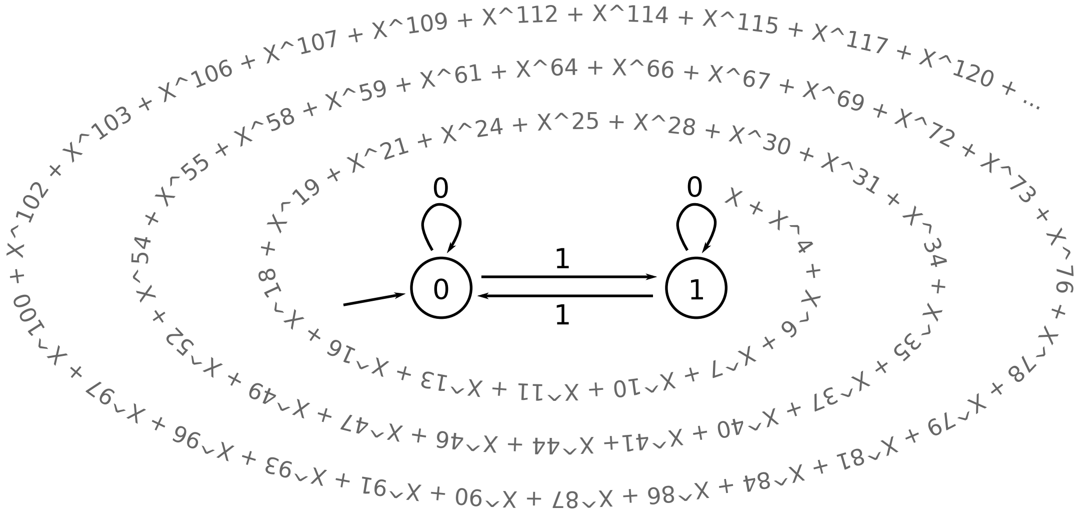
university of groningen

faculty of mathematics
and natural sciences

From finite automata to power series and back again

Anneroos Everts


0
0
1
1
0
1
X + X^4 + X^6 + X^7 + X^10 + X^11 + X^13 + X^16 + X^18 + X^19 + X^21 + X^24 + X^25 + X^28 + X^30 + X^31 + X^34 + X^35 + X^37 + X^40 + X^41+ X^44 + X^46 + X^47 + X^49 + X^52 + X^54 + X^55 + X^58 + X^59 + X^61 + X^64 + X^66 + X^67 + X^69 + X^72 + X^73 + X^76 + X^78 + X^79 + X^81 + X^84 + X^86 + X^87 + X^90 + X^91 + X^93 + X^96 + X^97 + X^100 + X^102 + X^103 + X^106 + X^107 + X^109 + X^112 + X^114 + X^115 + X^117 + X^120 + ...

Master Thesis in Mathematics

February 22, 2012

From finite automata to power series and back again

Summary

In this thesis we examine the steps of Christol's theorem and Ore's lemma, to find answers to the following two questions. Given a finite q-automaton over $\mathbb{F}_q$ with m states, what can we say about the algebraic degree of the corresponding algebraic power series over $\mathbb{F}_q$? Conversely, given an algebraic power series of algebraic degree d, can we find a bound on the number of states of generating minimal automaton?

We discuss some special cases and give answers to the questions above: Given a finite q-automaton with m states, the degree of the corresponding power series is at most $q^m - 1$. Conversely, given an algebraic power series F in $\mathbb{F}_q[[X]]$ that satisfies a polynomial of degree d, with coefficients in $\mathbb{F}_q[X]$ that have degree at most A, then we can bound the number of states of a minimal generating automaton with a bound that is doubly exponential in d.

Master Thesis in Mathematics
Author: Anneroos Everts
First supervisor: prof. dr. Jaap Top
Second supervisor: prof. dr. Holger Waalkens
Date: February 22, 2012

Institute of Mathematics and Computing Science
P.O. Box 407
9700 AK Groningen
The Netherlands

Contents

1
Introduction

Finite automata and power series are linked in an interesting way by Christol's theorem: a power series $F = \sum_{n \geq 0} a_n X^n$ over $\mathbb{F}_q$, with q a prime power, is algebraic over $\mathbb{F}_q(X)$ if and only if its coefficients a_n are generated by a finite automaton [Christol et al., 1980]. This theorem gives an interesting connection between finite automata, a subject from computer science, and the algebraic concept of formal power series. With this theorem we can transfer properties from finite automata to power series, and vice versa.

A finite automaton can be visualized as a graph with a finite number of nodes, which are called the states, with directed edges between them. An automaton takes a string of symbols as input. Starting from the initial state, it moves from state to state along the directed edges, according to the symbols it reads one by one. When the last symbol is read the automaton produces an output that corresponds to the last state reached. If an automaton takes as input the q-ary representation $(n)_q$ of a non-negative number n, we call it a q-automaton. Given a q-automaton for q a prime power, let a_n denote the output corresponding to $(n)_q$. Then we say that the q-automaton generates the power series $F = \sum_{n \geq 0} a_n X^n$ over $\mathbb{F}_q$. Christol's theorem states that this power series is algebraic.

In this thesis we answer the following two questions:

- Given a finite automaton with m states, what can we say about the algebraic degree of the corresponding power series?
- Conversely, given an algebraic power series of algebraic degree d, can we find a bound on the number of states of a minimal automaton that generates it?

We will give answers to these questions by closely following the steps in the proofs of Christol's theorem and a lemma by Ore.

For this thesis we used the excellent book *Automatic sequences, Theory, Applications, Generalizations* by Allouche and Shallit [2003] as the main reference. Most of the definitions, theorems and notation in Chapter 2 are adopted from this book. We used Chapter 3 of *Substitutions in Dynamics, Arithmetics and Combinatorics* by Fogg, Berthé, Ferenczi, Mauduit, and Siegel [2002], for the proof of Ore's lemma.

This thesis is organized as follows. In the first two sections of Chapter 2 we introduce finite automata and automatic sequences. In Section 2.3, we state and prove Christol's theorem. We briefly discuss Furstenberg's theorem on diagonals of rational multivariate power series in

Section 2.4, and use the result in two detailed examples in the Section 2.5. The main results of this thesis, the answers to the two questions above, are stated in Chapter 3. In Chapter 4 we summarize our results and give suggestions for further research.

2

Finite automata and automatic sequences

In this chapter we introduce the concepts of finite automata and automatic sequences. We start with some definitions and short examples in the first two sections, and we state and prove some lemmas. We use these lemmas to prove Christol's theorem in Section 2.3. In Section 2.4 we briefly discuss Furstenberg's theorem, and use it in Section 2.5 to construct two detailed examples.

2.1 Finite automata

A finite automaton is a model of computation, with a finite number of states and transitions. It takes as input a word: a string of symbols from a given alphabet. Starting from the initial state, it moves from state to state for every symbol it reads. When the last symbol is read, the finite automaton produces an output that corresponds to the last state reached.

Formally, a finite automaton is defined to be a 6-tuple $M = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ where

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

We can represent a finite automaton with a *transition diagram*, which is a directed graph where every vertex represents a state q_i, see for example Figure 2.1. The transition function δ is represented by directed edges that are labeled with a symbol from alphabet Σ. The initial state is indicated by an unlabeled arrow. Every vertex has a label q_i/a, where q_i is the name of the vertex and a is the output that corresponds to the state q_i, so $\tau(q_i) = a \in \Delta$. Sometimes the name q_i of a vertex is omitted, and only the symbol a for the output is given.

In this thesis, every automaton is a *reverse reading deterministic finite automaton with output*, as described in [Allouche and Shallit, 2003, Chapter 4]. This means that the automaton reads the symbols from right to left, as opposed to the more common forward reading. We have chosen to use reverse reading automata, because we use them in all of our proofs. The two definitions are equivalent for all results in this chapter [Allouche and Shallit, 2003, Chapter 5].

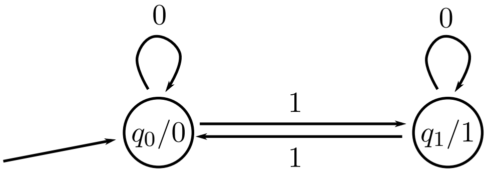


Figure 2.1: An example of an automaton with two states.

A standard example of a finite automaton is presented in Figure 2.1. Although this automaton has only two states, it is not trivial. Both the input and output alphabets are $\{0,1\}$. For every 0 that the automaton reads, the automaton stays in the same state. For every 1 it reads, it moves to the other state. So $\delta(q_0,1)=q_1$, $\delta(q_1,1)=q_0$ and $\delta(q_i,0)=q_i$ for $i\in\{0,1\}$. For example, for the input 101101, the automaton visits the states q_0,q_1,q_1,q_0,q_1,q_1 and ends in q_0, and hence the output is 0. We see that for a given input w the automaton gives output 0 if and only if w contains an even number of ones. Otherwise, the automaton ends in q_1 and gives output 1.

To link the input and output of an automaton directly, we extend the domain of δ. We define Σ^* to be the set of all finite words that can be made with symbols from the alphabet Σ, including the empty word ϵ. For example, for the alphabet $\Sigma_2=\{0,1\}$ we have $\Sigma_2^*=\{\epsilon,0,1,00,01,10,11,000,001,010,...\}$. We can now extend the domain of δ to $Q\times\Sigma^*$. First, define δ for the empty string ϵ, for all states $q\in Q$:

$$\delta(q,\epsilon)=q.$$

Next, for all $w\in\Sigma^*$, $a\in\Sigma$ and $q\in Q$, define

$$\delta(q,aw)=\delta(\delta(q,w),a).$$

With this extension of the domain of δ, $\tau(\delta(q_0,w))$ is the output generated by the automaton $(Q,\Sigma,\delta,q_0,\Delta,\tau)$ for a given input string $w\in\Sigma^*$. For example, if the input for the automaton in Figure 2.1 is 101011001, then the output is given by $\tau(\delta(q_0,101011001))=1$.

2.2 Automatic sequences

In this thesis, we focus on finite automata that take as input the representation of an integer n in base $k\geq 2$, so the input alphabet is $\Sigma_k=\{0,1,2,\ldots,k-1\}$. Such an automaton is called a *k-automaton*. Since every non-negative integer n can be expressed in a unique way as $n=\sum_{i=0}^{t}c_ik^i$ with $c_t\neq 0$ and $c_i\in\Sigma_k$ for $0\leq i\leq t$, we can define the *canonical base-k representation* of n as $(n)_k=c_tc_{t-1}\cdots c_1c_0\in\Sigma_k^*$. Conversely, given a string $w\in\Sigma_k^*$ of length $|w|$, $[w]_k$ denotes the non-negative number $n=\sum_{i=0}^{|w|-1}w_ik^i$. Clearly we have $[(n)_k]_k=n$, but $([w]_k)_k$ is in general not equal to w.

A k-automaton can be used to generate an infinite sequence $(a_n)_{n\geq 0}$, where a_n is the output that corresponds to the input $(n)_k$. This sequence is then called k-automatic. More formally:

Definition 1. An infinite sequence $\mathbf{a}=(a_n)_{n\geq 0}$ over a finite alphabet Δ is called *k-automatic* if there exists a k-automaton $M=(Q,\Sigma_k,\delta,q_0,\Delta,\tau)$ such that $a_n=\tau(\delta(q_0,w))$ for all $n\geq 0$ and all w with $[w]_k=n$.

Note that this definition implies that leading zeros do not make any difference in the output of M: if two words $w_1,w_2\in\Sigma_k^*$ satisfy $w_1=0^tw_2$, with 0^t a string of t zeros, then they represent the same number $[w_1]_k=[w_2]_k$, so the automaton M must satisfy $\tau(\delta(q_0,w_1))=\tau(\delta(q_0,w_2))$. A direct consequence is that each edge labeled with a 0 of such a k-automaton should connect two states with the same output label. We call automata with this property *leading zeros invariant.* Any k-automaton with this property generates a k-automatic sequence $(a_n)_{n\geq 0}$. If a sequence is k-automatic, then there exist both a forward and a reverse reading automaton, see [Allouche and Shallit, 2003, Chapter 5].

