university of groningen

faculty of mathematics and natural sciences

On the algebraicity and automaticity of generalized power series

Master Project Mathematics

July 2016

Student: H.K. Kruizinga

First supervisor: prof .dr. J. Top

Second supervisor: prof. dr. H. Waalkens

Abstract

In this thesis we will consider 'generalized power series' $\mathbb{F}_q((t^{\mathbb{Q}}))$. We look at the connection between the algebraic degree of such series over $\mathbb{F}_q(t)$ and the size of the finite automaton accepting the exponents. This was done before for the regular power series $\mathbb{F}_q((t))$, by examining the proof of Christol's theorem. With the recent expansion of this theorem, we try to obtain similar results for generalized power series. Given an automaton, we found a bound on the algebraic degree. For the other direction we only point out the steps to obtain a bound.

1 Introduction

In this thesis we consider a finite field $\mathbb{F}_q$, with q some power of a prime p. Christol proved that a power series $f = \sum_i f(i)t^i$, $f : \mathbb{Z}_{\geq 0} \to \mathbb{F}_q$ from $\mathbb{F}_q[[t]]$ is algebraic over $\mathbb{F}_q(t)$ if and only if f is automatic: there exists a finite automaton accepting the base q expansions of the support $\sup(f) = \{i | f(i) \neq 0\}$. [3]

There is a connection between the algebraic degree of a power series and the number of states of the automaton. Given the number of states, there is a bound on the algebraic degree. Reversely, given the polynomial $P(T) \in \mathbb{F}_q(t)[T]$, such that $P(f) = 0$, there is a bound on the number of states. [4]

Kedlaya stated and proved an expansion of Christol's theorem for 'generalized power series'. The support of these generalized power series is not just restricted to $Z_{\geq 0}$, but will be a (well-ordered) subset of $\mathbb{Q}$. Moreover, the set of generalized power series is an algebraically closed field, so this expansion gives a complete picture of elements algebraic over $\mathbb{F}_q(t)$. [8]

Our contribution to this subject is considering the connection between algebraic degree of a power series and the number of states of the corresponding automaton. Given a specific automaton, we find a bound on the algebraic degree. For the reverse, we did not find an explicit bound on the number of states. We will restrict ourselves to specific generalized power series and we will point out the difficulties in obtaining a bound.

2 Automata

In this section we introduce some automata theory. This and more on automata theory can be found in the book "Automatic Sequences" [1].

We start by defining a *deterministic finite automaton (DFA)*. It is a simple model of computation; given a string as input it either accepts or rejects it.

Definition 2.1. A *deterministic finite automaton (DFA)* is a 5-tuple $M = (Q, \Sigma, \delta, q_0, F)$, with

- Q is a finite set containing *states*,
- Σ is the finite *input alphabet*,
- $\delta : Q \times \Sigma \to Q$ is the *transition function*
- $q_0 \in Q$ is the *initial state*,

- $F \subset Q$ are the *accepting/final states*.

A *DFA with output (DFAO)* is a 6-tuple $M = (Q, \Sigma, \delta, q_0, \Delta, \tau)$, with

- Δ is the finite *output alphabet*,
- $\tau : Q \to \Delta$ is the *output function*.

Starting in the initial state q_0, a DFA moves to another state (according to δ) when a key from the input alphabet is pressed. A DFAO moreover gives an output when we are in a certain state.

Definition 2.2. Let Σ^* be the set of all finite strings consisting of elements of Σ. We can extend the domain of δ as follows: $\delta(q, \epsilon) = q$ for the empty string $\epsilon \in \Sigma^*$, and $\delta(q, aw) = \delta(\delta(q, w), a), \forall w \in \Sigma^*, a \in \Sigma$. Any subset of Σ^* is called a *language*.

Example 2.3. The alphabet that we will use most in examples is $\Sigma = \{0, 1\}$. Then $\Sigma^* = \{\epsilon, 0, 1, 00, 01, 10, 11, 000, 001, \dots\}$

Definition 2.4. A string $s \in \Sigma^*$ is accepted by a DFA M if and only if $\delta(q_0, s) \in F$. We denote the set of all strings accepted by M by $L(M)$: the *language accepted by* M. We call a language *regular* if it is accepted by a DFA.

Note. We say that a DFAO accepts a string s if and only if $\tau(\delta(q_0, s)) \neq 0$

In our case DFA(O)'s read strings from Σ^* from left to right. There also exist reverse reading DFA(O)'s, which read strings from right to left. They are connected by the following operation:

Definition 2.5. The *reversal operator* rev: $\Sigma^* \to \Sigma^*$ is given by

$$\text{rev}(s_1 s_2 \dots s_n) = s_n \dots s_2 s_1.$$

Lemma 2.6. *For a finite alphabet Σ we have that:*

- *A language $L \subset \Sigma^*$ is regular $\Leftrightarrow$ rev(L) (the reverse language) is regular.*
- *Let the DFAO with output alphabet Δ accepting a regular language L have n states, then the DFAO accepting rev(L) has at most $|\Delta|^n$ states.*

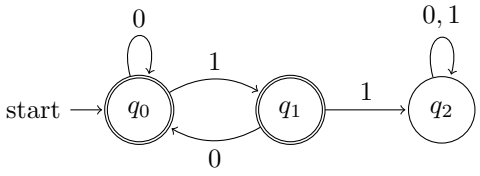
Proof. [1, Corollary 4.3.4/5] □

Remark. This means that a forward reading automaton does not (necessarily) have the same number of states as a reverse reading automaton accepting the same language.

We can represent a DFA by a *transition diagram*:

Definition 2.7. The transition diagram linked to a DFA $M = (Q, \Sigma, \delta, q_0, F)$ is given by an edge-labeled directed graph. It has vertex set Q and it has an edge from $q \in Q$ to $q' \in Q$ labeled by $s \in \Sigma$ if $\delta(q, s) = q'$. The initial state q_0 is represented by an arrow (with start) and the accepting states from F are drawn with double circles.

Example 2.8. Consider the DFA $M = (Q, \Sigma, \delta, q_0, F)$, with $Q = \{q_0, q_1, q_2\}$, $\Sigma = \{0, 1\}$ and $F = \{q_0, q_1\}$ given by the following transition diagram:



So δ is given by:

- $\delta(q_0, 0) = q_0$
- $\delta(q_1, 0) = q_0$
- $\delta(q_2, 0) = q_2$
- $\delta(q_0, 1) = q_1$
- $\delta(q_1, 1) = q_2$
- $\delta(q_2, 0) = q_2$

Starting in the initial state q_0, it will stay in q_0 if the automaton reads a 0, it will move to q_1 if the automaton reads a 1. In q_1 the automaton moves back to q_0 if it reads a 0 and it moves to q_2 if it reads a 1. Once the automaton is in state q_2 it will always remain there (also called a *sink state*). So the language accepted by M is given by all strings without two (or more) consecutive ones: $\{\epsilon, 1, 10, 100, 101, 1000, 1001, 1010, 10000, 10001, 10010, 10100, 10101, \dots\}$.

Definition 2.9. For a language $L \subset \Sigma^*$, an equivalence relation $\sim_L$ is defined by $x \sim_L y$ if and only if $xz \in L \Leftrightarrow yz \in L$, $\forall z \in \Sigma^*$.

Lemma 2.10. *(Myhill-Nerode theorem) The language L is regular if and only if Σ^* has finitely many equivalence classes under $\sim_L$. Furthermore, if L is regular, the number of states of the DFA corresponding to it equals the number of equivalence classes.*

Remark. The DFA $M = (Q, \Sigma, \delta, q_0, F)$ corresponding to a regular language L is given by

- Q is the set of equivalence classes under $\sim_L$.
- δ is then a function sending the equivalence class of $x \in \Sigma^*$ and some $s \in \Sigma$ to the equivalence class of xs.
- q_0 is the equivalence class of the empty string $\epsilon \in \Sigma^*$.
- F is the set of equivalence classes of elements of L.

