pmatrix} 0 & 0 & \cdots \end{pmatrix}$. We can clear denominators in the entries of $\mathbf{u}$ to obtain $\mathbf{u}' \in R^n$ such that $(\mathbf{u}')^T A = \begin{pmatrix} 0 & 0 & \cdots \end{pmatrix}$. Since the rows of A are linearly independent over R, $\mathbf{u}' = \mathbf{0}$. It follows that $\mathbf{u} = \mathbf{0}$, so the rows of A are linearly independent over K.

Now let A' be the reduced row echelon form of A (over K). Since the rows of A are linearly independent, then A' must have n linearly independent pivot columns, and the corresponding columns of A are also linearly independent. If we let B be the $n \times n$ submatrix of A with those columns, then $\det(B) \neq 0$. □

Recall that if M is a maximal ideal of a ring R, then $\overline{F}(X)$ denotes the power series over R/M obtained by reducing each coefficient of $F(X)$ modulo M.

Lemma 6.1.2. *Let R be an integral domain and let $F(X) \in R[[X]]$. Suppose there exists a set $\mathcal{S}$ of maximal ideals of R such that:*

(1) $\bigcap_{M \in \mathcal{S}} M = (0)$;

(2) There exists D such that for each $M \in \mathcal{S}$, $\overline{F}(X) \in (R/M)[[X]]$ is a rational function with numerator and denominator polynomials in $(R/M)[X]$ with degree at most D.

Then $F(X)$ is rational.

Proof. Write $F(X) = \sum_{n \geq 0} f_n X^n$ and let A be the $(D+1) \times \aleph_0$ matrix

$$A := \begin{pmatrix} f_1 & f_2 & \cdots \\ f_2 & f_3 & \cdots \\ \vdots & \vdots & \\ f_{D+1} & f_{D+2} & \cdots \end{pmatrix}.$$

We claim that every $(D+1) \times (D+1)$ submatrix of A has determinant 0. Suppose not. Then there exists a $(D+1) \times (D+1)$ submatrix B of A with nonzero determinant. Since $\bigcap_{M \in \mathcal{S}} M = (0)$, there is some $M \in \mathcal{S}$ such that $\det(B) \notin M$, that is, $\det(B) \not\equiv 0$

(mod M). However, by assumption, there exist $P(X), Q(X) \in R[X]$ with $Q(0) \notin M$ and $\deg(P(X)), \deg(Q(X)) \leq D$ such that

$$F(X)Q(X) \equiv P(X) \pmod{M}. \tag{6.1}$$

Write

$$Q(X) = q_0 + q_1 X + \cdots + q_D X^D.$$

If $n > D$, then it follows from (6.1) that the n-th coefficient in $F(X)Q(X)$ is 0 modulo M, that is

$$f_n q_0 + f_{n-1} q_1 + \cdots + f_{n-D} q_D \equiv 0 \pmod{M}$$

for all $n > D$. If we let

$$\mathbf{q} = \begin{pmatrix} q_D \\ v_{D-1} \\ \vdots \\ q_0 \end{pmatrix} \in (R/M)^{D+1},$$

then we have shown above that $\mathbf{q}^T A \equiv 0 \pmod{M}$, and so $\mathbf{q}^T B \equiv 0 \pmod{M}$, a contradiction.

By Lemma 6.1.1, the rows of A are linearly dependent over R, that is, there is a vector

$$\mathbf{s} = \begin{pmatrix} s_D \\ s_{D+1} \\ \vdots \\ s_0 \end{pmatrix} \in R^{D+1}$$

such that $\mathbf{s}^T A = 0$. Therefore

$$f_n s_0 + f_{n-1} s_1 + \cdots + f_{n-D} s_D = 0$$

for all $n > D$. Let

$$S(X) = s_0 + s_1 X + \cdots + s_D X^D;$$

we have shown that the n-th coefficient of $F(X)S(X)$ is 0 whenever $n > D$, so $F(X)S(X)$ is a polynomial in $R[X]$. □