As an example, consider the 2-automaton in Figure 2.1. This automaton generates the sequence

$$\mathbf{b} := (b_n)_{n\geq 0} = (0,1,1,0,1,0,0,1,1,0,0,1,0,1,1,0,\ldots).$$

This sequence is known as the Thue-Morse sequence [Thue, 1912]. Since multiplying a non-negative integer n by two is equivalent to adding a 0 to its binary representation, we see that $b_n = b_{2n}$. Similarly, multiplying n by two and adding one is equivalent to adding a 1 to $(n)_2$, so $b_{2n+1} = 1 - b_n$. Together with $b_0 = 0$, these recurrence relations are another way to define the Thue-Morse sequence.

2.2.1 Pointwise sum and product

Let $\mathbf{a} = (a_n)_{n\geq 0}$ and $\mathbf{b} = (b_n)_{n\geq 0}$ be two sequences over an alphabet Δ. If addition and multiplication are defined for Δ, then we can define their *pointwise sum* $(a_n + b_n)_{n\geq 0}$ and *pointwise product* $(a_n b_n)_{n\geq 0}$. A result of the following lemma is that if $\mathbf{a}$ and $\mathbf{b}$ are k-automatic, then so are their pointwise sum and pointwise product.

Lemma 2. *Let $\mathbf{a} = (a_n)_{n\geq 0}$ and $\mathbf{b} = (b_n)_{n\geq 0}$ be two k-automatic sequences over the finite alphabets Δ_1 and Δ_2 respectively. Let ρ be a function from $\Delta_1 \times \Delta_2$ into the finite alphabet Δ_3. Then the sequence $(\rho(a_n, b_n))_{n\geq 0}$ is also k-automatic.*

Proof. Since $\mathbf{a}$ and $\mathbf{b}$ are k-automatic, there are k-automata $M_1 = (Q_1, \Sigma_k, \delta_1, q_{0_1}, \Delta_1, \tau_1)$ and $M_2 = (Q_2, \Sigma_k, \delta_2, q_{0_2}, \Delta, \tau_2)$ that generate $\mathbf{a}$ and $\mathbf{b}$ respectively. Define

$$M_3 = (Q_1 \times Q_2, \Sigma_k, \delta_3, [q_{0_1}, q_{0_2}], \Delta_1 \times \Delta_2, \tau_3),$$

where δ_3 and τ_3 are defined as:

$$\begin{aligned}\delta_3([q_1, q_2], c) &= [\delta_1(q_1, c), \delta_2(q_2, c)] \ \in Q_1 \times Q_2\\ \tau_3([q_1, q_2]) &= [\tau_1(q_1), \tau_2(q_2)] \ \in \Delta_1 \times \Delta_2\end{aligned}$$

for all $q_1 \in Q_1$, $q_2 \in Q_2$ and $c \in \Sigma_k$. The k-automaton M_3 generates $\mathbf{a} \times \mathbf{b} = ([a_n, b_n])_{n\geq 0}$, which is hence k-automatic. Finally, the k-automaton

$$M_3' = (Q_1 \times Q_2, \Sigma_k, \delta_3, [q_{0_1}, q_{0_2}], \Delta_1 \times \Delta_2, \rho \circ \tau_3),$$

generates $\rho(\mathbf{a} \times \mathbf{b}) = (\rho(a_n, b_n))_{n\geq 0}$, so this sequence is k-automatic.

□

2.2.2 The k-kernel

The *k-kernel* $K_k(\mathbf{a})$ of an infinite sequence $\mathbf{a} = (a_n)_{n\geq 0}$ is defined to be the set of subsequences

$$K_k(\mathbf{a}) = \{(a_{k^i \cdot n + j})_{n\geq 0} : i \geq 0 \text{ and } 0 \leq j < k^i\}.$$

The k-kernel $K_k(\mathbf{a})$ can be finite or infinite, but it always contains the sequence $\mathbf{a}$ itself, since $\mathbf{a}$ corresponds to the subsequence with $i = j = 0$.

As an example, we consider the 2-kernel of Thue-Morse sequence $\mathbf{b} = (0, 1, 1, 0, 1, 0, 0, 1, \ldots)$. Besides $\mathbf{b}$ itself, the 2-kernel contains the subsequence corresponding to $i = 1$ and $j = 0$:

$$(b_{2n+0})_{n \geq 0} = (b_0, b_2, b_4, b_6, b_8, \ldots) = (0, 1, 1, 0, 1, \ldots).$$

This subsequence equals $\mathbf{b}$, since we already saw that $b_n = b_{2n}$. Using $b_{2n+1} = 1 - b_n$ we see that the subsequence corresponding to $i = j = 1$, given by

$$(b_{2n+1})_{n \geq 0} = (b_1, b_3, b_5, b_7, b_9, \ldots) = (1, 0, 0, 1, 0, \ldots),$$

is equal to $(1 - b_n)_{n \geq 0}$. The relations $(b_{2n})_{n \geq 0} = (b_n)_{n \geq 0}$ and $(b_{2n+1})_{n \geq 0} = (1 - b_n)_{n \geq 0}$ can be used to prove that these two subsequences are in fact the only two elements of $K_2(\mathbf{b})$.

Lemma 3. *Let $k \geq 2$. A sequence $\mathbf{a} = (a_n)_{n \geq 0}$ is k-automatic if and only if $K_k(\mathbf{a})$ is finite.*

Proof. $\Rightarrow$: Since $\mathbf{a}$ is k-automatic, there exists a k-automaton $(Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ such that $a_n = \tau(\delta(q_0, 0^t(n)_k))$ for all $n, t \geq 0$. Given $i \geq 0$ and $0 \leq j < k^i$, let $w \in \Sigma_k^*$ be the word such that $|w| = i$ and $[w]_k = j$. We will show that for these i and j, the subsequence $(a_{k^i n + j})_{n \geq 0}$ is generated by the k-automaton $(Q, \Sigma_k, \delta, q, \Delta, \tau)$, where $q = \delta(q_0, w)$. Since there are only finitely many choices for q, the finiteness of $K_k(\mathbf{a})$ follows.

With i, j, w and q as above, we have for $n > 0$ that $(k^i n)_k = (n)_k 0^i$, so

$$(k^i n + j)_k = (n)_k w.$$

Hence for $n > 0$ we have

$$\delta(q_0, (k^i n + j)_k) = \delta(q_0, (n)_k w) = \delta(\delta(q_0, w), (n)_k) = \delta(q, (n)_k).$$

For $n = 0$ we have $(k^i n + j)_k = (j)_k$ and $w = 0^t (j)_k$ for some $t \geq 0$. So we have for $n = 0$:

$$\delta(q_0, (k^i n + j)_k) = \delta(q_0, (j)_k) = \delta(q_0, 0^t (j)_k) = \delta(q_0, w) = q = \delta(q, (0)_k).$$

So we see that the subsequence $(a_{k^i n + j})_{n \geq 0}$ is generated by the k-automaton $(Q, \Sigma_k, \delta, q, \Delta, \tau)$, which completes this part of the proof.

$\Leftarrow$: We can partition Σ_k^* with the following equivalence relation: for $w, x \in \Sigma_k^*$ we have

$$w \equiv x \quad \Leftrightarrow \quad a_{k^{|w|} \cdot n + [w]_k} = a_{k^{|x|} \cdot n + [x]_k} \quad \text{for all } n \geq 0.$$

The number of equivalence classes is equal to the number of elements in the k-kernel of $\mathbf{a}$, and hence finite. We can use these equivalence classes as the states of a k-automaton $M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$, where

$$\begin{aligned} Q &= \{\, [x] : x \in \Sigma_k^* \,\} \\ \delta([x], c) &= [cx] \quad \forall\, c \in \Sigma_k, \\ \tau([w]) &= a_{[w]_k}, \\ q_0 &= [\epsilon]. \end{aligned}$$

Before we prove that M generates $\mathbf{a}$, we need to check that δ and τ are well-defined. So we have to check if $[w] = [x]$ implies $\delta([w], c) = \delta([x], c)$ for all $c \in \Sigma$, and $\tau([w]) = \tau([x])$. Firstly, if $[w] = [x]$ then

$$a_{k^{|w|} \cdot n + [w]_k} = a_{k^{|x|} \cdot n + [x]_k} \qquad \forall n \geq 0. \tag{2.1}$$

Since this holds for all n, it also holds for $n = km + c$, for all $m \geq 0$. This implies:

$$a_{k^{|cw|}\cdot m+[cw]_k} = a_{k^{|cx|}\cdot n+[cx]_k} \qquad \forall m \geq 0.$$

So $[cw] = [cx]$, hence $\delta([x], c) = \delta([w], c)$. Secondly, if $[w] = [x]$, then taking $n = 0$ in (2.1) gives us $a_{[w]_k} = a_{[x]_k}$. So $\tau([w]) = \tau([x])$, hence τ is also well-defined.

By induction on the length of w, we see that $\delta(q_0, w) = [w]$. Hence $\tau(\delta(q_0, w)) = \tau([w]) = a_{[w]_k}$ for all $w \in \Sigma^*$, so M generates $\mathbf{a}$.

□

Lemma 4. *For all $m \geq 1$, a sequence $\mathbf{a} = (a_n)_{n\geq 0}$ is k-automatic if and only if it is k^m-automatic.*

Proof. $\Rightarrow$: Suppose that $\mathbf{a}$ is k-automatic. By Lemma 3, we know that the k-kernel $K_k(\mathbf{a})$ is finite. Since $K_{k^m}(\mathbf{a})$ is a subset of $K_k(\mathbf{a})$, it follows that the k^m-kernel is finite. Using Lemma 3 again, we find that $\mathbf{a}$ is k^m-automatic.

$\Leftarrow$: Since $\mathbf{a}$ is k^m-automatic, it is generated by k^m-automaton $M = (Q, \Sigma_{k^m}, \delta, q_0, \Delta, \tau)$, so $a_n = \tau(\delta(q_0, (n)_{k^m}))$. The idea of the proof is to use this k^m-automaton as a basis to make a k-automaton $N = (Q', \Sigma_k, \delta', q_0, \Delta, \tau')$ that generates $\mathbf{a}$. We use the fact that for every $b \in \Sigma_{k^m}$ there is a unique string $b_{m-1} \dots b_1 b_0$ of length m in Σ_k^m such that $[b]_{k^m} = [b_{m-1} \dots b_1 b_0]_k$. For each state of M, we will replace the k^m outgoing edges by a tree of states, representing the choices of $b_0, b_1, \dots, b_{m-1}$. See Figure 2.2 for an example of a 4-automaton and a 2-automaton that generate the same sequence.

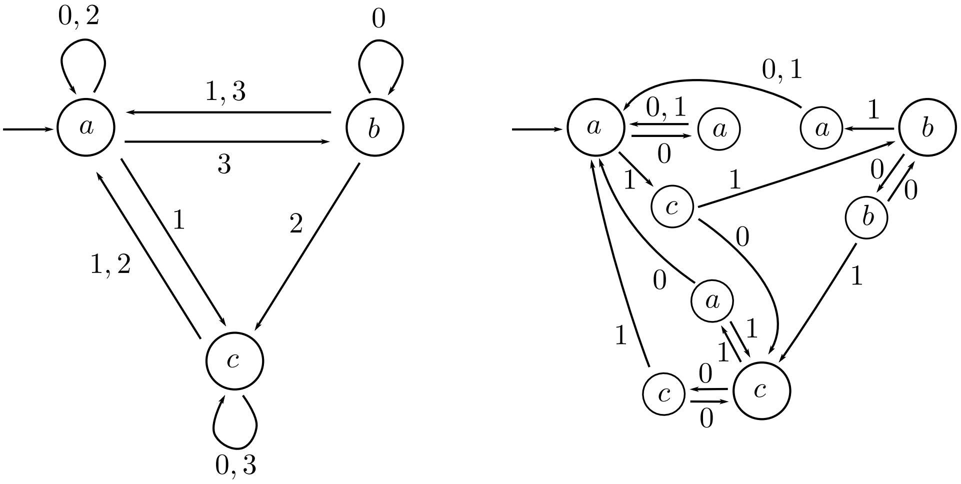


Figure 2.2: These 4-automaton and 2-automaton generate the same sequence.