Example 2.11. We use Example 2.8 to clarify this. Recall that the language is given by all strings without two (or more) consecutive ones. We have the following three equivalence classes:

$$\begin{aligned}\langle\epsilon\rangle &= \{s \in \{0,1\}^* | s_n = 0, s \text{ has no 2 consecutive ones or } s = \epsilon\},\\ \langle 1\rangle &= \{s \in \{0,1\}^* | s_n = 1, s \text{ has no 2 consecutive ones}\},\\ \langle 11\rangle &= \{s \in \{0,1\}^* | s \text{ has 2 consecutive ones}\},\end{aligned}$$

since for $s \in \langle\epsilon\rangle$, $sz \in L \Leftrightarrow z \in L$, and for $s \in \langle 1\rangle$, $sz \in L \Leftrightarrow z \in L, z_0 = 0$ and for $s \in \langle 11\rangle$, $sz \notin L, \forall z$.

The transition diagram from Example 2.8 indeed has three states: q_0 corresponding to $\langle\epsilon\rangle$, q_1 corresponding to $\langle 1\rangle$, q_0 and q_2 corresponding to $\langle 11\rangle$.

We need the following definitions: [8, Definition 7.1.1/2]

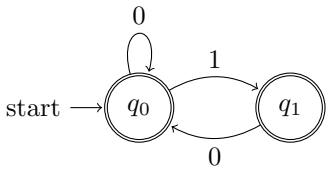
Definition 2.12. We call a state q *reachable* from q' if there exists an $s \in \Sigma$ such that $\delta(q', s) = q$. We call a DFAO *minimal* if it has no unreachable states.

Whenever a DFAO is not minimal, we can remove its unreachable states to obtain a minimal DFAO accepting the same language.

Definition 2.13. We call a state $q \in Q$ *relevant* if there is an accepting state reachable from q (i.e., $\exists s \in \Sigma^*$ such that $\delta(q, s) \in F$).

Hereafter, we omit the irrelevant states in transition diagrams. Whenever a DFA(O) 'normally' would go to a irrelevant state, it now just directly is rejected (or has output 0).

Example 2.14. The relevant states of the automaton in Example 2.8 are q_0 and q_1. Therefore we can omit the irrelevant state q_2 from the transition diagram, without losing information:



So when the automaton is in state q_1 and reads a 1, it will move to an invisible 'garbage state'.

3 Christol's theorem

Consider input alphabet $\Sigma_b = \{0, 1, \ldots, b-1\}$, then the *value* of a string $s = s_1 s_2 \ldots s_n \in \Sigma_b^*$ is given by:

$$v(s) = \sum_{i=1}^{n} s_i b^{n-i}.$$

With this notion we link a language L to a power series over $\mathbb{F}_q$ as follows:

$$\sum_{s \in L} \tau(s) t^{v(s)}$$

Example 3.1. The values of the language accepted in Example 2.8 are given by $0, 1, 2, 4, 5, 8, 9, 10, 16, 17, 18, 20, 21, \ldots$. Therefore the corresponding power series looks like:

$$f = 1 + t + t^2 + t^4 + t^5 + t^8 + t^9 + t^{10} + t^{16} + t^{17} + t^{18} + t^{20} + t^{21} + \ldots.$$

Definition 3.2. We call a function $f : \mathbb{Z}_{\geq 0} \to \Delta$ *b-automatic* if there is a DFAO $M = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ such that for any $n \in \mathbb{Z}_{\geq 0}$, $f(n) = \tau(v(n))$. We call a $S \subset \mathbb{Z}_{\geq 0}$ *b-regular* if its characteristic function $\mathbb{1}_S(s)$ is b-automatic.

The following theorem shows a connection between being algebraic and being automatic. [3]

Theorem 3.3 (Christol)**.** *For q a power of p and $\{a_i\}_{i=0}^{\infty}$ a sequence with $a_i \in \mathbb{F}_q$, the corresponding power series $\sum_{i=0}^{\infty} a_i t^i \in \mathbb{F}_q[[t]]$ is algebraic over $\mathbb{F}_q(t)$ if and only if $f(i) = a_i$ is p-automatic.*

The effectivity of this theorem is explored in [4], where, in contrast to this thesis, reverse reading automata are considered. The results are in the following two theorems:

Theorem 3.4. *Given a reverse reading DFAO with m states, the algebraic degree of the corresponding power series is at most $q^m - 1$.*

Theorem 3.5. *Given a power series algebraic of degree d: a zero of $P(T) = \alpha_0 + \alpha_1 T + \cdots + \alpha_d T^d$, $deg(\alpha_i) \leq A, \forall i = 1, ..., d$, the maximal number of states of the correspondig reverse reading DFAO is*

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

We consider again Example 2.8 to show this is indeed true:

Example 3.6. Let $q = 2$, and consider the automaton from Example 2.8, which accepted all strings with no more than one consecutive 1. To explore whether f is indeed algebraic over $\mathbb{F}_2(t)$, we look at the support of f^{2^i}. This is easy to look at since we work in $\mathbb{F}_2(t)$:

$$\begin{aligned}\sup(f) &= \{s | s \text{ has no 2 consecutive ones}\},\\ \sup(f^2) &= \{s0 | s \text{ has no 2 consecutive ones}\},\\ \sup(f^4) &= \{s00 | s \text{ has no 2 consecutive ones}\},\end{aligned}$$

for $s \in \Sigma^* = \{0, 1\}^*$. We get the following:

$$\begin{aligned}\sup(f) &= \{s | s \text{ has no 2 consecutive ones}\}\\ &= \{s0 | s \text{ has no 2 consecutive ones}\} \cup \{s01 | s \text{ has no 2 consecutive ones}\}\\ &= \sup(f^2) \cup \sup(tf^4).\end{aligned}$$

So we obtain $tf^4 + f^2 = f$, from which we see that f indeed is algebraic (of degree 3) over $\mathbb{F}_2(t)$: f is a root of $tx^3 + x + 1 = 0$.

4 Generalized power series

There exist also other series outside $\mathbb{F}_q[[t]]$ that are algebraic over $\mathbb{F}_q(t)$. As an example, consider:

Example 4.1 (Chevalley [2])**.** The series

$$x = t^{-1/p} + t^{-1/p^2} + t^{-1/p^3} + t^{-1/p^4} + \dots$$

satisfies $tx^p - tx - 1 = 0$. Here t^{-1/p^n} is and element of an algebraic closure of $\mathbb{F}_q(t)$ satisfying $(t^{-1/p^n})^{p^n} = t^{-1}$. Below we describe extensions of $\mathbb{F}_q(t)$

in which adding infinitely many such elements makes sense. Since we work in characteristic p, note that

$$x^p = t^{-1} + t^{-1/p} + t^{-1/p^2} + t^{-1/p^3} + \dots,$$

so $x^p = t^{-1} + x$ and we indeed get that x is a zero of $tx^p - tx + 1 \in \mathbb{F}_q(t)$.

The series given in Example 4.1 is a so-called generalized power series. These series where first introduced by Hahn in 1907. [5]

In order to define this, we first need the following definition:

Definition 4.2. A set $S \subset \mathbb{Q}$ is called well-ordered if and only if every nonempty subset of S has a least element.

Note. This is equivalent to saying that there do not exist infinitely descending sequences of elements of S.

Definition 4.3. *Generalized power series* are of the form $\sum_i f(i)t^i$, where the support of f is well-ordered and lies in $\mathbb{Q}$.

