Notting but structure

A search for sparse elements of finite compositional order in the Nottingham group

Master Thesis Mathematical Sciences
August 2022

Author: Mieke Wessel

Supervisor: Prof. Dr. Gunther Cornelissen
Second reader: Dr. Stefano Marseglia

Universiteit Utrecht

Contents

Introduction **3**

1 Prerequisites **4**

1.1 Nottingham group, p-automata and sparseness 4

1.2 Digraphs and linear algebra 9

1.3 Previous work 11

2 Automatic sparseness **12**

2.1 Identifying sparseness by the automaton 12

2.2 Determining the rank of sparseness 16

2.3 Bounds on the number of walks 17

2.4 An upper bound for non-sparse series 21

2.5 Non-sparse series of finite order 23

3 The structure of the support **26**

3.1 Structural lemmas and a greedy algorithm 26

3.2 Sparse series of order 2 33

3.3 Support relations 35

3.4 Generalisations to $p > 2$ 42

4 The power of a Galois group **47**

4.1 Field theoretic conditions for sparseness 47

4.2 Newton polygons 50

5 Further research **54**

Bibliography **55**

Appendix **57**

Conventions and notation

- By $\mathbb{N} = \{1, 2, \ldots\}$ we denote the set of natural numbers, i.e., all positive integers, not including zero.
- For any $N \in \mathbb{N}$ we write $[N]$ for the set $\{1, \ldots, N\} \subset \mathbb{N}$.
- For any set $S \subset \mathbb{R}$ and $a, b \in \mathbb{R}$ we write $aS + b$ for the set $\{as + b \mid s \in S\}$.
- When an index is omitted, we indicate this by writing a hat above the omitted part of the expression. For example, $a_0 + \ldots + a_{i-1}X^{i-1} + a_{i+1}X^{i+1} + \ldots + a_nX^n$ can then be written as $a_0 + ..\widehat{a_iX^i}.. + a_nX^n$.
- If $\sigma(t) : X \to X$ is a function, we write $\sigma(t)^n$ for the n-fold product of $\sigma(t)$ with itself and $\sigma(t)^{\circ n}$ for the n-fold composition of $\sigma(t)$ with itself. So $\sigma(t)^2 = \sigma(t)\sigma(t)$ and $\sigma(t)^{\circ 2} = \sigma(\sigma(t))$.
- Let P be a prime of some ring R and $x \in R$. We write $v_P(x)$ for the P-adic valuation of x. For example, if $R = \mathbb{F}_2[\![t]\!]$, $P = (t)$ and $x = t^5 + t^{10} + t^{15} + \ldots$, then $v_P(x) = 5$.

Introduction

The Nottingham group $\mathscr{N}_p$ was first defined in 1988 by D. Johnson and I. York [20], who both resided in Nottingham at the time. The combination of it being easy to calculate with and having interesting properties, has led to it becoming popular amongst group theorists [11]. It has been used, for instance, to refute conjectures about profinite groups [18]. Furthermore, number theorists have also taken an interest in this group, because it occurs as wild ramification group of $\mathbb{F}_p((t))$. In 1997 these two interpretations were brought together by the proof and implications of the following theorem [26]:

Theorem 0.1 (Camina [10])**.** *Every countably based pro-p group can be embedded, as a closed subgroup, in the Nottingham group.*

A consequence of this theorem is that every finite p-group must be a subgroup of the Nottingham group. However, so far not many explicit examples of elements of finite order have been found [11]. The goal of this thesis is to find more examples and then especially examples that are sparse. This property interests us because it forces the elements to have a relatively straightforward structure, which might be possible to generalise. Besides, explicit elements of the Nottingham group are also relevant in deformation theory (e.g. [3, p. 212]) and since sparse elements are relatively easy to implement in a computer program, this gives an extra motivation to look for such elements. An important invariant of elements of the Nottingham group is their depth. The only sparse series of finite order that were known before this thesis are of order 2 or 4 and have depths $2^\mu \pm 1$ [8] or 1 [4, 8], respectively.