More precisely, we start with the automaton M and delete its edges, but keep the states. Connect each state $q \in Q$ to k new states with edges labeled with $0, 1, \dots, k-1$ respectively, representing the k choices of b_0. Connect each just created state to its own k new states in the same way, so with edges labeled with the k choices of b_1. Continue this process, until for each $q \in Q$ a tree of depth $m - 1$ is created, with root q and k^{m-1} different leaves. Every leaf corresponds to a certain q and word $b_{m-2} \cdots b_1 b_0$. For each leaf, and each value of b_{m-1}, we connect the leaf to $\delta(q, b)$, where $[b]_{k^m} = [b_{m-1} \cdots b_1 b_0]_k$, and label the edge with the value of b_{m-1}. As a result, we have that $\delta(q, b) = \delta'(q, b_{m-1} \cdots b_1 b_0)$ for all $q \in Q$, $b \in \Sigma_{k^m}$ and $b_{m-1}, \dots, b_0 \in \Sigma_k$ such that $[b]_{k^m} = [b_{m-1} \cdots b_1 b_0]_k$. Furthermore, $\delta'(q, c)$ is defined for all $q \in Q'$ and all $c \in \Sigma_k$.

We extend the output function τ of$M$ to an output function τ' for N. First, let $\tau'(q) = \tau(q)$ for $q \in Q$. To make sure that leading zeros do not make any difference in the output, give every state in $Q' \setminus Q$ the same output as the state of Q to which it is connected by a path of zeros. Now all states $q \in Q'$ have an output label.

With induction on the length of w we have that $a_n = \tau'(\delta'(q_0, w))$ for all w such that $[w]_k = n$. Thus, this k-automaton N generates $\mathbf{a}$, which is therefore k-automatic.

□

2.3 Christol's Theorem

A formal power series $F(X) = \sum_{n\geq 0} a_n X^n$ in the ring $\mathbb{F}_q[[X]]$ of formal power series over $\mathbb{F}_q$, with $q = p^k$ for some prime p, corresponds to an infinite sequence $\mathbf{a} = (a_n)_{n\geq 0}$ over Σ_q. Christol's theorem states that a power series F is algebraic over $\mathbb{F}_q(X)$ if and only if the corresponding sequence is p-automatic. For example, the algebraic power series $\sum_{n\geq 0} X^n = \frac{1}{1+X}$ in $\mathbb{F}_2[[X]]$ corresponds to the sequence $(1)_{n\geq 0}$ over Σ_2, which is clearly 2-automatic.

To prove Christol's theorem we need the following $\mathbb{F}_q$-linear transformation. For $0 \leq r < q$, let Λ_r be the map from $\mathbb{F}_q[[X]]$ to $\mathbb{F}_q[[X]]$ defined by

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

Note that the sequence corresponding to $\Lambda_r(F)$ is a subsequence of $\mathbf{a}$ and is an element of the kernel $K_k(\mathbf{a})$.

Lemma 5. *Let F and G be two formal power series in $\mathbb{F}_q[[X]]$, then the following properties hold:*

$$\begin{aligned} &(a) \quad F(X) = \sum_{0\leq r<q} X^r \big(\Lambda_r F(X)\big)^q, \\ &(b) \quad \Lambda_r(F^q G) = F\Lambda_r(G), \qquad \forall\ 0 \leq r < q. \end{aligned}$$

Proof. (a): We have

$$\begin{aligned} F(X) &= \sum_{n\geq 0} a_n X^n = \sum_{0\leq r<q} \sum_{n\geq 0} a_{qn+r} X^{qn+r} = \sum_{0\leq r<q} X^r \sum_{n\geq 0} a_{qn+r} X^{qn} \\ &= \sum_{0\leq r<q} X^r \Big(\sum_{n\geq 0} a_{qn+r} X^n\Big)^q = \sum_{0\leq r<q} X^r \Lambda_r\big(F(X)\big)^q. \end{aligned}$$

(b): Use part (a) to write $G = \sum_{r=0}^{q-1} \Lambda_r(G)^q X^r$, then

$$F^q G = \sum_{r=0}^{q-1} \big(F\Lambda_r(G)\big)^q X^r.$$

In general, for a power series $B = \sum_{n\geq 0} b_n X^n$ and $r, s \in \{0, \ldots, q-1\}$ it holds that $\Lambda_r(B^q X^s) = \Lambda_r(\sum_{n\geq 0} b_n X^{qn+s}) = 0$ if $r \neq s$ and $\Lambda_r(B^q X^s) = B$ if $r = s$. So for $0 \leq r < q$ we have

$$\Lambda_r(F^q G) = \Lambda_r\big(\sum_{s=0}^{q-1} (F\Lambda_s(G))^q X^s\big) = \sum_{s=0}^{q-1} \Lambda_r\big((F\Lambda_s(G))^q X^s\big) = F\Lambda_r(G).$$

□

Note that the polynomial ring $\mathbb{F}_q[X]$ is contained in $\mathbb{F}_q[[X]]$. Hence, the field of fractions $\mathbb{F}_q(X)$ is contained in the field of fractions $\mathcal{Q}(\mathbb{F}_q[[X]]) = \mathbb{F}_q[[X]][\frac{1}{X}] =: \mathbb{F}_q((X))$. The latter consists of Laurent series $\sum_{n\geq n_0} a_n X^n$, with $n_0 \in \mathbb{Z}$. A power series $F \in \mathbb{F}_q[[X]]$ is *algebraic* over $\mathbb{F}_q(X)$ if there are polynomials $p_0, p_1, \ldots, p_d$ in $\mathbb{F}_q[X]$, not all zero, such that $\sum_{i=0}^{d} p_i F^i = 0$.

Lemma 6 (Ore)**.** *A formal power series $F \in \mathbb{F}_q[[X]]$ is algebraic over $F_q(X)$ if and only if there exists polynomials $A_0, \ldots, A_t$ in $\mathbb{F}_q[X]$, not all zero, such that*

$$A_0 F + A_1 F^q + A_2 F^{q^2} + \ldots + A_t F^{q^t} = 0.$$

Furthermore we can suppose that $A_0 \neq 0$.

Proof. We will follow the proofs as presented in [Fogg et al., 2002, Chapter 3] and [Allouche and Shallit, 2003, Chapter 12]. Since the sufficiency is clear, we only have to prove the necessity. Let $F = \sum_{n\geq 0} a_n X^n$ be an algebraic formal power series, then there exist a polynomial $P \in \mathbb{F}_q[X][T]$ such that $P(F) = 0$. Let $d = \deg P$, and perform Euclidean division of T^{q^i} by P in the ring $\mathbb{F}_q(X)[T]$ for $0 \leq i \leq d$: there are polynomials Q_i and R_i in $\mathbb{F}_q(X)[T]$ such that

$$T^{q^i} = Q_i P + R_i,$$

with $\deg_T(R_i) < d$. Since $R_0, \ldots, R_d$ are $d+1$ polynomials of degree at most $d-1$ in T, they are linearly dependent. So there are polynomials $A_0, \ldots, A_d \in \mathbb{F}_q[X]$ such that $A_0 R_0 + A_1 R_1 + \ldots + A_d R_d = 0$. Using $R_i = T^{q^i} - Q_i P$ we obtain

$$\sum_{i=0}^{d} A_i T^{q^i} = P \cdot \sum_{i=0}^{d} A_i Q_i.$$

Since F is a zero of P, it is also a zero of the left-hand side, so we find that $A_0 F + A_1 F^q + A_2 F^{q^2} + \ldots + A_t F^{q^t} = 0$.

To prove that there is such a relation with $A_0 \neq 0$, assume that we have $A_0 F + A_1 F^q + A_2 F^{q^2} + \ldots + A_t F^{q^t} = 0$, with t minimal. Let j be the smallest non-negative integer such that $A_j(X) \neq 0$ and assume that $j > 0$. Using property (a) of Lemma 5 we have

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

Since $A_j \neq 0$, it follows that there is an r for which $\Lambda_r(A_j) \neq 0$. For this r, using Λ_r on $\sum_{i=j}^{t} A_i F(X)^{q^i} = 0$ gives us

$$0 = \sum_{i=j}^{t} \Lambda_r(A_i F(X)^{q^i}) = \sum_{i=j}^{t} \Lambda_r(A_i) F(X)^{q^{i-1}},$$

where we use property (b) of Lemma 5 in the second equality. This gives us a new relation for $F, F^q, \ldots, F^{q^{t-1}}$, where the coefficient in front of $F^{q^{j-1}}$ is nonzero. This contradicts the minimality of j, hence $j = 0$.

□

Originally, the last part of this lemma is not stated in Ore's lemma. With Ore's lemma and the linear transformation Λ_r we are now ready to prove Christol's theorem.

Theorem 7 (Christol)**.** *Let $q = p^n$ for a prime p and a positive integer n and let $\mathbf{a} = (a_n)_{n\geq 0}$ be a sequence over $\mathbb{F}_q$. Then $\mathbf{a}$ is p-automatic if and only if the formal power series $\sum_{n\geq 0} a_n X^n$ is algebraic over $\mathbb{F}_q(X)$.*

Proof. We follow the proof as presented in [Allouche and Shallit, 2003, Chapter 12].

$\Rightarrow$: Since $\mathbf{a}$ is p-automatic, by Lemma 4 it is also q-automatic. By Lemma 3 we know that the q-kernel $K_q(\mathbf{a})$ is finite. Let $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(s)}$ be the s elements of $K_q(\mathbf{a})$, with $\mathbf{a}^{(1)} = \mathbf{a}$. Let $F_i = \sum a_n^{(i)} X^n$ be the formal power series corresponding to $\mathbf{a}^{(i)}$ for $1 \leq i \leq s$. Using property (b) of Lemma 5, we can rewrite each F_i as

$$F_i = \sum_{r=0}^{q-1} \Lambda_r(F_i)^q X^r.$$

The sequence corresponding to $\Lambda_r(F_i) = \sum_{n\geq 0} a_{qn+r}^{(i)} X^n$ is an element of the kernel $K_q(\mathbf{a})$ for all $0 \leq r \leq q-1$ and $1 \leq i \leq s$. Since $K_q(\mathbf{a})$ is finite, this implies that the power series F_i belongs to the vector space spanned by $F_1(X)^q, \dots, F_s(X)^q$ over $\mathbb{F}_q(X)$. Similarly, we find that

$$F_i^q = \sum_{n\geq 0} a_n^{(i)} X^{qn} = \sum_{r=0}^{q-1} \sum_{n\geq 0} a_{qn+r}^{(i)} X^{q(qn+r)} = \sum_{r=0}^{q-1} X^{qr} (\sum_{n\geq 0} a_{qn+r}^{(i)} X^n)^{q^2}.$$

So F_i^q, and hence F_i, belongs to the vector space spanned by $F_1(X)^{q^2}, \dots, F_s(X)^{q^2}$ over $\mathbb{F}_q(X)$, for all $1 \leq i \leq s$. By continuing this argument, and choosing $i = 1$, we find that $F_1, F_1^q, F_1^{q^2}, \dots, F_1^{q^s}$ belong to the vector space spanned by $F_1(X)^{q^{s+1}}, F_2(X)^{q^{s+1}}, \dots, F_s(X)^{q^{s+1}}$ over $\mathbb{F}_q(X)$. The dimension of this vector space is at most s, so the $s+1$ power series $F_1, F_1^q, F_1^{q^2}, \dots, F_1^{q^s}$ are linearly dependent, hence F_1 is algebraic.