Example 4.4. The support of x in Example 4.1 is given by the set

$$S = \{-1/p, -1/p^2, -1/p^3, -1/p^4, \dots\} = \{-1/p^n | n \in \mathbb{Z}_{>0}\}.$$

Any subset of S is given by $\{-1/p^{n_i} | n_1 < n_2 < \dots \in \mathbb{Z}_{>0}\}$ and has a least element: $-1/p^{n_1}$. So the support S of x is well-ordered, which means that x is a generalized power series.

Theorem/Definition 4.5. The set of generalized power series is a field and will be denoted by $\mathbb{F}_q((t^{\mathbb{Q}}))$.

For this to be true, we need to find out whether generalized power series are closed under addition, multiplication and taking inverses.

Consider the generalized power series $f = \sum_i f(i)t^i$ and $g = \sum_i g(i)t^i$ both with well-ordered support S_f and S_g. Then $f + g$ is given by:

$$f + g = \sum_i (f(i) + g(i))t^i.$$

So its support is contained in $S_f \cup S_g$ and the (subset of a) union of two well-ordered sets is again well-ordered. [9, Lemma 2.9]

The product fg is given by:

$$fg = \sum_i f(i)t^i \cdot \sum_j g(j)t^j = \sum_k \big(\sum_{i+j=k} f(i)g(j) \big) t^k.$$

This has support contained in $S_f + S_g$, which is also well-ordered. [9, Lemma 2.9].

Moreover, any $f \in \mathbb{F}_q((t^{\mathbb{Q}}))$ we can write as $at^m(1 - f')$, $f' \in \mathbb{F}_q((t^{\mathbb{Q}}))$ with $f'(0) = 0$ such that $\big(\sum_{n=0}^{\infty}(f')^n\big)(1 - f') = 1$, so $a^{-1}t^{-m}\sum_{n=0}^{\infty}(f')^n$ defines an inverse for f. It has support contained in $\cup_{n=0}^{\infty}\mathrm{sup}(f')^n$, where $\mathrm{sup}(f') = \{s - m | m = \min_{s \in S_f} s, s \in S_f \backslash \{m\}\}$ is well-ordered, so again well-ordered. [9, Lemma 2.10]

Theorem 4.6. *The field $\overline{\mathbb{F}_q}((t^{\mathbb{Q}}))$ is algebraically closed.*

Proof. [6, Proposition 1]. □

So all series algebraic over $\mathbb{F}_q(t)$ are elements of $\mathbb{F}_q((t^{\mathbb{Q}}))$. For $\mathbb{F}_q((t^{\mathbb{Q}}))$ we will state an expansion of Christol's theorem. Before we do this, we need some definitions in order to relate generalized power series to automata theory.

5 Automata with radix point

For our purposes we want an automaton accepting finite strings linked to elements in $\mathbb{Q}$. We will construct such automata by extending the input alphabet $\Sigma' = \{0, 1, ..., b-1\}$ we used before, to $\Sigma = \{0, 1, ..., b-1, .\}$ (where we added a radix point).

Definition 5.1. A *valid base b expansion* $s = s_1 \dots s_n \in \Sigma^*$ must satisfy $s_1 \neq 0 \neq s_n$ and $\exists! k$ such that s_k is the radix point. The *value of a valid base b expansion* is given by

$$v(s) = \sum_{i=1}^{k-1} s_i b^{k-1-i} + \sum_{i=k+1}^{n} s_i b^{k-i}. \tag{1}$$

Each element $i \in S_p := \{m/p^n | m, n \in \mathbb{Z}_{\geq 0}\} = \mathbb{Z}[p^{-1}]_{\geq 0}$ can be written in the above form, so we can denote the valid base b expansion of i by $s(i) = s_1 \dots s_n$, with s_i as in equation (1). So this gives a bijection between valid base b expansions and S_p.

Definition 5.2. Let M be a DFAO with input alphabet $\Sigma = \{0, 1, ..., b-1, .\}$. We call a state $q \in Q$ *preradix* if there is a base b expansion $s \in \Sigma'^*$ (without radix point), such that $q = \delta(q_0, s)$. Moreover, define Q_{pre} to be the set of preradix states. Equivalently, we call a state $q \in Q$ *postradix* if there is a valid base b expansion $s \in \Sigma^*$ (with radix point), such that $q = \delta(q_0, s)$. Also, define Q_{post} to be the set of postradix states.

Note. A relevant state is either preradix of postradix, since a DFA only accepts valid base b expansions, which have exactly one radix point.

Again, we need to define what it means to be automatic: now for generalized power series:

Definition 5.3. We call a function $f : \mathbb{Q}_{\geq 0} \to \delta$ *b-automatic* if there is a DFAO $M = (Q, \Sigma, \delta, q_0, F)$ (accepting valid base b expansions), such that for any $i \in \mathbb{Q}_{\geq 0}$, $f(i) = \tau(s(i))$. We call a $S \subset \mathbb{Q}_{\geq 0}$ *b-regular* if its characteristic function $\mathbb{1}_S(s)$ is b-automatic.

Note. A function $f : S_b \to \Delta$ is b-automatic if and only if $f^{-1}(d)$ is b-regular for all $d \in \Delta$.

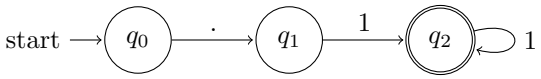
Note. We restrict ourselves to automata that only can accept strings linked to elements in $\mathbb{Q}_{\geq 0}$. So we do not cover all generalized power series (see Example 4.1). But we can build automata for such series indirectly, using the following notion.

Definition 5.4. For $f : \mathbb{Q} \to \mathbb{F}_q$ a function with well-ordered support S, we call $\sum_i f(i)t^i$ *p-quasi-automatic* if:

- There exist integers a, b with $a > 0$ such that $aS + b \subset \{m/p^n | m, n \in \mathbb{Z}_{\geq 0}\} =: S_p$ (the nonnegative p-adic rationals).
- For such a, b, the function $f_{a,b} : S_p \to \mathbb{F}_q$, given by $f_{a,b}(x) = f((x-b)/a)$ is p-automatic.

Example 5.5. The automata corresponding to the series in Example 4.1 is p-quasi-automatic: we have that $f(i) = 1 \Leftrightarrow i = -1/p^n$, for $n \in \mathbb{Z}_{>0}$. So for $a = b = 1$, we obtain $f_{a,b}(i) = 1 \Leftrightarrow i = (p^n - 1)/p^n$, for $n \in \mathbb{Z}_{>0}$.

Consider $p = 2$, then the valid base 2 expansion of $i = (p^n - 1)/p^n$ is given by $s(i) = .\underbrace{11\ldots1}_{n\text{-times}}$. This is given by the following DFA:



since it only accepts strings of the form $.1^*$.

Definition 5.6. A DFA M with input alphabet $\Sigma = \{0, 1, \ldots, b-1, .\}$ is called *well-formed* if the language accepted by M only consists of valid base b expansions of elements of a subset of S_b. It is called *well-ordered* if moreover that subset is well-ordered.

For a DFA M to be well-formed will mean the following:

- There is no zero transition from q_0 (i.e.: $\delta(q_0, 0)$ is no relevant state), since valid base b expansions do not start with 0.
- Preradix states have output zero (i.e.: $\tau(q) = 0, \forall q \in Q_{\text{pre}}$), since $\delta(q_0, s) \in Q_{\text{post}}$ for all valid base b expansions s.
- There is no zero transition to an accepting state (i.e.: if $\tau(q) \neq 0$, then $\delta(q', 0) \neq q, \forall q' \in Q$), since valid base b expansions do not end with a 0.

For a DFA M to be well-ordered, we get some additional restraints: [8, section 7]

Let $G = (V, E)$ be a directed graph, with vertex set V and (directed) edge set E.