Example 6.1.3. Let $R = \mathbb{Q}[Y]$ and let $F(X) = \sum_{n=0}^{\infty} YX^n \in R[[X]]$. In order to apply Lemma 6.1.2, we choose $\mathcal{S} = \{\langle Y - k\rangle : a \in \mathbb{N}\}$. Then $\bigcap_{M\in\mathcal{S}} M = (0)$. For $k \in \mathbb{N}$, let $F(X) \equiv \sum_{n=0}^{\infty} kX^n \equiv \sum_{n=0} \pmod{\langle Y - k\rangle}$. Therefore

$$F(X) \equiv \frac{k}{1-X} \pmod{\langle Y - k\rangle}.$$

If we pick $D = 2$, then $F(X)$ and $\mathcal{M}$ satisfy the conditions of Lemma 6.1.2, and we see that

$$F(X) = \frac{Y}{1-X} \in \mathbb{Q}(X).$$

6.2 Proof of the Main Result

The proof will be divided into several parts.

Claim 1. *There exists a finitely generated $\mathbb{Z}$-algebra $R \subset K$ such that $\phi(X) \in XR[[X]]$, and so $F(X) \in XR[[X]]$. Moreover, R is Jacobson.*

Proof of Claim 1. We can write $\phi(X) = \frac{P(X)}{Q(X)}$ where $P(X), Q(X) \in K[X]$, $Q(0) \neq 0$ and $P(0) = 0$. Let

$$P(X) = \sum_{i=0}^{m} u_i X^i \text{ and } Q(X) = \sum_{j=0}^{n} v_j X^j.$$

We have $u_0 = 0$ and $v_0 \neq 0$. Let $R = \mathbb{Z}[u_1, \ldots, u_m, v_0, \ldots, v_n, v_0^{-1}]$, that is, R is the $\mathbb{Z}$-algebra generated by $\{u_1, \ldots, u_m, v_0, \ldots, v_n, v_0^{-1}\}$. Since $\mathbb{Z}$ is Jacobson, then R is Jacobson by the Nullstellensatz.

Now

$$\begin{aligned}\phi(X) &= \frac{u_1X + \cdots + u_mX^m}{v_0 + \cdots + v_nX^n} \\ &= \frac{u_1X + \cdots + u_mX^m}{v_0(1 + v_0^{-1}v_1X + \cdots + v_0^{-1}v_nX^n)} \\ &= \frac{v_0^{-1}X(u_1 + \cdots + u_mX^{m-1})}{1 + C(X)}\end{aligned}$$

where $C(X) \coloneqq v_0^{-1}v_1X + \cdots + v_0^{-1}v_nX^n$. Thus

$$\phi(X) = v_0^{-1}X(u_1 + \cdots + u_mX^{m-1})(1 - C(X) + C(X)^2 - \cdots),$$

and so $F(X) \in XR[[X]]$. □

Claim 2. *Let R be defined as in the proof of Claim 1. If M is a maximal ideal of R, then R/M is a finite field.*

Proof of Claim 2. First we will show that

$$R \cong \mathbb{Z}[t_1, \ldots, t_s]/I,$$

where $s = m + n + 1$, $t_1, \ldots, t_s$ are indeterminants, and I is an ideal of $\mathbb{Z}[t_1, \ldots, t_s]$. To see this, note that

$$\begin{aligned} \phi : \mathbb{Z}[t_1, \ldots, t_s] &\to R \\ H(t_1, \ldots, t_s) &\mapsto H(u_1, \ldots, u_m, v_0, \ldots, v_n, v_0^{-1}) \end{aligned}$$

is a surjective ring homomorphism, so if we let $I = \ker(\phi)$, we have $R \cong \mathbb{Z}[t_1, \ldots, t_s]/I$ by the isomorphism theorem.

Let M be a maximal ideal of R. Since $R \cong \mathbb{Z}[t_1, \ldots, t_s]/I$, then by Corollary 2.4.11 there is a unique ideal M' of $\mathbb{Z}[t_1, \ldots, t_s]$ such that $R/M \cong \mathbb{Z}[t_1, \ldots, t_s]/M'$. This implies that M' is a maximal ideal of $\mathbb{Z}[t_1, \ldots, t_s]$.