$\Leftarrow$: The converse implication is a bit more involved. Let $F = \sum_{n\geq 0} a_n X^n$ be an algebraic power series, with corresponding sequence $\mathbf{a} = (a_n)_{n\geq 0}$. The idea of the proof is to make a finite set $\mathcal{H}$ that contains power series of a certain form, such that F is an element of $\mathcal{H}$, and that for all $0 \leq r \leq q-1$ we have $\Lambda_r(\mathcal{H}) \subset \mathcal{H}$. This implies that the power series corresponding to the elements of $K_q(\mathbf{a})$ are all contained in $\mathcal{H}$, so $K_q(\mathbf{a})$ is finite. Hence $\mathbf{a}$ is q-automatic, and by Lemma 4 it is also p-automatic. In the rest of the proof we will construct the set $\mathcal{H}$ and prove that $\mathcal{H}$ is stable under Λ_r.

Let F be of algebraic degree t over $\mathbb{F}_q(X)$. By Ore's Lemma there are polynomials $f_0, f_1, \dots, f_t \in \mathbb{F}_q[X]$, with $f_0 \neq 0$, such that

$$\sum_{i=0}^{t} f_i F^{q^i} = 0. \tag{2.2}$$

Define $G := F/f_0$, then equation (2.2) gives us

$$G = \sum_{i=1}^{t} g_i G^{q^i},$$

with $g_i = -f_i f_0^{q^i-2}$ for $1 \leq i \leq t$, and define

$$N = \max(\deg(f_0), \deg(g_1), \dots, \deg(g_k)).$$

Let $\mathcal{H}$ be the finite set of formal power series of the form

$$\sum_{i=0}^{t} h_i G^{q^i}, \tag{2.3}$$

where the h_i are polynomials in $\mathbb{F}_q[X]$ with $\deg(h_i) \leq N$. For any element $H = \sum_{i=0}^{t} h_i G^{q^i}$ in $\mathcal{H}$ and any $0 \leq r \leq q-1$, we have

$$\Lambda_r(H) = \Lambda_r\Big(h_0 G + \sum_{i=1}^{t} h_i G^{q^i}\Big) = \Lambda_r\Big(\sum_{i=1}^{t}(h_0 g_i + h_i)G^{q^i}\Big) = \sum_{i=1}^{t} \Lambda_r(h_0 g_i + h_i) G^{q^{i-1}},$$

where we used property (b) of Lemma 5 in the last equality. Since the degree of the polynomials $h_0 g_i + h_i$ is at most $2N$, we have for all $1 \leq i \leq t$ that

$$\deg(\Lambda_r(h_0 g_i + h_i)) \leq 2N/q \leq N.$$

Hence, $\Lambda_r(H)$ is an element of $\mathcal{H}$ for all $H \in \mathcal{H}$ and $0 \leq r < q$, so $\mathcal{H}$ is stable under Λ_r. Furthermore, $\mathcal{H}$ contains $F = f_0 G$, which completes the proof. □

If the sequence corresponding to a power series F is generated by a finite automaton M, then we say that M generates the power series F. With Christol's theorem, we can prove algebraic statements about power series with the use of automata. We give two corollaries of Christol's theorem, starting with Corollary 8 on Hadamard products. To prove this without automata theory is much more involved, see [Furstenberg, 1967]. Define for two formal power series $F = \sum_{n\geq 0} a_n X^n$ and $G = \sum_{n\geq 0} b_n X^n$ in $\mathbb{F}_q[[X]]$, the *Hadamard product* of F and G as $F \odot G = \sum_{n \geq 0} a_n b_n X^n$.

Corollary 8. *If two power series F and G are algebraic over $F_q(X)$, then so is their Hadamard product $F \odot G$.*

Proof. Since F and G are algebraic over $\mathbb{F}_q(X)$, by Christol's theorem the sequences $\mathbf{a} = (a_n)_{n\geq 0}$ and $\mathbf{b} = (b_n)_{n \geq 0}$ are both p-automatic. By Lemma 2, their pointwise product $(a_n b_n)_{n\geq 0}$ is also p-automatic. Hence, by Christol's theorem, the Hadamard product $F \odot G = \sum_{n\geq 0} a_n b_n X^n$ is algebraic over $\mathbb{F}_q(X)$. □

For a power series $F = \sum_{n \geq 0} a_n X^n$ over $\mathbb{F}_q$ and an element $\alpha \in \mathbb{F}_q^*$, define F_α as

$$F_\alpha = \sum_{\substack{n \geq 0 \\ a_n = \alpha}} X^n.$$

So the coefficients of F_α only take the values 0 and 1, and we have $F = \sum_{\alpha \in \mathbb{F}_q^*} \alpha F_\alpha$.

Corollary 9. *A power series F over $\mathbb{F}_q$ is algebraic if and only if the power series F_α is algebraic for each $\alpha \in \mathbb{F}_q^*$.*

Proof. $\Leftarrow$: Since F can be written as the pointwise sum $F = \sum_{\alpha \in \mathbb{F}_q^*} \alpha F_\alpha$ and all the F_α are algebraic, by Lemma 2 we know that F is also algebraic.

$\Rightarrow$: If $F = \sum_{n \geq 0} a_n X^n$ is algebraic, then there is a q-automaton $M = (Q, \Sigma_q, \delta, q_0, \mathbb{F}_q, \tau)$ that generates $\mathbf{a} = (a_n)_{n\geq 0}$. For every $\alpha \in \mathbb{F}_q^*$, define the new output function

$$\tau_\alpha(q) = \begin{cases} 1 & \text{if } \tau(q) = \alpha, \\ 0 & \text{otherwise.} \end{cases}$$

The q-automaton $\widetilde{M} = (Q, \Sigma_q, \delta, q_0, \mathbb{F}_q, \tau_\alpha)$ generates the sequence corresponding to F_α. Hence the power series F_α is algebraic over $\mathbb{F}_q(X)$ for each $\alpha \in \mathbb{F}_q^*$. □

2.4 Furstenberg's Theorem

Let G be an element of the ring $\mathbb{F}_q((X, Y))$ of Laurent series in two variables over $\mathbb{F}_q$, so

$$G(X, Y) = \sum_{\substack{m \geq m_0 \\ n \geq n_0}} g_{m,n} X^m Y^n,$$

with $g_{m,n} \in \mathbb{F}_q$ and $m_0, n_0 \in \mathbb{Z}$. The *diagonal* $\mathcal{D}(G)$ of G is the formal Laurent series in one variable defined by

$$\mathcal{D}(G) = \sum_{k \geq \max\{m_0, n_0\}} g_{k,k} X^k.$$

Diagonals and Hadamard products are linked in the following sense:

$$\left(\sum_{n \geq 0} a_n X^n\right) \odot \left(\sum_{n \geq 0} b_n X^n\right) = \sum_{n \geq 0} a_n b_n X^n = \mathcal{D}\left(\left(\sum_{n \geq 0} a_n X^n\right)\left(\sum_{n \geq 0} b_n Y^n\right)\right),$$

which follows straightforwardly from the definitions of diagonals and Hadamard products. Using Corollary 8 on this relation gives us that if F and G are two algebraic power series, then the diagonal of the algebraic power series $F \cdot G$ in two variables algebraic too. The following theorem by Furstenberg shows that in general the diagonal of a two-dimensional rational Laurent series is algebraic.

Theorem 10 (Furstenberg). *A formal Laurent series $F = \sum_{n \geq n_0} a_n X^n$ over a finite field $\mathbb{F}_q$ is algebraic if and only if it is the diagonal of a rational Laurent series in two variables, i.e. $F = \mathcal{D}(G)$ for an element $G = \sum_{m,n \geq 0} g_{m,n} X^m Y^m$ of $\mathbb{F}_q(X, Y) \subset \mathbb{F}_q((X, Y))$.*

Proof. We will sketch the idea of the proof here. For a complete proof, see [Allouche and Shallit, 2003, Chapter 12, 14] or [Furstenberg, 1967].

$\Leftarrow$: Christol's theorem can be generalized to the multidimensional case, where we use multivariate power series and multidimensional arrays. In the two-dimensional case, Christol's theorem states: a formal power series $G = \sum_{m,n \geq 0} g_{m,n} X^m Y^n$ is algebraic over $\mathbb{F}_q(X, Y)$, with $q = p^k$, if and only if the corresponding double sequence $\mathbf{g} = (g_{m,n})_{m,n \geq 0}$ is p-automatic. A double sequence can be seen as an infinite matrix, and it is p-automatic if there exist a finite p-automaton $M = (Q, \Sigma, \delta, q_0, \Delta, \tau)$, with $\Sigma = \Sigma_p \times \Sigma_p$, that generates $\mathbf{g}$. This automaton M takes as input a string of pairs of symbols, so for example $([w_k, v_k], \ldots, [w_1, v_1], [w_0, v_0])$, and reads it pair by pair. It produces the output $g_{m,n}$ for $m = [w_k \cdots w_0]_p$ and $n = [v_k \cdots v_0]_p$. The concept of kernel can also be defined for the two-dimensional case. The p-kernel of the double sequence $\mathbf{g} = (g_{m,n})_{m,n \geq 0}$ is a set of infinite submatrices:

$$K_p(\mathbf{g}) = \{(g_{p^i m + j, p^i n + l})_{m,n \geq 0} : i \geq 0, 0 \leq j < p^i, 0 \leq l < p^i\}.$$

Like in the one-dimensional case, the p-kernel of a double sequence is finite if and only if the double sequence is p-automatic.

We now have generalizations to the multidimensional case of the previous lemmas and theorems, so we can start with the actual proof. If G is a rational Laurent series, the corresponding double sequence $\mathbf{g} = (g_{m,n})_{m,n\geq 0}$ is p-automatic, and hence $K_p(\mathbf{g})$ is finite. Let F be the diagonal of G, then we have $a_n = g_{n,n}$, so the kernel of $\mathbf{a}$ can be written as

$$K_p(\mathbf{a}) = \{(a_{p^i n+j})_{n\geq 0} : i \geq 0, 0 \leq j < p^i\} = \{(g_{p^i n+j, p^i n+j})_{n\geq 0} : i \geq 0, 0 \leq j < p^i\}.$$

So the infinite sequences in $K_p(\mathbf{a})$ are diagonals of the infinite matrices in $K_k(\mathbf{g})$ with $j = l$. Hence, $K_p(\mathbf{a})$ is finite too, so $\mathbf{a}$ is p-automatic and F is algebraic.

$\Rightarrow$: If F is an algebraic power series, then Ore's lemma gives us that there are polynomials $B_j(X)$ not all equal to zero, such that

$$B_0 F + B_1 F^q + \cdots + B_t F^{q^t}. \tag{2.4}$$

The idea of this part of the proof is to construct from this equation a rational power series G in two variables such that $F = \mathcal{D}(G)$. The approach to find this G, as presented in [Furstenberg, 1967] and [Allouche and Shallit, 2003, Chapter 12], is rather long and not so transparent, so we leave out the rest of the proof.

□

2.5 Examples

Furstenberg's theorem can be used to construct an algebraic power series in one variable, by taking the diagonal of a rational Laurent series in two variables. In this section we discuss two similar examples, in which we consider the diagonal of a Laurent series. In the first example we find a quite trivial diagonal, but in the second example we find a more interesting diagonal and construct a generating automaton, find a algebraic relation and compute the kernel.