Definition 5.7. A directed graph G is called a *rooted saguaro* with *root* $v \in V$ if the following holds:

- Each vertex lies on at most one minimal cycle.
- There exist directed paths from v to each vertex of G.

A saguaro can have more than one root (it happens when v lies on a cycle). We call a minimal cycle of a rooted saguaro a *lobe*.

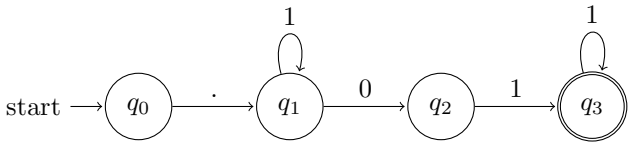
Definition 5.8. Let G be a rooted saguaro. A *proper b-labeling* of G is given by a function $\ell : E \to \{0, ..., b-1\}$ satisfying:

1. For $v, w, x \in V$ and $vw, vx \in E$ we have $\ell(vw) \neq \ell(vx)$.

2. For $v, w, x \in V$ and $vw \in E$ on a lobe, whereas $vx \in E$ not on a lobe, then $\ell(vw) > \ell(vx)$.

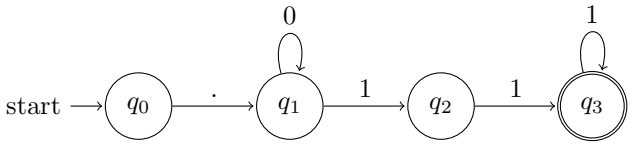
Theorem 5.9. *Let the DFA M be minimal and well-formed. Then M is well-ordered if and only if for each relevant postradix state q, the subgraph G_q (consisting of relevant states reachable from q) is a rooted saguaro with root q, equipped with a proper b-labeling.*

Example 5.10. Let M be given by:



Note that M indeed only accepts valid base 2 expansions, so it is well-formed. Also, it is well-ordered: Consider for example G_{q_1}. It is a rooted saguaro and has a proper 2-labeling.

Example 5.11. Now we consider a DFA M that is not well-ordered:



Note that we just swapped two labels from the previous example. This transition diagram fails to have a proper 2-labeling as it does not satisfy the second condition of Definition 5.8.

This automaton accepts the sequence $\{.11, .011, .0011, .00011, \dots\}$, an infinitely descending sequence, so the support of f is indeed not well-ordered.

6 The theorem

Kedlaya proved the following theorem as an expansion of Christol's theorem:

Theorem 6.1 (Kedlaya)**.** *For q a power of p and $f : \mathbb{Q} \to \mathbb{F}_q$ a function with well-ordered support S, the corresponding generalized power series $\sum_{i \in S} f(i)t^i \in \mathbb{F}_q((t^{\mathbb{Q}}))$ is algebraic over $\mathbb{F}_q(t)$ if and only if it is p-quasi-automatic.*

Example 6.2. In Example 5.5, f is indeed both 2-automatic and algebraic.

For Theorem 6.1 we want to find the analogue of Theorems 3.4 and 3.5, which we will (try to) do in the following two sections.

Note. Theorems 3.4 and 3.5 are about the number of states of a reverse reading automaton, whereas we consider the number of states of a forward reading automata. By Lemma 2.6 this is not equivalent,

7 From automatic to algebraic

In this section, we consider p-quasi-automatic generalized power series h given by $\sum_i h(i)t^i \in \mathbb{F}_q((t^{\mathbb{Q}}))$, which will correspond to a DFAO or $q-1$ (different) DFA's. We will follow along with the proof of the proposition below [8, Lemma 5.1.2] to find a bound on the algebraic degree of this series given the number of states of the automaton.

Proposition 7.1. *A p-quasi-automatic generalized power series $x = \sum_i x_i t^i \in \mathbb{F}_q((t^{\mathbb{Q}}))$ is algebraic over $\mathbb{F}_q(t)$.*

We will use the effective variant of [8, Lemma 3.3.5]

Lemma 7.2. *Let $K \subset L$ be fields of characteristic p. Consider the matrices $A, B \in K^{n\times n}$ (with at least one of them invertible) and some column vectors $\mathbf{w} \in K^n$, $\mathbf{v} \in L^n$. Let σ denote the p-th power frobenius map. If $A\mathbf{v}^\sigma + B\mathbf{v} = \mathbf{w}$, then the entries of $\mathbf{v}$ are algebraic of degree at most $p^{n^2+n} - 1$ over K.*

Proof. If A is invertible, we obtain the following relations for $i \in \mathbb{Z}_{\geq 0}$:

$$\mathbf{v}^{\sigma^i} = U_i\mathbf{v} + \mathbf{w}_i,$$

for some $U_i \in K^{n\times n}$, $\mathbf{w}_i \in K^n$. Such vectors span a vector space of dimension at most n^2+n, so for some $m \leq n^2+n$ there exist nontrivial $c_0, c_1, ..., c_m \in K$ such that

$$c_0\mathbf{v} + c_1\mathbf{v}^\sigma + \cdots + c_m\mathbf{v}^{\sigma^m} = 0.$$

So each entry of $\mathbf{v}$ is algebraic of degree at most $p^m - 1 \leq p^{n^2+n} - 1$.
If A is not invertible, then by assumption, B must be invertible. In this case we obtain the following relations for $i \in \mathbb{Z}_{\geq 0}$:

$$\mathbf{v}^{\sigma^{-i}} = U_i\mathbf{v} + \mathbf{w}_i,$$

where the entries of U_i and $\mathbf{w}_i$ lie in $K_i := \{x \in L | x^{\sigma^i} \in K\}$. Note that $K_i \subset K_{i+1}$. Such vectors again span a vector space of dimension at most n^2+n, so for some $m \leq n^2+n$ there exist nontrivial $c_0, c_1, ..., c_m \in K_m$ such that

$$\begin{aligned} c_0\mathbf{v} + c_1\mathbf{v}^{\sigma^{-1}} + \cdots + c_m\mathbf{v}^{\sigma^{-m}} &= 0, \\ \Rightarrow c_0^{\sigma^m}\mathbf{v}^{\sigma^m} + c_1^{\sigma^m}\mathbf{v}^{\sigma^{m-1}} + \cdots + c_m^{\sigma^m}\mathbf{v} &= 0. \end{aligned}$$

So each entry of $\mathbf{v}$ is algebraic of degree at most $p^m - 1 \leq p^{n^2+n} - 1$.

□

We will also use the effective version of [8, Lemma 4.2.2]:

Lemma 7.3. *Let $\sum_i x_i t^i \in \mathbb{F}_q((t^{\mathbb{Q}}))$ be algebraic over $\mathbb{F}_q(t)$ of degree d, and $a, b \in \mathbb{Z}$ such that $a > 0$ and $gcd(a, p) = 1$. Then $\sum_i x_{ai+b}t^i$ is algebraic over $\mathbb{F}_q(t)$ of degree ad.*

Proof. We will prove this in two steps, first we consider $a = 1$, then we consider $b = 0$. Applying these two in succession will complete the proof.

For $a = 1$, let $\sum_i x_i t^i$ be a root of $P(z)$ over $\mathbb{F}_q(t)$, then $\sum_i x_{i+b} t^i = \sum_i x_i t^{i-b}$ is a root of $P(zt^b)$ which obviously has the same degree.