Now, by the Nullstellensatz, $N \coloneqq M' \cap \mathbb{Z}$ is a maximal ideal of $\mathbb{Z}$ and $\mathbb{Z}[t_1, \ldots, t_s]/M'$ is a finite extension of $\mathbb{Z}/N$. Since $\mathbb{Z}/N$ is a finite field for any maximal ideal $N \subset \mathbb{Z}$, we can conclude that R/M is finite as well. □

We can now proceed to prove that $F(X)$ is either rational or transcendental, completing the proof of our main result.

Proof. Suppose that $F(X)$ is algebraic over $K(X)$. Then there exist $A_0(X), \ldots, A_s(X) \in K[X]$ with $A_s(X) \neq 0$ such that

$$A_0(X) + A_1(X)F(X) + \cdots + A_s(X)F(X)^s = 0. \tag{6.2}$$

We wish to show that $F(X)$ satifies a relation of this type with coefficient polynomials in $R[X]$. Let K_0 denote the field of fractions of R, then $F(X) \in K_0[[X]]$. Let $\mathcal{B}$ be a basis for the extension K/K_0. We claim that $\mathcal{B}$ is linearly independent over $K_0[[X]]$. To see this, let $\beta_1, \ldots, \beta_l \in \mathcal{B}$ and $G_1(X), \ldots, G_l(X) \in K_0[[X]]$ such that

$$\sum_{i=1}^{l} G_i(X)\beta_i = 0 \tag{6.3}$$

in $K[[X]]$. For each $1 \leq i \leq l$ we write

$$G_i(X) = \sum_{n=0}^{\infty} a_n^{(i)} X^n$$

where $a_n^{(i)} \in K_0$ for all $n \geq 0$. Then for each $n \geq 0$ we have $\sum_{i=1}^{l} a_n^{(i)} \beta_i = 0$ by (6.3). Since $\mathcal{B}$ is linearly independent over K_0, then $a_n^{(i)} = 0$ for every $1 \leq i \leq l$ and $n \geq 0$. Thus $G_i(X) = 0$ for all $1 \leq i \leq l$, and so $\mathcal{B}$ is linearly independent over $K_0[[X]]$.

Since $F(X)$ is algebraic over $K(X)$, we can write

$$A_0(X) = a_0 + a_1 X + \cdots + a_e X^e.$$

Then for each i we can write

$$a_i = \sum_{\beta \in \mathcal{B}} a_i^{(\beta)} \beta,$$

where $a_i^{(\beta)}$ is in K_0 for all $\beta \in \mathcal{B}$ and is nonzero for only finitely many β. Thus

$$\begin{aligned} A_0(X) &= \sum_{i=0}^{e} \sum_{\beta \in \mathcal{B}} a_i^{(\beta)} \beta X^i \\ &= \sum_{\beta \in \mathcal{B}} \left(\sum_{i=0}^{e} a_i^{(\beta)} X^i \right) \beta \\ &= \sum_{\beta \in \mathcal{B}} A_0^{(\beta)}(X) \beta, \end{aligned}$$

where

$$A_0^{(\beta)} := a_0^{(\beta)} + \cdots + a_e^{(\beta)} X^e.$$

By the same construction we can write

$$A_i(X) = \sum_{\beta \in \mathcal{B}} A_i^{(\beta)}(X) \beta$$

for all $1 \leq i \leq s$ where $A_i^{(\beta)}(X)$ is in $K_0[X]$ for all $\beta \in \mathcal{B}$ and is nonzero for only finitely many β.

We can now rewrite (6.2) as

$$\sum_{i=0}^{s} \sum_{\beta \in \mathcal{B}} A_i^{(\beta)}(X) \beta F(X)^i = 0,$$

and so

$$\sum_{\beta\in\mathcal{B}}\left(\sum_{i=0}^{s} A_i^{(\beta)}(X)F(X)^i\right)\beta = 0.$$

Since $\sum_{i=0}^{s} A_i^{(\beta)}(X)F(X)^i \neq 0$ for at most finitely many $\beta \in \mathcal{B}$ and we have shown that $\mathcal{B}$ is linearly independent over $K_0[[X]]$, we see that

$$\sum_{i=0}^{s} A_i^{(\beta)}(X)F(X)^i = 0 \tag{6.4}$$

for all $\beta \in \mathcal{B}$, and so $F(X)$ is algebraic over $K_0(X)$. By choosing $\beta \in \mathcal{B}$ such that $A_0^{(\beta)}(X), \ldots, A_s^{(\beta)}(X)$ are not all zero and clearing denominators in (6.4), we obtain the relation

$$B_0(X) + B_1(X)F(X) + \cdots + B_t(X)F(X)^t = 0,$$

where $B_0(X), \ldots, B_t(X) \in R[X]$ and $B_t(X) \neq 0$, as desired.