2.5.1 Diagonal of $(1 + X + Y)^{-1}$

Consider the following power series in $\mathbb{F}_2(X, Y)$:

$$F_1(X, Y) = (1 + X + Y)^{-1} = \sum_{n\geq 0} (X + Y)^n = \sum_{n\geq 0} \sum_{m=0}^{n} \binom{n}{m} X^m Y^{n-m}.$$

To find its diagonal, we need the coefficients in front of the monomials $X^m Y^{n-m}$ for which $m = n - m$, so $n = 2m$. Hence,

$$\mathcal{D}(F_1) = \mathcal{D}\left(\sum_{n\geq 0} \sum_{m=0}^{n} \binom{n}{m} X^m Y^{n-m}\right) = \sum_{k\geq 0} \binom{2k}{k} X^k \in \mathbb{F}_2((X)).$$

We will see that $\mathcal{D}(F_1) = 1$, by using a theorem of Legendre [Legendre, 1808]. Let s_k denote the number of ones in the binary representation of a non-negative number k. Legendre's theorem states that for i such that $2^i | k!$ but $2^{i+1} \nmid k!$, it holds that $i = k - s_k$. So let j be such that $2^j \mid \binom{2k}{k} = \frac{(2k)!}{k!k!}$ and $2^{j+1} \nmid \binom{2k}{k}$. Then we have $j = 2k - s_{2k} - (k - s_k + k - s_k) = -s_{2k} + 2s_k = s_k$. This implies that $j \geq 1$ for $k > 0$, so $\binom{2k}{k} \equiv 0 \bmod 2$ for $k > 0$, hence $\mathcal{D}(F_1) = 1$. This power series has not an interesting kernel or automaton, so we move on to the next example.

2.5.2 Diagonal of $(1+X+Y^2)^{-1}$

We now consider $F_2(X,Y) = (1+X+Y^2)^{-1}$. Rewrite F_2 as

$$F_2(X,Y) = \sum_{n\geq 0}(X+Y^2)^n = \sum_{n\geq 0}\sum_{m=0}^{n}\binom{n}{m}X^{n-m}Y^{2m}.$$

To compute $\mathcal{D}(F_2)$, we need the coefficients of F_2 in front of the monomials $X^{n-m}Y^{2m}$ with $n-m=2m$, so $n=3m$. We find

$$\mathcal{D}(F_2) = \sum_{k\geq 0}\binom{3k}{k}X^{2k} = 1+X^2+X^4+X^8+X^{10}+X^{16}+X^{18}+X^{20}+X^{32}+\dots. \tag{2.5}$$

We use Legendre's theorem again to see when $\binom{3k}{k} \equiv 1 \bmod 2$. For $k \geq 0$, write $\binom{3k}{k} = \frac{(3k)!}{k!(2k)!}$, and let j_k be such that $2^{j_k} \mid \binom{3k}{k}$ but $2^{j_k+1} \nmid \binom{3k}{k}$, then

$$j_k = 3k - s_{3k} - (k-s_k) - (2k - s_{2k}) = s_k + s_{2k} - s_{3k}.$$

So $j_k = 0$, hence $\binom{3k}{k} \equiv 1 \bmod 2$, if and only if $s_k + s_{2k} = s_{3k}$. This last equation means that there are no carries when we add the binary representations of k and $2k$. So k and $2k$ can not have a 1 on the same place in their binary representations, hence $(k)_2$ can not have any consecutive ones. If this holds for k, it also holds for $2k$, since $(2k)_2 = (k)_2 0$, so $\binom{3k}{k} \equiv \binom{6k}{2k} \bmod 2$.

Let $\mathbf{a} = (a_k)_{k\geq 0}$ be the sequence corresponding to $F = \mathcal{D}(F_2) = \sum_{k\geq 0} a_k X^k$, with

$$a_k = \begin{cases} \binom{3(k/2)}{k/2} = \binom{3k}{k} & \text{for even } k, \\ 0 & \text{for odd } k. \end{cases}$$

We have $a_k = 1$ if and only if k is even and has no consecutive ones in its binary representation. With this description, we can find a 2-automata for $(a_k)_{k\geq 0}$, see Figure 2.3. Once the automaton gets in state q_3 it can never get out, so it will produce output 0. This exactly happens when $(k)_2$ ends with a 1 or $(k)_2$ contains consecutive ones. If the automaton is in any of the other states, the output is 1. This automaton is leading zeros invariant and generates $\mathbf{a}$.

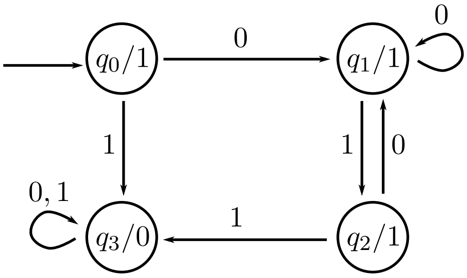


Figure 2.3: A 2-automaton with four states that generates the sequence corresponding to $F = \mathcal{D}((1+X+Y^2)^{-1})$.

Since $F = \mathcal{D}(F_2)$ is algebraic over $\mathbb{F}_2(X)$, there is a polynomial P with coefficients in $\mathbb{F}_2(X)$

such that $P(F) = 0$. To find this polynomial P, we start by computing F^2:

$$\begin{aligned} F^2 &= \Big(\sum_{k\geq 0}\binom{3k}{k}X^{2k}\Big)^2 = \sum_{k\geq 0}\binom{3k}{k}X^{4k} = \sum_{k\geq 0}\binom{6k}{2k}X^{4k} \\ &= \sum_{\substack{k\geq 0\\ k\equiv 0 \bmod 2}}\binom{3k}{k}X^{2k}. \end{aligned}$$

If $k \equiv 3 \bmod 4$ then the binary representation of k ends with two ones, so $\binom{3k}{k} = 0$ for these k. Using this, we find

$$\begin{aligned} F + F^2 &= \sum_{k\geq 0}\binom{3k}{k}X^{2k} + \sum_{\substack{k\geq 0\\ k\equiv 0 \bmod 2}}\binom{3k}{k}X^{2k} \\ &= \sum_{\substack{k\geq 0\\ k\equiv 1 \bmod 2}}\binom{3k}{k}X^{2k} = \sum_{\substack{k\geq 0\\ k\equiv 1 \bmod 4}}\binom{3k}{k}X^{2k}. \end{aligned}$$

With similar calculations as for F^2 we compute X^2F^4:

$$\begin{aligned} X^2F^4 &= X^2 \sum_{\substack{k\geq 0\\ k\equiv 0 \bmod 4}}\binom{3k}{k}X^{2k} = \sum_{\substack{k\geq 0\\ k\equiv 0 \bmod 4}}\binom{3k}{k}X^{2(k+1)} \\ &= \sum_{\substack{k\geq 0\\ k\equiv 1 \bmod 4}}\binom{3(k-1)}{k-1}X^{2k}. \end{aligned}$$

For $k \equiv 1 \bmod 4$, the binary representations of k and $k-1$ end with 01 and 00 respectively, the rest of digits are the same. So for $k \equiv 1 \bmod 4$, k has consecutive ones if and only if $k-1$ has, so $\binom{3k}{k}$ and $\binom{3(k-1)}{k-1}$ have the same value in $\mathbb{F}_2$. So we see that F satisfies

$$F + F^2 + X^2F^4 = \sum_{\substack{k\geq 0\\ k\equiv 1 \bmod 4}}\left(\binom{3k}{k} + \binom{3(k-1)}{k-1}\right)X^{2k} = 0.$$

Hence, F is a zero of the irreducible polynomial $P(T) = X^2T^3 + T + 1 = 0 \in \mathbb{F}_2(X)[T]$.

Since $\mathbf{a}$ is 2-automatic, the kernel must be finite. Let $\mathbf{a}^{(1)} = \mathbf{a}$ be the first element of $K_2(\mathbf{a})$, and split $\mathbf{a}^{(1)}$ in two subsequences, $\mathbf{a}^{(2)} = (a_{2k})_{k\geq 0}$ and $\mathbf{a}^{(3)} = (a_{2k+1})_{k\geq 0}$, which are both elements of $K_2(\mathbf{a})$. We now have three elements in the 2-kernel:

$$\begin{aligned} \mathbf{a}^{(1)} &= \mathbf{a} = (1,0,1,0,1,0,0,0,1,0,1,0,0,0,0,0,1,0,1,0,1,\ldots), \\ \mathbf{a}^{(2)} &= (a_{2k})_{k\geq 0} = (1,1,1,0,1,1,0,0,1,1,1,\ldots), \\ \mathbf{a}^{(3)} &= (a_{2k+1})_{k\geq 0} = (0,0,0,0,0,0,0,0,0,0,\ldots). \end{aligned}$$

If we keep repeating the splitting of the sequences, we obtain all elements of $K_2(\mathbf{a})$. We know that $\mathbf{a}^{(1)}$ splits into $\mathbf{a}^{(2)}$ and $\mathbf{a}^{(3)}$, and that the zero sequence $\mathbf{a}^{(3)}$ splits in two copies of itself. So we only need to see what happens if we split $\mathbf{a}^{(2)} = (a_{2k})_{k\geq 0} = (\binom{6k}{2k})_{k\geq 0} = (\binom{3k}{k})_{k\geq 0}$. Let $\mathbf{b}$ and $\mathbf{c}$ be the ‘even’ and ‘odd’ subsequences of $\mathbf{a}^{(2)}$:

$$\begin{aligned}\mathbf{b} &= (\binom{3(2k)}{2k})_{k\geq 0} = (\binom{3k}{k})_{k\geq 0} = \mathbf{a}^{(2)},\\ \mathbf{c} &= (\binom{3(2k+1)}{2k+1})_{k\geq 0}.\end{aligned}$$

If k is odd, then $(2k+1)_2$ ends with two 1's, so $\binom{3(2k+1)}{2k+1} \equiv 0 \bmod 2$. If k is even, $\binom{3(2k+1)}{2k+1} \equiv \binom{3k}{k} \bmod 2$. So we see that $\mathbf{c} = \mathbf{a}$. Hence $K_2(\mathbf{a})$ consists of the three elements $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}$ and $\mathbf{a}^{(3)}$.

In the proof of Lemma 3 we construct an automaton using the elements of the kernel. This shows that there is a 2-automaton with just three states that generates $\mathbf{a}$. The 2-automaton that we create with this procedure is the automaton in Figure 2.4. We named the state on the left q_3, because this automaton can also be obtained by merging the states q_0 and q_2 of the automaton in Figure 2.3.

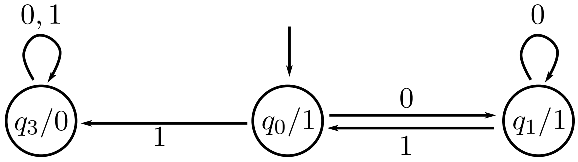


Figure 2.4: A 2-automaton with only three states that generates the sequence corresponding to $F = \mathcal{D}(F_2)$.

3
Effectivity and bounds

Consider a q-automaton over $\mathbb{F}_q$ with m states. It describes a formal power series F over $\mathbb{F}_q$, which is algebraic according to Christol's theorem. What can we say about the algebraic degree of F? The other way around, consider a formal power series in $\mathbb{F}_q[[X]]$ that satisfies a polynomial $P = \alpha_0 + \alpha_1 T + \ldots + \alpha_d T^d$ over $\mathbb{F}_q[X]$. What can we say about the size of a corresponding q-automaton? By closely examining the steps in the proofs of Christol's theorem and Ore's lemma, we can answer these questions. The results are summarized in Theorem 12 in Section 3.1 and Theorem 14 in Section 3.2. Both theorems are followed by special cases and remarks.

For both theorems that follow, we need the following lemma, which is a direct consequence of the proof of Lemma 3.