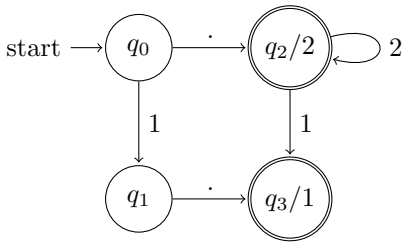
For $b = 0$, let $\tau : \sum x_i t^i \mapsto \sum x_i t^{ai}$. Then let $\sum_i x_i t^i$ be a root of $P(z) = \sum_{j=0}^d c_j z^j$ over $\mathbb{F}_q(t)$. So $\sum_i x_{ai} t^i = \sum_i x_i t^{i/a}$ is a root of $\sum_{j=0}^d c_j{}^{\tau^{-1}} z^j$ over $\mathbb{F}_q(t^{1/a})$. So it is also a root of $\sum_{j=0}^d c_j z^{aj}$ over $\mathbb{F}_q(t)$ □

Let $h(i) = x_i$, so $h = \sum_i h(i)t^i$. By definition there are integers $a > 0$ and b such that $a \cdot \sup(h) + b \subset \{m/p^n | m, n \in \mathbb{Z}_{\geq 0}\} =: S_p$. Choose a and b such that a is as small as possible. Then $h_{a,b} : S_p \mapsto \mathbb{F}_q$, given by $h_{a,b}(x) = h((x-b)/a)$, is p-automatic. Let M be the DFAO corresponding to $h_{a,b}$.

In order to follow the prove, we are going to split M into $q-1$ DFA's: For each $\alpha \in \mathbb{F}_q$, define $S_\alpha = \{j \in S_p | h((j-b)/a) = \alpha\}$, so $a \cdot \sup(h) + b = \cup_{\alpha \in \mathbb{F}_q} S_\alpha$. Note that each S_α is p-regular, so $\sum_{j \in S_\alpha} t^j$ is p-automatic. The DFA corresponding to $\sum_{j \in S_\alpha} t^j$ is given by M^α constructed as follows: Make all states of M with output α accepting states: F^α, all other states become rejecting states. Finally, only keep the essential states: Q^α. So $M^\alpha = (Q^\alpha, \delta, \Sigma, q_0, F^\alpha)$. For obtaining a bound on the algebraic degree, we need to know the size of the following types of states:

- $q \in Q^\alpha_{\text{pre}}$ such that q is reachable from a cycle, the size of which we denote by m_α.
- $q \in Q^\alpha_{\text{post}}$ such that a lobe is reachable from q, with size n_α
- The longest chain of states $q \in Q^\alpha_{\text{post}}$ such that no lobe is reachable from any of them, with size l_α.

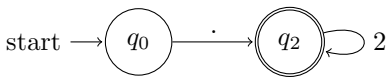
Example 7.4. Let $q = 3$ and the automaton M be given by:



Here $q_2/2$, $q_3/1$ means that $\tau(q_2) = 2$ and $\tau(q_3) = 1$. So the generalized power series is given by

$$f = t + \sum_{i=1}^{\infty} t^{(3^i-2)/3^i} + 2 + \sum_{i=1}^{\infty} 2t^{(3^i-1)/3^i} =$$
$$t + 2 + t^{1/3} + 2t^{2/3} + t^{7/9} + 2t^{8/9} + t^{25/27} + 2t^{26/27} + \dots.$$

Now M_1 is given by the DFA with the same states, but the only accepting state is q_3. And M_2 is given by the following DFA:



With this construction we can follow the proof of the following lemma in order to obtain a bound on the algebraic degree of h. [8, Lemma 5.1.1]

Lemma 7.5. *For S a p-regular subset of S_p, $\sum_{i\in S} t^i \in \mathbb{F}_p((t^{\mathbb{Q}}))$ is algebraic over $\mathbb{F}_p(t)$.*

Theorem 7.6. *The p-quasi-automatic generalized power series $\sum_i h(i)t^i$, where its corresponding DFAO has one relevant radix transition, is algebraic of degree at most*

$$a \prod_{\alpha\in\mathbb{F}_q^*} p^{{m_\alpha}^2+m_\alpha+{n_\alpha}^2+n_\alpha+l_\alpha}, \tag{2}$$

where

- *a is as in Definition 5.4,*
- *$m_\alpha = |\{q \in Q^\alpha_{pre} | q$ is reachable from a cycle$\}|$,*
- *$n_\alpha = |\{q \in Q^\alpha_{post} | a$ lobe is reachable from $q\}|$,*
- *l_α is the size of longest chain of states $q \in Q^\alpha_{post}$ such that no lobe is reachable from any of them.*

Note. For a DFAO M, with m preradix states, n postradix ones and R radix transitions, the algebraic degree of the corresponding generalized power series is at most:

$$a\big((p^{m^2+m+n^2+2n})^R\big)^{q-1}.$$

Remark. We use q for two different things, namely $q = p^k$ and for states $q \in Q$. It should be clear from the context which one we mean, otherwise we write p^k instead of q.

Proof. To each state $q \in Q^\alpha$ we assign a (generalized) power series: To preradix states q a power series $f(q) \in \mathbb{F}_p[[t]]$ and to postradix states q a generalized power series $g(q) \in \mathbb{F}_p[[t^{\mathbb{Q}}]]$, with support in $S_p \cap [0, 1)$.

First, consider $q \in Q^\alpha_{\text{pre}}$. Define $T_q = \{n \in \mathbb{N} | \delta(q_0, s(n)) = q\}$. Now set $f(q) = \sum_{n\in T_q} t^n \in \mathbb{F}_p[[t]]$. These $f(q)$ have the following relations:

$$\begin{aligned} f(q) &= \sum_{\delta(q',d)=q} t^d f^p(q') \text{ for } q \neq q_0, \\ f(q_0) &= 1+ \sum_{\delta(q',d)=q_0} t^d f^p(q'). \end{aligned} \tag{3}$$

Note that $f(q) = \sum_{n\in T_q} t^n$ gives a polynomial when there are no cycles in all paths from q_0 to q. So for such states q, $f(q)$ is algebraic of degree 1 over $\mathbb{F}_q(t)$. Now, consider the set of states $\{q_1, ...q_{m_\alpha}\}$ with a cycle on a path from q_0 to q. We can rewrite the system of equations (3) to:

$$f(q_i) = r(q_i) + \sum_{\delta(q_j,d)=q_i} t^d f^p(q_j),$$

where

$$r(q_i) = \sum_{q\notin\{q_1,\dots,q_m\},\delta(q,d)=q_i} t^d f^p(q) \in \mathbb{F}_q(t).$$

When $q_0 \in \{q_1, ..., q_{m_\alpha}\}$, we will just stick with the system of equations (3) as then $Q^\alpha_{\text{pre}} = \{q_1, ..., q_{m_\alpha}\}$.

This is of the form

$$\mathbf{f} = A\mathbf{f}^\sigma + \mathbf{r}.$$

So we can apply Lemma 7.2 to conclude that the entries from $\mathbf{f}$ (the $f(q_i)$) are algebraic of degree at most $p^{m_\alpha^2+m_\alpha} - 1$ over $\mathbb{F}_q(t)$

Now we consider $q \in Q^\alpha_{\text{post}}$. Define $V_q = \{x \in S_p \cap [0,1) | \delta(q, s'(x)) \in F^\alpha\}$ and set $g(q) = \sum_{x \in V_q} t^x$. These $g(q)$ satisfy the following relations:

$$\begin{aligned} g^p(q) &= \sum_{\delta(q,d)=q'} t^d g(q') = \sum_{d=0}^{p-1} t^d g(\delta(q,d)) \text{ for } q \notin F^\alpha, \\ g^p(q) &= 1 + \sum_{\delta(q,d)=q'} t^d f(q') = 1 + \sum_{d=0}^{p-1} t^d g(\delta(q,d)) \text{ for } q \in F^\alpha. \end{aligned} \tag{4}$$

Note that $g(q) = \sum_{n \in V_q} t^n$ gives a finite expression when there are no lobes on all paths from q to a final state. Let l_q denote the maximum length of such paths. Then for these q, we have that $g(q) \in \mathbb{F}_p[t^{1/p^{l_q}}]$, so $g(q)$ is algebraic of degree p^{l_q}.