Now if M is a maximal ideal of R and $B_1(X), \ldots, B_t(X)$ are not all 0 modulo M, then $\overline{F}(X) \in (R/M)[[X]]$ is algebraic over $(R/M)(X)$. In particular, if b is the leading coefficient of $B_t(X)$, then $\overline{F}(X)$ is algebraic over $(R/M)(X)$ whenever $b \notin M$.

We now wish to use Lemma 6.1.2. Let D equal the larger of the degrees of the numerator and denominator polynomials of $\phi(X)$ and let $\mathcal{S}$ be the set of all maximal ideals M of R not containing b such that d is not a power of $\mathrm{char}(R/M)$. If every maximal ideal of R is in $\mathcal{S}$, then $\bigcap_{M\in\mathcal{S}} M = J(R) = (0)$. Otherwise, if M is maximal in R and $M \notin \mathcal{S}$, then $d \in M$ or $b \in M$. Setting $I := \bigcap_{M\in\mathcal{S}} M$ and letting $\mathcal{M}$ be the collection of all maximal ideals of R, we see that

$$\begin{aligned}
(0) &= J(R) \\
&= \left(\bigcap_{M\in\mathcal{S}} M\right) \cap \left(\bigcap_{M\in\mathcal{M}\setminus\mathcal{S}} M\right) \\
&= I \cap \left(\left(\bigcap_{\substack{M\in\mathcal{M}\\ b\in M}} M\right) \cap \left(\bigcap_{\substack{M\in\mathcal{M}\\ d\in M}} M\right)\right) \\
&\supset I \cap ((b) \cap (d)) \\
&\supset I(db).
\end{aligned}$$

Since R is an integral domain, it follows that $I = (0)$.

For any maximal ideal M of R, R/M is a finite field by Claim 2, and we have shown above that $\overline{F}(X) \in (R/M)[[X]]$ is algebraic over $(R/M)(X)$ for any $M \in \mathcal{S}$. Thus, if $M \in \mathcal{S}$, then $\overline{F}(X)$ is rational by Theorem 5.3.1 with numerator and denominator polynomials having degree at most D. Therefore $F(X)$ is rational by Lemma 6.1.2. □

Bibliography

[AS03] Jean-Paul Allouche and Jeffrey Shallit. *Automatic sequences: Theory, applications, generalizations*. Cambridge University Press, Cambridge, 2003.

[ASST01] S. D. Adhikari, N. Saradha, T. N. Shorey, and R. Tijdeman. Transcendental infinite sums. *Indagationes Mathematicae*, 12(1):1 – 14, 2001.

[Büc60] J. Richard Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik Grundlagen Math.*, 6:66–92, 1960.

[Car58] Pierre Cartier. Questions de rationalité des diviseurs en géométrie algébrique des diviseurs en géométrie algébrique. *Bull. Soc. Math. France*, 86:177–251, 1958.

[Chr79] Gilles Christol. Ensembles presque périodiques k-reconnaissables. *Theoret. Comput. Sci.*, 9(1):141–145, 1979.

[CKMFR80] G. Christol, T. Kamae, M. Mendès France, and G. Rauzy. Suites algébriques, automates et substitutions. *Bull. Soc. Math. France*, 108(4):401–419, 1980.

[Cob69] Alan Cobham. On the base-dependence of sets of numbers recognizable by finite automata. *Math. Systems Theory*, 3:186–192, 1969.

[Cob72] Alan Cobham. Uniform tag sequences. *Math. Systems Theory*, 6:164–192, 1972.

[dP51] A. de Polignac. *Recherches nouvelles sur les nombres premiers*. Bachelier, 1851.

[EHM05] Gove Effinger, Kenneth Hicks, and Gary L. Mullen. Integers and polynomials: comparing the close cousins $\mathbf{Z}$ and $\mathbf{F}_q[x]$. *Math. Intelligencer*, 27(2):26–34, 2005.