In Chapter 1 all relevant definitions and propositions about the Nottingham group, sparseness and automata are given and previous work is reviewed. Chapter 2 is devoted to giving a new proof of Cobham's Theorem in the language of digraphs. The theorem concerns a characterisation of automatic sparse elements and the new proof can be used to determine a growth constant in both the sparse and non-sparse case. Chapter 3 examines if the finite order condition imposes certain structural conditions on the elements. Especially when the order is 2 this turns out to be the case. These conditions are then used to make smart guesses for other series of order 2 that might be sparse, which works for depths $2^\mu - 3$. The other guesses are not all sparse, as is proved in Chapter 4, based on a Galois theoretic condition for sparseness combined with Newton polygons.

Chapter 1

Prerequisites

In this chapter the prerequisites to understand the rest of the thesis are given. The first section contains all specific terminology and some examples. The second section focuses on the basics of algebraic graph theory and is only relevant to follow the proofs in Chapter 2. It can, therefore, be skipped if the reader is either already familiar with the subject or not as interested in the specifics. The third section gives a more extensive recap of previously known results than was given in the introduction.

1.1 Nottingham group, p-automata and sparseness

Definition 1.1. Let p be a prime number. The *Nottingham group* $\mathscr{N}_p$ is the subgroup of $\mathrm{Aut}(\mathbb{F}_p[\![t]\!])$ consisting of the elements σ such that $\sigma(t) = t + O(t^2)$. The group operation is composition.

Remark. Note that an element of $\mathrm{Aut}(\mathbb{F}_p[\![t]\!])$ is completely determined by the image of t. Thus the same holds for $\mathscr{N}_p$.

That $\mathscr{N}_p$ meets all the group axioms follows quite straightforwardly from the fact that $\mathrm{Aut}(\mathbb{F}_p[\![t]\!])$ is a group. However, for intuitive understanding, it is still useful to calculate some examples of inverses and compositions.

Examples. Note that the identity of $\mathscr{N}_p$ is t. All examples are in $\mathscr{N}_2$.

- Define $\sigma_{K,1}(t) := t + t^2 + t^3 + t^4 + \ldots$. Then, $\sigma_{K,1}$ is its own inverse. This can be deduced by calculating $\sigma_{K,1}(t + t^2 + \ldots + t^n) = t + t^{n+1} + O(t^{n+2})$ for each $n \in \mathbb{N}$. Alternatively a more combinatorial proof can be given. We may write $\sigma_{K,1}(\sigma_{K,1}(t)) = a_1 t + a_2 t^2 + a_3 t^3 + \ldots$ and note that a_i equals the number of compositions of i. For example we get $a_3 = 4$, because $3 = 1 + 1 + 1 = 1 + 2 = 2 + 1 = 3$. Hence, $a_i = 2^{i-1}$, which is zero modulo 2 for all $i > 1$ and $a_1 = 1$.
- Define $\tau(t) := t + t^2 + t^4 + t^8 + \ldots$. Then, $\tau^{-1}(t) = t + t^2$ is its inverse. We can calculate this by first taking $\tau^{-1}(t) = t + O(t^2)$. Since $\tau(t) = t + t^2 + \ldots$, we see that the next term of τ^{-1} should be t^2. Using that we are in characteristic 2 we find $(t + t^2)^{2^n} = t^{2^n} + t^{2^{n+1}}$ and thus, $\tau(t + t^2) = t + 2t^2 + 2t^4 + 2t^8 + \ldots = t$.

- We can also calculate some compositions of these elements.

  $$\sigma_{K,1}(\tau^{-1}(t)) = t + 2t^2 + 3t^3 + 5t^4 + 8t^5 + 13t^6 + 21t^7 + \ldots = t + t^3 + t^4 + t^6 + t^7 + \ldots,$$
  $$\tau^{-1}(\sigma_{K,1}(t)) = t + 2t^2 + t^3 + 2t^4 + t^5 + 2t^6 + t^7 + \ldots = t + t^3 + t^5 + t^7 + \ldots.$$

  From this we conclude that $\mathscr{N}_2$ is non-abelian.