Define $l = \max_q\{l_q\}$ and let the set $\{q_1, ..., q_n\}$ denote the states with a lobe on a path from q_i to a final state, then we can rewrite the system of equations (4) to

$$g(q_i) = s(q_i) + \sum_{\delta(q_i,d)=q_j} t^d g^p(q_j),$$

for

$$\begin{aligned} s(q_i) &= \sum_{q \notin \{q_1,...,q_n\}, \delta(q,d)=q_i} t^d g^p(q) \in \mathbb{F}_q[t^{1/p^{l_{q_i}}}] \text{ for } q \notin F^\alpha, \\ s(q_i) &= 1 + \sum_{q \notin \{q_1,...,q_n\}, \delta(q,d)=q_i} t^d g^p(q) \in \mathbb{F}_q[t^{1/p^{l_{q_i}}}] \text{ for } q \in F^\alpha. \end{aligned}$$

This is of the following form:

$$\mathbf{g}^\sigma = B\mathbf{g} + \mathbf{s}.$$

So we again can apply Lemma 7.2 to conclude that the elements from g are algebraic of degree at most $p^{n_\alpha^2+n_\alpha} - 1$ over $\mathbb{F}_p[t^{1/q^l}]$. It follows that the entries of g are algebraic of degree at most $(p^{n_\alpha^2+n_\alpha} - 1)p^{l_\alpha}$ over $\mathbb{F}_q(t)$.

Let $\overrightarrow{q_r} \in Q_{\text{pre}}$ denote the state directly before the radix transition and let $\overleftarrow{q_r} \in Q_{\text{post}}$ denote the state directly after it (i.e.: $\delta(\overrightarrow{q_r}, .) = \overleftarrow{q_r}$.

So the algebraic degree of $\sum_{j \in S_\alpha} t^j = f(\overrightarrow{q_r})g(\overleftarrow{q_r})$ is at most

$$p^{m_\alpha^2+m_\alpha} p^{n_\alpha^2+n_\alpha} p^{l_\alpha},$$

(we do not substract 1, since when $m_\alpha = 0$, the algebraic degree of $f(\overrightarrow{q_r})$ is 1, similar for n_α).

We have an automaton for each S_α, for $\alpha \in \mathbb{F}_q$. So in case $q = 2$ we would be finished here. However, for $q \neq 2$, we have $q-1$ different automata M_α each with degree

$$p^{{m_\alpha}^2+m_\alpha+{n_\alpha}^2+n_\alpha+l_\alpha}.$$

So

$$\sum_i h((i-b)/a)t^i = \sum_{\alpha\in\mathbb{F}_q} \alpha\Big(\sum_{j\in S_\alpha} t^j\Big)$$

is algebraic of degree at most

$$\prod_{\alpha\in\mathbb{F}_q^*} p^{{m_\alpha}^2+m_\alpha+{n_\alpha}^2+n_\alpha+kl_\alpha}.$$

Note that we have that $\gcd(a,p) = 1$, since we chose a as small as possible: If however $p|a$, then

$$ai+b = \frac{m}{p^n} \Rightarrow \frac{ai}{p} + \lceil b/p \rceil = \frac{m-bp^n}{p^{n+1}} + \lceil b/p \rceil = \frac{m-bp^n+\lceil b/p\rceil p^{n+1}}{p^{n+1}} \in S_p.$$

So a was not minimal. Now by Lemma 7.3, we have obtained the bound in equation (2).

□

Example 7.7. In the simplest case: $q = 2$, $a = 1$, the bound from Theorem 7.6 is given by:

$$2^{m^2+m+n^2+n+l}.$$

For Example 5.5 we get $f(q_0) = 1$ and $g(q_1) = t^{1/2} + t^{3/4} + t^{7/8} + \ldots$, $g(q_2) = 1 + t^{1/2} + t^{3/4} + t^{7/8} + \ldots$. So the equations from (3) and (4) do hold. We have $m = 0, n = 2, l = 0$, so the bound is given by 2^6, much too high for the actual algebraic degree, which is algebraic of degree 2 (by Example 4.1).

Note. For a p-quasi-automatic generalized power series $\sum_i h(i)t^i$, generated by a DFAO with $R > 1$ radix transitions, we construct automata for each radix transition. Each one has one radix transition and preradix states such that the radix transition is reachable from them and postradix states reachable from the radix transition. When we take the sum of the corresponding generalized power series of these R DFAO's we obtain $\sum_i h(i)t^i$. This yields also the following corollary.

Corollary 7.8. *A p-quasi-automatic generalized power series $\sum_i h(i)t^i$ is algebraic of degree at most*

$$a \prod_{\alpha\in\mathbb{F}_q^*} (p^{{m_\alpha}^2+m_\alpha+{n_\alpha}^2+n_\alpha+l_\alpha})^{R_\alpha}, \tag{5}$$

where R_α is the number of (relevant) radix transitions of M_α.

Example 7.9. In Example 7.4, we have $a = 1$ since f is 3-automatic. Moreover, $m_1 = 0, n_1 = 1, l_1 = 1, R_1 = 2$ and $m_2 = 0, m_2 = 1, l_2 = 0, R_2 = 1$. So the bound from (5) is given by $(3^3)^2 \cdot 3^2 = 3^8$.

When we divide the DFAO into two DFAO's with each one radix transition, we get the bound $3^3 \cdot 3^0 \cdot 3^0 = 3^3$

Remark. We can not compare the bounds in this section with the bound from Theorem 3.4 as that bound is about reverse reading automata and we consider forward reading automata.

However, if we consider reverse reading automata, we can follow the same strategy as in the proof above and moreover we can use Theorem 3.4 to obtain a bound for the number of states of reverse reading automata. This bound will be even nicer:

$$a \prod_{\alpha \in \mathbb{F}_q^*} (p^{m'_\alpha} - 1) p^{{n_\alpha}^2 + n_\alpha + l_\alpha},$$

where m'_α is the total number of preradix (or, since we have a reverse reading DFA: postradix) states.

8 From algebraic to automatic

In this section we consider an algebraic generalized power series of degree k. Note that there are infinitely many (generalized) power series algebraic over $\mathbb{F}_q(t)$, but finitely many automata of a given size. So the algebraic degree will not give enough information for a bound on the size of the corresponding automaton. In [4, Theorem 14] this is solved by taking in account the degree of the coefficients, so we can expect that we will also need that here.

We need the following definition.

Definition 8.1. For $c \in \mathbb{Z}_{>0}$, call

$$T_c := \{n - \sum_{i=1}^{\infty} b_i/p^i | n \in \mathbb{Z}_{>0}, b_i \in \{0, 1, \ldots, p-1\}, \sum b_i \leq c\} \subset S_p$$

Note. (A subset of) T_c is well-ordered.