Lemma 11. *If an infinite sequence $\mathbf{a} = (a_n)_{n\geq 0}$ is k-automatic, then there is a (reverse reading) k-automaton M with $|K_k(\mathbf{a})|$ states, that generates $\mathbf{a}$. Furthermore, there is no k-automaton that generates $\mathbf{a}$ with less than $|K_k(\mathbf{a})|$ states.*

Proof. In the second part of the proof of Lemma 3, we create a k-automaton M that generates $\mathbf{a}$, with exactly $|K_k(\mathbf{a})|$ states. Suppose there is a k-automaton $\widetilde{M}$ that generates $\mathbf{a}$ with t states, such that $t < |K_k(\mathbf{a})|$. From the first part of the proof of Lemma 3, we see that $|K_k(\mathbf{a})|$ is bounded by the number of states of $\widetilde{M}$, so $|K_k(\mathbf{a})| \leq t$. This leads to a contradiction, so there is no k-automaton that generates $\mathbf{a}$ with less than $|K_k(\mathbf{a})|$ states. □

For a given power series or sequence, we say that an automaton is *minimal* if the number of states equals the size of the corresponding kernel.

3.1 From a q-automaton to an algebraic power series

Theorem 12. *Let M be a leading zeros invariant q-automaton over $\mathbb{F}_q$ with m states, where q is a prime power. Let $F = \sum a_n X^n \in \mathbb{F}_q[[X]]$ be the corresponding formal power series. Then the algebraic degree of F is at most $q^m - 1$.*

Proof. Let s denote the number of elements in the kernel of $\mathbf{a} = (a_n)_{n\geq 0}$, by Lemma 11 we have $s \leq m$. In the first part of the proof of Christol's theorem, we find that $F, F^q, F^{q^2}, \ldots, F^{q^s}$ are linearly dependent. Assuming that F is nonzero, we find that the degree of algebraicity of F is at most $q^s - 1$, which is in turn bounded by $q^m - 1$. □

This exponential bound implies that a power series corresponding to a relatively simple finite automaton with 4 states over $\mathbb{F}_2$, has an algebraic degree that we can only bound by $2^4 - 1 = 15$. However, we know from the first part of the proof of Christol's theorem that there is a relationship between just 5 monomials, namely F, F^2, F^4, F^8 and F^{16}. In some special cases this can help us find a polynomial P such that $P(F) = 0$, but in general such a polynomial P can be hard to find. If the automaton is not minimal and if we can compute the size s of the kernel, then we can bound the algebraic degree of the corresponding power series by $q^s - 1$.

To investigate the tightness of the bound of Theorem 12, we explore the infinite class of power series F_m that satisfy $F_m^{2^m-1} + (X+1) = 0$ for $m \geq 1$. The result is summarized in the next proposition.

Proposition 13. *For $m \geq 1$, let F_m be the power series solution of $F_m^{2^m-1} + (X+1) = 0$ with $F(0) = 1$. There exists a reverse reading 2-automaton over $\mathbb{F}_2$ that generates F_m with $m+2$ states. Moreover, there are no 2-automata that generate F_m with less than $m+2$ states.*

Proof. We write $F_m = \sum_{n\geq 0} a_n X^n$ and use the equation $F_m^{2^m} + XF_m + F_m = 0$ to get

$$\sum_{n\geq 0} a_n X^{2^m n} + \sum_{n\geq 0} a_n X^{n+1} + \sum_{n\geq 0} a_n X^n = 0.$$

So we see that $a_0 = 1$ and for $n \geq 1$ we have

$$a_n = \begin{cases} a_{n-1} & \text{for } n \not\equiv 0 \bmod 2^m, \\ a_{n-1} + a_{n/2^m} & \text{for } n \equiv 0 \bmod 2^m. \end{cases} \tag{3.1}$$

This recurrence relation defines the unique power series with $F_m(0) = 1$ that satisfies $P(F_m) = 0$. Before we present an explicit expression for the power series F_m, note that the power series $(1+X)^{-1} = \sum_{n\geq 0} X^n$ can be written as $\prod_{j\geq 0}(1 + X^{2^j})$, since every n can uniquely be written as a sum of different powers of 2. If we remove the terms where j is a positive multiple of m, then we get the power series

$$F_m = \frac{\prod_{j\geq 0}(1 + X^{2^j})}{\prod_{k\geq 1}(1 + X^{2^{km}})} = \prod_{k\geq 0}(1 + X^{2^{km}})^{-1}. \tag{3.2}$$

Clearly $F(0) = 1$ and $P(F_m) = 0$ since

$$F_m^{2^m} = \prod_{k\geq 0}\left((1 + X^{2^{km}})^{-1}\right)^{2^m} = \prod_{k\geq 1}(1 + X^{2^{km}})^{-1} = (1+X)F_m.$$

Let $(n)_2 = c_t \dots c_1 c_0$ be the binary representation of an integer n, then (3.2) gives us that $a_n = 1$ if and only if $c_{jm} = 0$ for all $j \geq 1$. A 2-automaton $M = (Q, \Sigma_2, \delta, q_0, \Sigma_2, \tau)$ for F_m can be constructed easily, see Figure 3.1 for the case that $m = 6$. Let $Q = \{q_0, q_1, \dots, q_m, q_{m+1}\}$, where q_0 is the initial state and q_{m+1} the sink state. We define the transition function as

$$\begin{aligned} \delta(q_i, a) &= q_{i+1} && \text{for } 0 \leq i < m \text{ and } a \in \{0,1\}, \\ \delta(q_m, 0) &= q_1, \\ \delta(q_m, 1) &= q_{m+1}, \\ \delta(q_{m+1}, a) &= q_{m+1} && \text{for } a \in \{0,1\}. \end{aligned}$$

The output for the sink state q_{m+1} is 0, and $\tau(q_i) = 1$ for $1 \leq i \leq m$. We see that if for some

$j > 0$ the digit c_{jm} of $(n)_2$ equals 1, then the automaton produces output 0. Otherwise, the automaton produces output 1. So this automaton generates F_m.

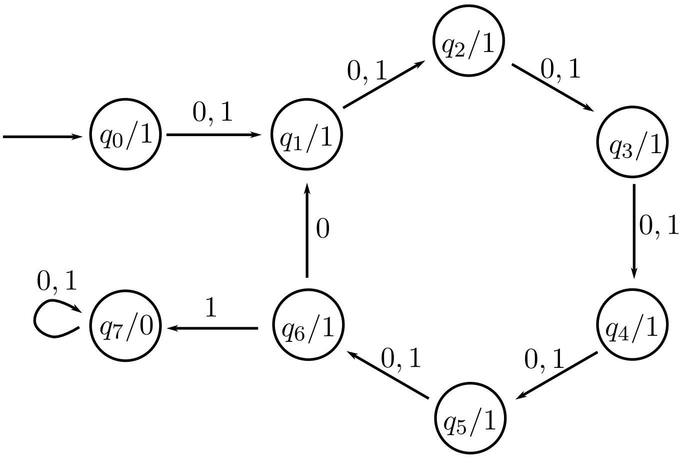


Figure 3.1: This 2-automaton generates the power series F_m of Proposition 13 for $m = 6$.

To see that F_m cannot be generated by an automaton with less states, we have to look at the kernel of the corresponding sequence $\mathbf{a} = (a_n)_{n\geq 0}$. Using 3.1 we see that

$$a_0 = a_1 = \ldots = a_{2^m-1} = 1 \quad \text{and} \quad a_{2^m} = a_{2^m+1} = \ldots = a_{2^{m+1}-1} = 0.$$

So the sequence $\mathbf{a}$ starts with 2^m ones, followed by 2^m zeros. Hence, for $0 \leq i \leq m$, the subsequence $(a_{2^i n})_{n\geq 0}$ starts with 2^{m-i} ones, followed by the same number of zeros. Furthermore, the subsequence $(a_{2^{m+1}n+2^m})_{n\geq 0}$ starts with a zero. So we found $m+2$ subsequences that begin differently, hence the kernel of $\mathbf{a}$ contains at least $m + 2$ elements. Using Lemma 11, this completes the proof.

□

So we found an infinite set of (minimal) automata of increasing size m that generate power series of algebraic degree $q^{m-2} - 1$. So this confirms that the bound of Theorem 12 should be exponential. However, we have not yet found an example that makes this bound tight.

We will now consider the very specific example that $q = 2$ and $m = 2$, so we look at all leading zeros invariant 2-automata over $\mathbb{F}_2$ with only two states. We will find the algebraic degrees of the corresponding power series and compare them to the bound of Theorem 12.

There are $2^6 = 62$ different 2-automata with 2 states over $\mathbb{F}_2$, but not all of them are leading zeros invariant. An automaton M has this property if and only if every edge labeled with a 0 connects two states with the same output, so $\tau(\delta(q_i, 0)) = \tau(q_i)$ for all states q_i. If both states of M have the same label, then the corresponding power series is 0 or $(1 + X)^{-1}$. So, we continue with the case that M has two states with different output labels and let F be the corresponding power series. Then $F + (1 + X)^{-1}$ is generated by an automaton similar to M, but with swapped output labels. Once we have a minimal polynomial of F, it is trivial to find a minimal polynomial for $F + (1+X)^{-1}$, which has the same degree. So we assume that the two states of M, q_0 and q_1, have output labels $\tau(q_0) = 0$ and $\tau(q_1) = 1$. Since these output labels are different, $\delta(q_i, 0) = q_i$ has to hold for $i = 0, 1$ to make sure that the automaton is leading zeros invariant. There are 4 possibilities to choose $\delta(q_0, 1)$ and $\delta(q_1, 1)$. When $\delta(q_0, 1) = q_0$, then the automaton generates $F = 0$. When $\delta(q_0, 1) = q_1$, then there are two possible automata left, see Figure 3.2.

The automaton in Figure 3.2a generates the power series $F_1 = \sum_{n\geq 1} X^n = X(1 + X)^{-1}$,

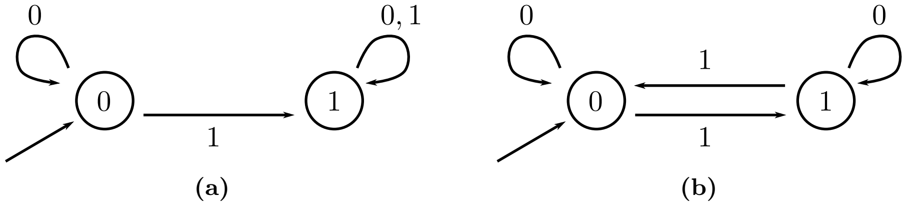


Figure 3.2: Two non-trivial, leading zeros invariant 2-automata with 2 states.

which satisfies $(X+1)F_1 - X = 0$. The complementary power series $F_2 = F_1 + (1+X)^{-1} = 1$ satisfies $F_2 - 1 = 0$. The automaton in Figure 3.2b is the same as in Figure 2.1, and generates the Thue-Morse sequence $\mathbf{b} = (0,1,1,0,1,0,0,1,1,0,0,1,0,1,1,0,\ldots)$. Let $F_3 = \sum_{n\geq 0} s_n X^n$ be the corresponding power series, where s_n counts the number of ones in $(n)_2$. We want to find the minimal polynomial of F_3, so we start by computing F_3^2:

$$F_3^2 = \sum_{n\geq 0} s_n X^{2n} = \sum_{n\geq 0} s_{2n} X^{2n}$$

where we used $s_n = s_{2n}$. Since $s_{2n} + 1 = s_{2n+1}$ we have

$$XF_3^2 = X\sum_{n\geq 0} s_{2n}X^{2n} = \sum_{n\geq 0} s_{2n}X^{2n+1} = \sum_{n\geq 0} s_{2n+1}X^{2n+1} + \sum_{n\geq 0} X^{2n+1}.$$

Hence

$$(1+X)F_3^2 = \sum_{n\geq 0} s_{2n}X^{2n} + \sum_{n\geq 0} s_{2n+1}X^{2n+1} + \sum_{n\geq 0} X^{2n+1} = F_3 + X(1+X^2)^{-1}.$$

So we find $(1+X)F_3^2 + F_3 + X(1+X^2)^{-1} = 0$, or equivalently $(1+X)^3F_3^2 + (1+X^2)F_3 + X = 0$. It is easy to show that the corresponding complementary power series $F_4 = F_3 + (1+X)^{-1} = \sum_{n\geq 0}(s_n+1)X^n$ satisfies the same polynomial.