[Eil74] Samuel Eilenberg. *Automata, languages, and machines. Vol. A*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York, 1974. Pure and Applied Mathematics, Vol. 58.

[Eis95] David Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. With a view toward algebraic geometry.

[Gal10] Joseph A Gallian. *Contemporary abstract algebra*. Brooks/Cole, Cengage Learning, Belmont, CA, 7th ed edition, 2010.

[Gol51] Oscar Goldman. Hilbert rings and the Hilbert Nullstellensatz. *Mathematische Zeitschrift*, 54(2):136–140, 1951.

[Han82] G. Hansel. A propos d'un théorème de Cobham. In Dominique Perrin, editor, *Actes de la Fête des Mots*, pages 55–59. Greco de Programmation, CNRS, Rouen, 1982.

[Har77] Robin Hartshorne. *Algebraic geometry*. Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.

[Hay65] D. R. Hayes. A Goldbach theorem for polynomials with integral coefficients. *The American Mathematical Monthly*, 72(1):pp. 45–46, 1965.

[Hil93] David Hilbert. Über die vollen invariantensysteme. *Mathematische Annalen*, 42(3):313–373, 1893.

[HJ99] Karel Hrbacek and Thomas Jech. *Introduction to set theory*, volume 220 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, third edition, 1999.

[HLPPA09] Xiang-Dong Hou, Sergio R. López-Permouth, and Benigno R. Parra-Avila. Rational power series, sequential codes and periodicity of sequences. *J. Pure Appl. Algebra*, 213(6):1157–1169, 2009.

[Huf54] D. A. Huffman. The synthesis of sequential switching circuits. I, II. *J. Franklin Inst.*, 257:161–190, 275–303, 1954.

[Irv04] Ronald S. Irving. *Integers, polynomials, and rings*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 2004. A course in algebra.

[LN97] Rudolf Lidl and Harald Niederreiter. *Finite fields*, volume 20 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 1997. With a foreword by P. M. Cohn.

[Mea55] George H. Mealy. A method for synthesizing sequential circuits. *Bell System Tech. J.*, 34:1045–1079, 1955.

[Moo56] Edward F. Moore. Gedanken-experiments on sequential machines. In *Automata studies*, Annals of mathematics studies, no. 34, pages 129–153. Princeton University Press, Princeton, N. J., 1956.

[Mor21] Harold Marston Morse. Recurrent geodesics on a surface of negative curvature. *Trans. Amer. Math. Soc.*, 22(1):84–100, 1921.

[MP90] Warren McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 52:99–115, 1990. 10.1007/BF02459570.

[Nak51] Tadasi Nakayama. A remark on finitely generated modules. *Nagoya Math. J.*, 3:139–140, 1951.

[Per90] Dominique Perrin. Finite automata. In *Handbook of theoretical computer science, Vol. B*, pages 1–57. Elsevier, Amsterdam, 1990.

[Pol08] Paul Pollack. A polynomial analogue of the twin prime conjecture. *Proc. Amer. Math. Soc.*, 136(11):3775–3784, 2008.

[Pol11] Paul Pollack. The exceptional set in the polynomial Goldbach problem. *Int. J. Number Theory*, 7(3):579–591, 2011.

[Pro51] Eugène Prouhet. Mémoire sur quelques relations entre les puissances des nombres. *C. R. Acad. Sci. Paris*, 33:225, 1851.

[Roy90] Ranjan Roy. The discovery of the series formula for π by Leibniz, Gregory and Nilakantha. *Math. Mag.*, 63(5):291–306, 1990.

[RW06] Michel Rigo and Laurent Waxweiler. A note on syndeticity, recognizable sets and Cobham's theorem. *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS*, (88):169–173, 2006.

[Thu12] Axel Thue. Über die gegenseitige lage gleicher teile gewisser zeichenreihen. *Norskevid. Selsk. Skr. I. Mat. Nat. Kl.*, 1:1–67, 1912.

[Zor35] Max Zorn. A remark on method in transfinite algebra. *Bull. Amer. Math. Soc.*, 41(10):667–670, 1935.