We will restrict ourselves to series supported on (a subset of) $\mathcal{T}_c := T_c \cap (0, 1]$. Since $\sum_{i=1}^{\infty} b_i/p^i < 1$, we just look at

$$\mathcal{T}_c = T_c \cap (0, 1] = \{1 - \sum_{i=1}^{\infty} b_i/p^i | b_i \in \{0, 1, \ldots, p-1\}, \sum b_i \leq c\} \subset S_p$$

Example 8.2. An easy case in $\mathbb{F}_2$ is given by $c = 1$:

$$\mathcal{T}_1 = \{1\} \cup \{1 - 1/2^n | n \in \mathbb{Z}_{>0}\} = \{1\} \cup \{(2^n - 1)/2^n | n \in \mathbb{Z}_{>0}\}.$$

When $c = 2$ it will be given by:

$$\begin{aligned} \mathcal{T}_2 &= \mathcal{T}_1 \cup \{1 - \sum_{i=1}^{\infty} b_i/p^i | b_i \in \{0, 1, \ldots, p-1\}, \sum b_i = 2\} \\ &= \{1\} \cup \{(2^n - 1)/2^n | n \in \mathbb{Z}_{>0}\} \cup \{(2^m - 2^n - 1)/2^m | m > n \in \mathbb{Z}_{>0}\}. \end{aligned}$$

Note. Consider the base p expansion of the elements of $\mathcal{T}_c$. Base p expansions of $\sum_{i=1}^{\infty} b_i/p^i$ look like $.b_1 b_2 b_3 \ldots b_{d-1} b_d$, which is a finite expression since $\sum b_i \leq c$. So $1 - \sum_{i=1}^{\infty} b_i/p^i$ corresponds to:

$$.\overline{(p - b_1 - 1)}\overline{(p - b_2 - 1)}\overline{(p - b_3 - 1)} \ldots \overline{(p - b_{d-1})}\overline{(p - b_d)},$$

where we work in modulo p, as the sum of these two base p expansions add up to 1.

When $p = 2$, the elements of $\mathcal{T}_c$ correspond to exactly the base 2 expansions with at most $c-1$ zeros. For example, the series $f_{a,b} = t^{1/2}+t^{3/4}+t^{7/8}+t^{15/16}+\ldots$ from Example 5.5 is supported on $\mathcal{T}_1$ and consists of all base 2 expansions without zeros.

We consider the following theorem: [7, Proposition 5.2.2]

Theorem 8.3. *A generalized power series $x = \sum_i x_i t^i \in \mathbb{F}_q((t^{\mathbb{Q}}))$ is algebraic over $\mathbb{F}_q((t))$ if and only if:*

- *$\exists a, b, c \geq 0$ such that $\sum_i x_{(i-b)/a} t^i$ is supported on T_c.*
- *For some (hence all) of such a, b, c, $\exists M, N$ such that every sequence*

$$c_n = x_{(m-b-\sum_{i=1}^{j-1} b_i p^{-i} - p^{-n} \sum_{i=j}^{\infty} b_i p^{-i})/a}, \tag{6}$$

with $m, j \in \mathbb{Z}_{\geq 0}$ and $b_i \in \{0, 1, \ldots, p-1\}$, such that $\sum b_i \leq c$, has period length dividing N after at most M terms.

For series supported on $\mathcal{T}_c = T_c \cap (0, 1]$, we then have the following: [8, Lemma 5.2.4]:

Lemma 8.4. *A generalized power series $x = \sum_i x_i t^i \in \mathbb{F}_q((t^{\mathbb{Q}}))$ with support in $\mathcal{T}_c$ for some $c \in \mathbb{Z}_{>0}$ and algebraic over $\mathbb{F}_q((t))$ is p-automatic. (Moreover, Proposition 7.1 then implies that it is algebraic over $\mathbb{F}_q(t)$.)*

Proof. Since x is supported on $\mathcal{T}_c$, we can take $a = 1, b = 0$ in 8.3. So every sequence

$$c_n = x_{1-\sum_{i=1}^{j-1} b_i p^{-i} - p^{-n} \sum_{i=j}^{\infty} b_i p^{-i}},$$

with $j \in \mathbb{Z}_{\geq 0}$ and $b_i \in \mathbb{F}_2$, such that $\sum b_i \leq c$, has period length dividing N after at most M terms.

We define an equivalence relation on S_p as follows: For $i, j \in S_p$, $i \sim j$ if and only if we can obtain the base q expansion of j by replacing each string of $M + u + vN$ consecutive zeros a string of $M + u + wN$ consecutive zeros, where $u, v, w \in \mathbb{Z}_{\geq \not\sim}$. Note that this is indeed a proper equivalence relation. Each equivalence class has a shortest element: the one where no string of $M + N$ consecutive zeros occurs. (If there is such a string, we can replace it by M consecutive zeros and we get a shorter element in the same equivalence class). Notice that when $i \sim j$, then $x_{1-i} = x_{1-j}$, since c_n is periodic of length M after N terms. Now, by Theorem 2.10, the function $f : S_p \to \mathbb{F}_q$ given by $f(i) = x_{1-i}$ is q-automatic.

This also means that $g(i) = x_i$ is q-automatic: use a finite state transducer. [1, Theorem 6.8.6] □

We now consider a family of functions $\{g_c\}$, constructed as follows: In $\mathbb{F}_2((t^{\mathbb{Q}}))$ consider g_c, generated by an automaton accepting every string $.s_1 s_2 s_3 \ldots$ with exactly $c - 1$ zeros. So the automaton in Example 5.5 corresponds to g_1 as it accepts strings of the form $.1^*$.

Proposition 8.5. *g_c is algebraic of degree 2^c.*

Proof. We prove this using induction: For $c = 1$ we have $g_1^2 + tg_1 + t = 0$. Now suppose it hold for $c = d$, then we have for $\epsilon \neq s \in \{0,1\}^*$:

$$\begin{aligned}\sup(g_{d+1}) &= \{.s | s \text{ has exactly } d \text{ zeros}\} \\ &= \{.0s | s \text{ has exactly } d-1 \text{ zeros}\} \cup \{.1s | s \text{ has exactly } d \text{ zeros}\}.\end{aligned}$$

So

$$\begin{aligned}\sup(g_{d+1}^2) &= \{.s | s \text{ has exactly } d-1 \text{ zeros}\} \cup \{1.s | s \text{ has exactly } d \text{ zeros}\} \\ &= \sup(g_d) \cup \sup(tg_{d+1}).\end{aligned}$$

Which gives us the following relation

$$g_{d+1}^2 + tg_{d+1} + g_d = 0, \tag{7}$$

Since g_d was algebraic of degree 2^d, g_{d+1} is now algebraic of degree (at most) 2^{d+1}.

Let g_d be a zero of

$$a_d^{(d)}(t)x^{2^d} + a_d^{(d-1)}(t)x^{2^{d-1}} + \cdots + a_d^{(1)}(t)x^2 + a_d^{(0)}(t)x + a_d(t),$$

with $a_d(t), a_d^{(0)}(t), ..., a_d^{(d)}(t) \in \mathbb{F}_2(t)$. Then using the relation from equation 7 in the first and last equality sign gives

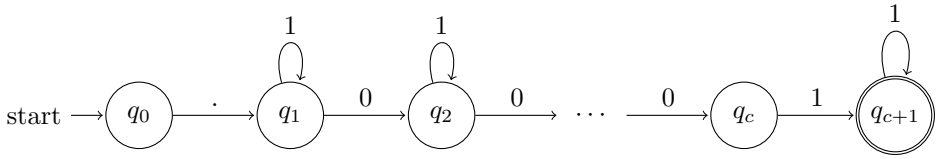
$$\begin{aligned}a_d^{(d)}(g_{d+1}^{2^{d+1}} + t^{2^d} g_{d+1}^{2^d}) = a_d^{(d)} g_d^{2^d} = a_d^{(d-1)} g_d^{2^{d-1}} + \cdots + a_d^{(1)} g_d^2 + a_d^{(0)} g_d + a_d \\ = a_d^{(d-1)}(g_{d+1}^{2^d} + t^{2^{d-1}} g_{c+1}^{2^{d-1}}) + \cdots + a_d^{(1)}(g_{d+1}^4 + t^2 g_{d+1}^2) + a_d^{(0)}(g_{d+1}^2 + tg_{d+1}) + a_d,\end{aligned}$$

So the coefficients of a polynomial, where g_{d+1} is a zero, will look like

$$\begin{aligned}a_{d+1}^{(d+1)} &= a_d^{(d)} \\ a_{d+1}^{(d)} &= a_d^{(d)} t^{2^d} + a_d^{(d-1)} \\ a_{d+1}^{(d-1)} &= a_d^{(d-1)} t^{2^{d-1}} + a_d^{(d-2)} \\ &\vdots \\ a_{d+1}^{(1)} &= a_d^{(1)} t^2 + a_d^{(0)} \\ a_{d+1}^{(0)} &= a_d^{(0)} t \\ a_{d+1} &= a_d\end{aligned}$$

Since We have $a_1^{(1)}(t) = 1$ and $a_1(t) = t$, we conclude that $a_c^{(c)}(t) = 1$ and $a_c(t) = t$ for all c. So the polynomial is Eisenstein, which proves the proposition. □

In general, an automaton corresponding to g_c will look like:



So the automaton has $c+2$ (essential) states.