We have now found the six different power series over $\mathbb{F}_2$ that are generated by a 2-automaton over $\mathbb{F}_2$ with 2 states. They all have an algebraic degree of at most 2. The bound of Theorem 12, which is 3 in this case, is in this case not tight.

3.2 From an algebraic power series to a q-automaton

Given a power series F of algebraic degree d, can we give an upper bound for the number of states of a minimal automaton that generates it? No, since the algebraic degree d is not enough to give such a bound. To see this, consider a formal power series F of algebraic degree 1, so $F - r(X) = 0$ for a rational function $r(X) \in \mathbb{F}_q(X)$. There are infinitely many choices for this $r(X)$, which are all generated by different automata, but there are only finitely many q-automata of a given size. So the algebraic degree d of F is not enough to find an upper bound for the size of a minimal automaton. Let $P(T) = \alpha_0 + \alpha_1 T + \ldots + \alpha_d T^d$ be a polynomial such that $P(F) = 0$, then we can give such a bound, if we know the maximum degree of the coefficients of P. The result is stated in the following theorem.

Theorem 14. *Let $F = \sum_{n\geq 0} a_n X^n$ be an algebraic power series that is a zero of $P(T) = \alpha_0 + \alpha_1 T + \ldots + \alpha_d T^d$, with $\alpha_i \in \mathbb{F}_q[X]$ and $\alpha_d \neq 0$. There exists a q-automaton that generates*

F with at most

$$C := q^{(d+1)\left(\left(q^d-1\right)\cdot A\cdot\left(\frac{q(q^d-1)}{q-1}-d^2+d\right)+1\right)}$$

states, where $A = \max_{0\leq i\leq d}(\deg \alpha_i)$.

Proof. We will closely follow the steps in the proofs of Christol's theorem and Ore's lemma, as presented in Section 2.3, to see how we can bound the number of states of a minimal generating automaton.

We start with revisiting the proof of Christol's theorem. The set $\mathcal{H}$ contains elements of the form $\sum_{i=0}^{d} h_i G^{q^i}$, with $h_i \in \mathbb{F}_q[X]$ and $\deg(h_i) \leq N$, for $0 \leq i \leq d$. So $\mathcal{H}$ has at most $q^{(d+1)(N+1)}$ elements. The q-kernel of $\mathbf{a} = (a_n)_{n\geq 0}$ corresponds to a subset of $\mathcal{H}$, so we have that $|K_q(\mathbf{a})| \leq q^{(d+1)(N+1)}$. Using Lemma 11, we find that there exists a q-automaton that generates $\mathbf{a}$ with at most $q^{(d+1)(N+1)}$ states. In the rest of this proof we will find an upper bound for N that depends on q, d and the maximum degree A of the coefficients α_i of P.

We now turn to the proof of Ore's lemma. Recall that the polynomials $R_i \in \mathbb{F}_q(X)[T]$ are the result of the Euclidean division $T^{q^i} = Q_iP + R_i$ for $0 \leq i \leq d$, with $Q_i \in \mathbb{F}_q(X)[T]$ and $\deg_T R_i < d$. Let $R_{i,j} \in \mathbb{F}_q(X)$ be the coefficient in front of T^j in R_i, for $0 \leq i \leq d$ and $0 \leq j < d$:

$$R_i = R_{i,0} + R_{i,1}T + R_{i,2}T^2 + \ldots + R_{i,d-1}T^{d-1} \qquad 0 \leq i \leq d.$$

We want to express each $R_{i,j}$ as a rational function in the α_i's. Since we have $T^{q^i} \equiv R_i \bmod P$ in $\mathbb{F}_q(X)[T]/(P)$, we can use the relation

$$T^d \equiv -\frac{\alpha_0}{\alpha_d} - \frac{\alpha_1}{\alpha_d}T - \ldots - \frac{\alpha_{d-1}}{\alpha_d}T^{d-1} \bmod P.$$

$q^i - d + 1$ times on T^{q^i} to find that each $R_{i,j}$ can be written as a homogeneous polynomial of degree $q^i - d + 1$ in the polynomials $\alpha_0, \ldots, \alpha_d$, divided by $\alpha_d^{q^i-d+1}$.

The next step in Ore's lemma is finding $f_0, f_1, \ldots, f_d \in \mathbb{F}_q[X]$ such that $f_0R_0 + \ldots + f_dR_d = 0$. We want to express $f_0, f_1, \ldots, f_d$ as a rational function in the $R_{i,j}$'s, and hence as a rational function in the α_i's. Let $f = (f_0, f_1, \ldots, f_d)$, and let $R = (R_{i,j})_{i,j}$ be the $(d+1) \times d$ matrix that has the coefficients of R_i on row i for $0 \leq i \leq d$. Note that both indices i and j start counting from 0. We can assume R has rank d. If this is not the case, we can leave some rows of R out and replace them by new rows with coefficients in $\mathbb{F}_q$, to obtain a matrix S of rank d and with the same height as R, such that if f is a solution of $fS = 0$, then we also have $fR = 0$.

Let M_i be the $d \times d$ submatrix of R that is obtained by removing the ith row of R. Since we assume that the rank of R is d, at least one of these submatrices M_i has rank d. Let $|M_i|$ denote the determinant of the submatrix M_i, then the non-zero row vector

$$\widetilde{f} = (|M_0|, -|M_1|, |M_2|, \ldots, (-1)^d|M_d|)$$

satisfies $\widetilde{f}R = 0$, since for $0 \leq j \leq d-1$

$$(\widetilde{f}R)_j = fR_{\bullet,j} = R_{0,j}|M_0| - R_{1,j}|M_1| + R_{2,j}|M_2| + \ldots + (-1)^d R_{d,j}|M_d| = \det([R_{\bullet,j}|R]) = 0,$$

where $R_{\bullet,j}$ denotes the jth column of R.

The maximum degree of the numerators in the ith row of the matrix R is at most A^{q^i-d+1}, since each $R_{i,j}$ can be written as a homogeneous polynomial of degree $q^i - d + 1$ in $\alpha_0, \ldots, \alpha_d$, divided by $\alpha_d^{q^i-d+1}$. Since the determinant $|M_i|$ consists of products of d polynomials $R_{i,j}$ that are all pairwise in a different row, each $\widetilde{f}_i = (-1)^i|M_i|$ can be written as a homogeneous polynomial of degree $(q^1 - d + 1) + (q^2 - d + 1) + \ldots + (q^d - d + 1)$ in $\alpha_0, \ldots, \alpha_d$, divided by $\alpha_d^{(q^1-d+1)+(q^2-d+1)+\ldots+(q^d-d+1)}$. We remove these denominators of $\widetilde{f}$ to obtain a polynomial solution f of $fR = 0$, with

$$\begin{aligned} \deg f_i &\leq A \cdot ((q^1 - d + 1) + (q^2 - d + 1) + \ldots + (q^d - d + 1)) \\ &= A \cdot \left(\frac{q(q^d - 1)}{q - 1} - d^2 + d \right) =: B. \end{aligned}$$

for $0 \leq i \leq d$.

Returning to Christol's theorem, we see that $N := \max\{\deg(f_0), \deg(g_1), \ldots, \deg(g_d)\}$, where $g_i = -f_i f_0^{q^i-2}$. We now are in a position to bound N:

$$\begin{aligned} N &= \max\{\deg(f_0), \deg(f_1) + (q^1 - 2)\deg(f_0)), \ldots, \deg(f_d) + (q^d - 2)\deg(f_0)\} \\ &\leq B + (q^d - 2)B \\ &= (q^d - 1) \cdot A \cdot \left(\frac{q(q^d - 1)}{q - 1} - d^2 + d \right). \end{aligned}$$

Using this bound for N in the bound $q^{(d+1)(N+1)}$ that we found in the beginning of this proof, we find that the number of states of a minimal generating automaton is bounded by

$$C = q^{(d+1)\left((q^d-1)\cdot A \cdot \left(\frac{q(q^d-1)}{q-1} - d^2 + d\right) + 1\right)}.$$

□

Since $q/(q-1) \leq 2$ for $q \geq 2$, and $-d^2 + d \leq 0$ for $d \geq 1$, the bound C in Theorem 14 can be bounded by the slightly simpler expression

$$q^{(d+1)(2A(q^d-1)^2+1)}.$$

Note that it does not matter if the polynomial P in Theorem 14 is irreducible or not, since the proof works for all polynomials P with $P(F) = 0$. In some cases, this gives us the opportunity to use a different polynomial P, with different degree d and height A, such that we can find a smaller bound C.

The bound of Theorem 14 seems rather loose, and we have not found examples that make this bound tight. In the rest of this section we will investigate the case that F is a rational power series that satisfies a polynomial with linear coefficients, so $d = A = 1$. In this case the bound of Theorem 14 is $q^{2(q-1)q+2}$.

We will first consider the case where $q = 2$. Then F is a rational power series over $\mathbb{F}_2$ that satisfies $s(X)F - r(X) = 0$ for certain linear polynomials $r, s \in \mathbb{F}_2[X]$, with $s(0) \neq 0$. Since $s(X) \in \{1, 1 + X\}$, and $r(X) \in \{0, 1, X, 1 + X\}$, we only have the following possibilities for F:

$$0, 1, X, 1 + X, \frac{1}{1 + X}, \frac{X}{1 + X}.$$

The kernels of the corresponding sequences are respectively of size $1, 2, 3, 3, 1$ and 2. Hence, by Lemma 11 there are 2-automata of these sizes that generate these sequences. The corresponding automata are easy to find, see Figure 3.3. So in the specific case that $A = d = 1$ and $q = 2$, a power series F can always be generated by a 2-automaton with at most 3 states. This degree is relatively small compared to the bound $C = 64$ that Theorem 14 gives.

The next proposition will consider the case where $A = d = 1$ and q is prime.

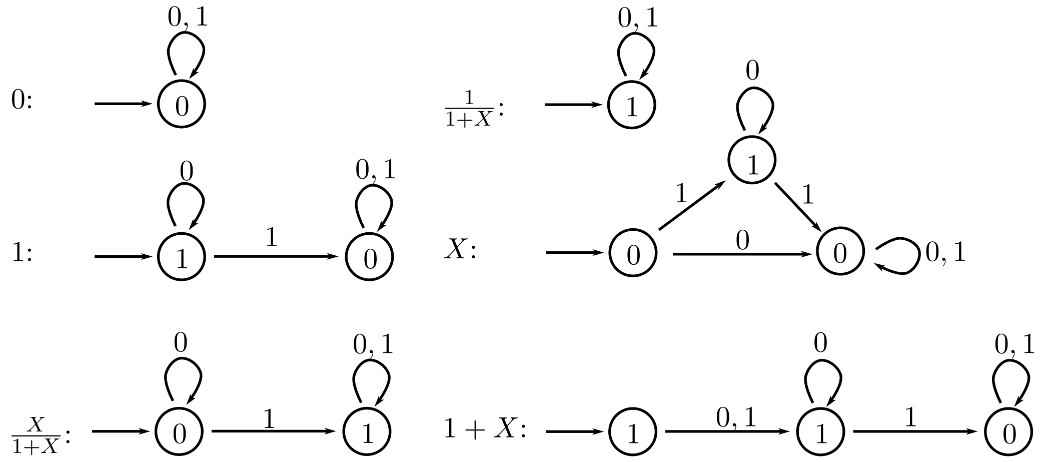


Figure 3.3: These 2-automata over $\mathbb{F}_2$ generate all possible rational power series over $\mathbb{F}_2$ that satisfy $s(X)F - r(X) = 0$, with $\deg(r), \deg(s) \leq 1$.