The family of functions g_c shows that the bound on the algebraic degree, given the number of states is at least exponential. (See the previous section)

To get a bound on the number of states of M_g for a generalized power series g algebraic over $\mathbb{F}_q(t)$ we need to know:

- M, N and c as in Theorem 8.3,
- The number of equivalence classes (from the proof above): this will give the maximal number of states of M_f (by Theorem 2.10). This bound will depend on $M+N$ and c.
- The number of states of an automaton M_g accepting all strings s, given an automaton M_f accepting all strings $1-s$.

The first problem is to find M, N. The answer for this problem must be sought in [7]. It corresponds to being 'twist-recurrent' of some order.

Definition 8.6. A sequence c_n as in Theorem 8.3 is called *twist-recurrent* of order k if there exist $d_0, ..., d_k \in \mathbb{F}_p$ such that:

$$d_0 c_n + d_1 c_{n+1}^p + \cdots + d_k c_{n+k}^{p^k},$$

for all $n \geq 0$.

Note. When c_n is twist-recurrent of order k, it has period at most $p^k - 1$ after at most k terms. When all sequences c_n from (6) are twist-recurrent of order k, we get that $N \leq (p^k-1)!$ and $M \leq k$.

The second problem will give a high bound: even if $q=2$, the bound on the number of states will be:

$$\frac{M+N}{M+N-1}\big((M+N)^c - 1\big) + (M+N)^c = \frac{M+N}{M+N-1}\big((2M+2N-1)(M+N)^{c-1} - 1\big).$$

Since we consider the strings with at most c ones and at most $M+N$ consecutive zeros (where the strings with c ones do not end on 0).

The third problem raises the bound yet even more. For this we have to look into the proof of [1, 4.3.6].

This procedure only gives a bound on the number of states for generalized power series supported on $\mathcal{T}_c$. For any other generalized power series $y \in \mathbb{F}_q((t^{\mathbb{Q}}))$, we can write y as an $\mathbb{F}_q((t))$-linear ination of generalized power series supported on $\mathcal{T}_c$. This is used in the proof of [8, Proposition 5.2.7]. Finding (the size of) a basis of this is a step that will increase the bound a lot. Moreover Lemma 7.2 is needed and also will increase the bound. The proof also uses Christol's theorem (from algebraic to automatic) since it was linear over $\mathbb{F}_q((t))$. So a forward reading version of Theorem 3.5 has to be used, for which we do not know a better option than combining Theorem 3.5 and Lemma 2.6. Besides having to know the algebraic degree of these different generalized power series, another problem involved here is finding the degrees of the coefficients

of the polynomial in $\mathbb{F}_q(t)[T]$ Combining all steps seems to give a complicated bound that probably is not very tight.

It is difficult to find examples of generalized power series with low algebraic degree and the corresponding automaton having many states. The example below is one of the better ones we could find.

Example 8.7. Consider

$$f = 1 + \sum_{i \in \{1,2,3,5\} \bmod 7} t^{(2^i-1)/2^i}.$$

The support is thus given by

$$\sup(f) = \{0\} \cup \{(2^i - 1)/2^i | i = 1, 2, 3, 5 \bmod 7\},$$

so

$$\begin{aligned}
\sup(f^2) &= \{0, 1\} \cup \{2(2^i - 1)/2^i | i = 2, 4, 6, 3 \bmod 7\}, \\
\sup(f^4) &= \{0, 2, 3\} \cup \{4(2^i - 1)/2^i | i = 4, 1, 5, 6 \bmod 7\}, \\
\sup(f^8) &= \{0, 4, 6, 7\} \cup \{8(2^i - 1)/2^i | i = 1, 2, 3, 5 \bmod 7\}.
\end{aligned}$$

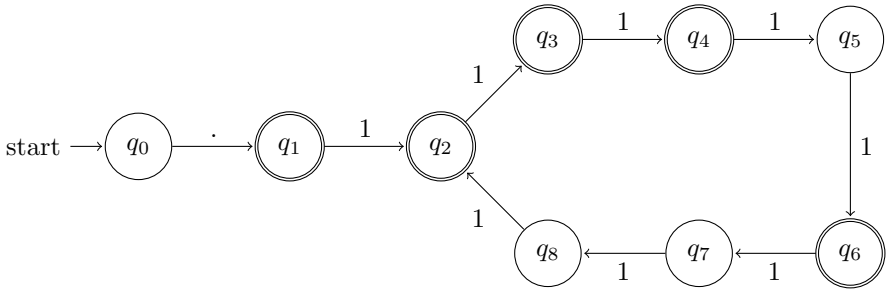
Notice that

$$\sup(f^8) + \sup(t^4 f^4) + \sup(t^7 f) = \{0, 7\}.$$

So f is a root of

$$x^8 + t^4 x^4 + t^7 x + t^7 + 1.$$

The corresponding automaton has at least 9 states:



9 Conclusion

For generalized power series we looked at the algebraic degree over $\mathbb{F}_q(t)$ and the size of the finite automaton accepting the exponents. As was done before for regular power series, we can obtain bounds from the generalized version of Christol's theorem. We found a bound on the algebraic degree of a generalized power series given the automaton generating the support. This result is stated in Theorem 7.6 and Corollary 7.8.

We pointed out the steps in finding a bound for the other direction: obtaining a bound on the number of states given the algebraic degree. A main step is using Lemma 8.4 about generalized power series with restraints on the support. The bound on the number of states for an automaton corresponding to these specific generalized power series will already be high. For all generalized power series the bound will be even worse, as Christol's theorem is again used in the proof of Kedlaya's theorem. This means that we can expect that the bound is a lot worse than the bound for regular power series

References

[1] Jean-Paul Allouche and Jeffrey Shallit. *Automatic sequences: theory, applications, generalizations*. Cambridge university press, 2003.

[2] Claude Chevalley. *Introduction to the theory of algebraic functions of one variable*. Number 6. American Mathematical Soc., 1951.

[3] Gilles Christol. Ensembles presque périodiques k-reconnaissables. *Theoretical Computer Science*, 9(1):141–145, 1979.

[4] Anneroos Everts. From finite automata to power series and back again. Master's thesis, Rijksuniversiteit Groningen, 2012.

[5] Hans Hahn. Über die nichtarchimedischen größensysteme. In *Hans Hahn Gesammelte Abhandlungen Band 1/Hans Hahn Collected Works Volume 1*, pages 445–499. Springer, 1995.

[6] Kiran S Kedlaya. Power series and p-adic algebraic closures. *arXiv preprint math/9906030*, 1999.

[7] Kiran S Kedlaya. Algebraic generalized power series and automata. *arXiv preprint math/0110089*, 2001.

[8] Kiran S Kedlaya. Finite automata and algebraic extensions of function fields. *Journal de théorie des nombres de Bordeaux*, 18(2):379–420, 2006.

[9] Donald S Passman. *The algebraic structure of group rings*. Courier Corporation, 2011.