Proposition 15. *Let p be a prime and let F be a power series in $\mathbb{F}_p[[X]]$ that satisfies $s(X)F - r(X) = 0$ for some $s, r \in \mathbb{F}_p[X]$ with $\deg(s), \deg(r) \leq 1$ and $s(0) \neq 0$. Then there is a p-automaton that generates F with at most $\max\{p, 4\}$ states.*

Proof. Since $s(0) \neq 0$ we can rewrite $F = r(X)/s(X)$ as

$$F = \frac{a + bX}{1 - cX} = a + \frac{bX + acX}{1 - cX} = a + (b + ac)X \sum_{n \geq 0} c^n X^n,$$

where $a, b, c \in \mathbb{F}_p$. If $c \neq 0$, we define $k = \frac{b+ac}{c}$ and rewrite F as

$$F = a + k \sum_{n \geq 0} c^{n+1} X^{n+1} = a + \sum_{n \geq 1} k c^n X^n.$$

So the sequence $\mathbf{a} = (a_n)_{n \geq 0}$ that corresponds to F satisfies

$$a_n = \begin{cases} a & \text{for } n = 0, \\ kc^n & \text{for } n \neq 0. \end{cases}$$

To compute the size of the kernel, note that we have $c^p = c$, so for $n \geq 0$, $i \geq 0$ and $0 \leq j < p^i$ we have

$$a_{p^i \cdot n + j} = \begin{cases} a & \text{for } n = j = 0 \\ kc^{p^i \cdot n + j} & \text{otherwise} \end{cases} = \begin{cases} a & \text{for } n = j = 0 \\ kc^{n+j} & \text{otherwise} \end{cases} = a_{n+j}.$$

As a result, the p-kernel of $\mathbf{a}$ consists of the subsequences $(a_{n+j})_{n \geq 0}$ for $j \geq 0$. When not both j and n are zero, we have that $a_{n+j} = a_{n+j+p-1}$. As a result, the p-kernel contains at most p elements.

When $c = 0$, then F is just a linear function, and the p-kernel consists of at most four elements: $(a, b, 0, 0, \ldots), (a, 0, 0, \ldots), (b, 0, 0, \ldots)$ and $(0, 0, 0, \ldots)$. Hence, the kernel of $\mathbf{a}$ exists of at most $\max\{p, 4\}$ elements. By Lemma 11, there is a p-automaton that generates F with at most $\max\{p, 4\}$ states.

□

Note that for the power series as described in Proposition 15, Theorem 14 gives a bound of $C = p^{2(p(p-1)+1)}$. If we reconsider the set of examples in Proposition 13, Theorem 14 gives us a bound of $C = 2^{2^m(1-1/2(-2+2^{2^m})(4-2^{2^m}-3\cdot 2^m+4^m))}$, while the power series are actually generated by 2-automata of size $m+2$. In both cases, the bound of Theorem 14 is much too high.

4
Conclusion

Christol's theorem gives us the opportunity to go from finite automata to power series and back again. The goal of this master's project was to make this theorem more effective by finding bounds for both directions. We found that the size of the kernel of a sequence is essential: it equals the number of states of a minimal (reverse reading) automaton that generates the sequence.

The first question of this thesis was: Given a finite automaton with m states, what can we say about the algebraic degree of the corresponding power series? We found an answer by looking at the first part of the proof of Christol's theorem. Our result is summarized in Theorem 12 and states: Given a finite q-automaton with m states, the degree of the corresponding power series is at most $q^m - 1$. We found an infinite class of examples that show that this bound should indeed be exponential, but we have not found an example that makes the bound tight.

The second question was: Given an algebraic power series of algebraic degree d, can we find a bound on the number of states of a minimal automaton that generates it? This question was a bit naive, since there are infinitely many power series over $\mathbb{F}_q$ of degree d, but there are only a finite number of finite q-automata of a given size. So the algebraic degree of F is not enough to bound the number of states of a minimal automaton. If F satisfies a polynomial $\alpha_0 + \alpha_1 F + \cdots + \alpha_d F^d = 0$, then we also need the maximym degree A of the coefficients α_i to find a bound on the number of states of a minimal generating automaton. The bound that we found after examining the steps the proof of Ore's lemma and of the second part of the proof Christol's theorem is doubly exponential in d, and can be found in Theorem 14. We investigated the case that $A = d = 1$ and q is prime. We found that we can bound the algebraic degree in this case by $\max\{q, 4\}$, which is much smaller bound than the bound given by Theorem 14.

Further research could focus on finding examples to make both bounds tight, or finding ways to lower these bounds. Since we used in this thesis right-to-left reading automata, and our results are applicable to these automata, the link with forward reading automata has still to be explored. Furthermore, we have found the algebraic degree of the power series $\mathcal{D}((1+X+Y^a)^{-1})$ for $a = 1, 2$, and we found corresponding minimal automata. What happens for $a > 2$ can be a subject for further research.

Appendix: Forward reading 2-automata over $\mathbb{F}_2$ with two states

In early stages of the work on this thesis, we considered all *forward reading* 2-automata over $\mathbb{F}_2$ with two states, including 2-automata that are not leading zeros invariant. For each automaton we found the corresponding power series and its minimal polynomial. This process is similar to what we did for reverse reading 2-automata with two states in the end of Section 3.1. Here, we describe all forward reading 2-automata with two states that are leading zeros invariant, which means in this case that $\delta(q_0, 0) = q_0$.

There are 32 leading zeros invariant 2-automata over $\mathbb{F}_2$ with two states, since there are 2^3 possibilities to define δ and 2^2 ways to label the states. If both output labels are 0, then the formal power series that corresponds to the automaton equals $F_0 = 0$, no matter how δ is defined. Similarly, if both states are labeled 1, then the corresponding power series is $F_1 = \sum_{n\geq 0} X^n = (1+X)^{-1}$. So we only have to consider the case where the 2-automaton has two different labels. Note that if F is the power series corresponding to such a 2-automaton M, then swapping the output labels of M results in the power series $F + (1+X)^{-1}$. Once we have a minimum polynomial for F, it is easy to find the minimum polynomial for $F + (1+X)^{-1}$. So we will only consider the 2-automata where $\tau(q_0) = 1$ and $\tau(q_1) = 0$.

In 4 cases of the 8 possibilities to define δ we have $\delta(q_0, 0) = q_0 = \delta(q_0, 1)$, which means that the automaton will always stay in state q_0 with output label 1 and hence will generate F_1. The four remaining automata are shown in figure 1, and are called M_3, M_4, M_5 and M_6. Let $F_i = \sum_{n\geq 0} a_n^{(i)} X^n$ be the power series corresponding to automaton i, for $i = 3, 4, 5, 6$. Clearly, we have $F_3 = 1$, since $a_n^{(3)} = 1$ if only if $n = 0$. So the minimum polynomial of F_3 is $T - 1 = 0$. For M_4 we have that $a_n^{(4)} = 1$ if and only if $(n)_2$ ends with a zero. So $F_4 = \sum_{n\geq 0} X^{2n} = \frac{1}{1+X^2}$, which has as minimum polynomial $(1 + X^2)T - 1 = 0$.

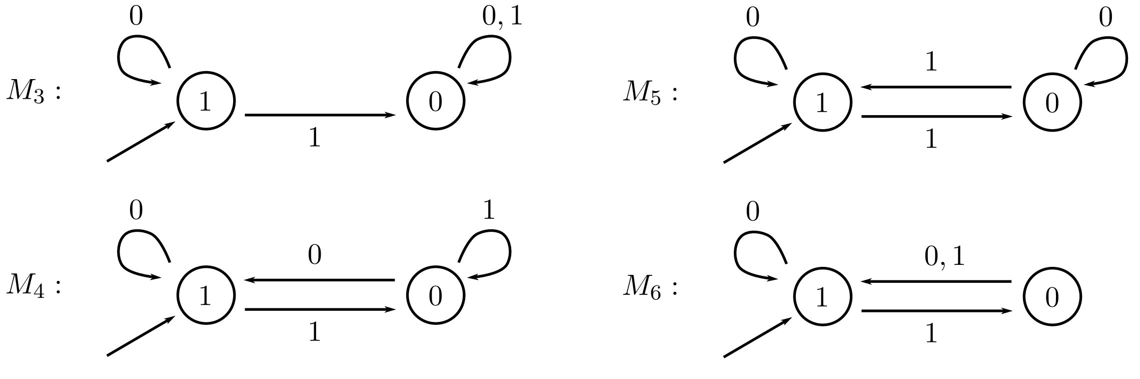


Figure 1: The remaining four non-trivial 2-automata with 2 states with $\tau(q_0) = 1$ and $\tau(q_1) = 0$.

The 2-automaton M_5 is similar to the automaton in figure 2.1, but with the labels swapped. The automaton M_5 generates the other Thue-Morse sequence: $a_n^{(5)} = 1$ if and only if the binary represenation of n contains an *even* number of ones. Note that it does not matter how this automaton reads the input, for both directions it produces the same sequence. Let s_n be the number of ones in $(n)_2$, then we can express F_5 as $F_5 = \sum_{n\geq 0}(s_n+1)X^n$. We compute

$$\begin{aligned}
(1+X)F_5^2 &= (1+X)\sum_{n\geq 0}(s_n+1)X^{2n} = \sum_{n\geq 0}(s_n+1)X^{2n} + \sum_{n\geq 0}(s_n+1)X^{2n+1} \\
&= \sum_{n\geq 0}(s_{2n}+1)X^{2n} + \sum_{n\geq 0}(s_{2n+1}+1)X^{2n+1} + \sum_{n\geq 0}X^{2n+1} = F_5 + \sum_{n\geq 0}X^{2n+1} \\
&= F_5 + \frac{X}{1+X^2},
\end{aligned}$$

so F_5 is a zero of $(1+X)^3T^2+(1+X^2)T+X=0$.

By looking at automaton M_6 we see that $a_n^{(6)}=1$ if and only if $(n)_2$ ends with an even number of ones, so we can express the corresponding power series as $F_6 = \sum_{i\geq 0}(\sum_{n\geq 0}X^{2n})^{4^i}X^{4^i-1}$. Consider

$$XF_6^2 = X\Big(\sum_{i\geq 0}\big(\sum_{n\geq 0}X^{2n}\big)^{4^i}X^{4^i-1}\Big)^2 = \sum_{i\geq 0}\big(\sum_{n\geq 0}X^{2n}\big)^{4^{i+1}}X^{4^{i+1}-1} =: \sum_{n\geq 0}b_nX^n,$$

then we see that $b_n=1$ if and only if $(n)_2$ ends with an odd number of ones. So $XF_6^2+F_6=\frac{1}{1+X}$, hence F_6 is a zero of $X(1+X)T^2+(1+X)T+1=0$.

Hence just like for reverse reading automata, all forward reading 2-automata with two states that are leading zeros invariant, generate power series with algebraic degree at most 2.

Bibliography

Jean-Paul Allouche and Jeffrey O. Shallit. *Automatic Sequences - Theory, Applications, Generalizations.* Cambridge University Press, 2003. ISBN 978-0-521-82332-6.

G. Christol, T. Kamae, M. Mendès France, and G. Rauzy. Suites algébriques, automates et substitutions. *Bull. Soc. Math.France*, 108:401–419, 1980.

N.P. Fogg, V. Berthé, S. Ferenczi, C. Mauduit, and A. Siegel. *Substitutions in Dynamics, Arithmetics and Combinatorics.* Springer, 2002. ISBN 978-3-540-44141-0.

Harry Furstenberg. Algebraic functions over finite fields. *Journal of Algebra*, 7:271–277, 1967.

John E. Hopcroft and Jeffrey D. Ullman. *Introduction to Automata Theory, Languages and Computation.* Addison-Wesley Publishing Company, 1979. ISBN 0-201-02988-X.

A.M. Legendre. Essai sur la théorie des nombres. *Chez Courcier, Imprimeur-Libraire pour les Mathématiques, quai des Augustins*, 57:8–10, 1808.

A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. *Norske vid. Selsk. Skr. Mat. Nat. Kl.*, 1:1–67, 1912.