Definition 1.2. We say a power series in $\mathbb{F}_p[\![t]\!]$ has *finite compositional order* N if $N \in \mathbb{N}$ is the smallest integer such that $\sigma^{\circ N}(t) = t$.

In previous example we saw that $\sigma_{K,1}$ was its own inverse. Therefore, the finite compositional order of $\sigma_{K,1}$ is equal to 2. Any conjugate of $\sigma_{K,1}$, such as $\tau \circ \sigma_{K,1} \circ \tau^{-1}$, also has order 2, but these are (in general) quite cumbersome to calculate. However, the next proposition shows that any power series of finite compositional order must be in $\mathscr{N}_p$ and have order some power of p.

Proposition 1.3. *Let $\sigma(t) = a_0 + a_1 t + a_2 t^2 + \ldots \in \mathbb{F}_p[\![t]\!]$ be of finite compositional order $N > 1$. Then $a_0 = 0$, $a_1 = 1$, σ has infinitely many terms and $N = p^r$ for some $r \in \mathbb{N}$.*

Proof. That a_0 must equal 0 holds because otherwise composition of power series is not well defined. Infinitely many terms are needed, because if $\sigma(t)$ were a polynomial of degree $n > 1$, the degree of $\sigma^{\circ N}(t)$ would be n^N instead of 1. The fact that $a_1 = 1$ and $N = p^r$ can be proved by analysing the coefficients of $\sigma^{\circ N}$, as follows.
Suppose $\sigma(t) = a_1 t + a_{d-1} t^{d-1} + O(t^d)$ with $a_{d-1} \neq 0$. It is easy to calculate that $\sigma(t)^{\circ N} = a_1^N t + N a_1^{N-1} a_{d-1} t^{d-1} + O(t^d)$. This must equal t, so p must divide N. Now suppose that $N = q \cdot p^r$ with $q \in \mathbb{N}$ and coprime to p. Then the series $\sigma(t)^{\circ p^r}$ has order q, which tells us that it is either the identity or p divides q. We assumed q to be coprime to p and may conclude $q = 1$ and N is a power of p. We also need $a_1^N = a_1^{p^r} \equiv 1 \mod p$, which by Fermat's little theorem implies $a_1 = 1$. □

Another consequence of the proposition is that no polynomial has finite compositional order. Therefore τ^{-1}, as defined in previous example, has infinite order. The same must then hold for τ, even though this element does have infinitely many terms.

One way to construct elements of the Nottingham group with a lot of structure is by using p-automata.

Definition 1.4. Let p be a prime. A *p-automaton* is a finite directed multigraph for which the following hold:

- each vertex is labelled by an element of $\mathbb{F}_p$.
- one vertex is additionally labelled with 'Start'.
- each vertex has p outgoing edges, that are all labelled with a different element of $\{0, 1, \ldots, p-1\}$.
- an edge labelled with 0 connects two vertices with the same label.
- each vertex can be reached from the start vertex. (This is called *accessibility.*)

From a p-automaton we can, in a natural way, construct a sequence. If for a certain sequence there exists such a p-automaton we call that sequence automatic. This is made precise in the following definition.

Definition 1.5. Let $k \in \mathbb{Z}$, p a prime and $(b_i)_{i \geq 0}$ a (finite) base-p expansion of k, in symbols $k = \sum_{i=0}^{n} b_i p^i$. Suppose we are given a p-automaton. From the 'Start' vertex consecutively follow the edges labelled by $b_0, b_1, \ldots, b_n$ and suppose we end up at vertex v. Define a_k to equal the label of v in $\mathbb{F}_p$. Then $(a_i)_{i \geq 0}$ is called an *automatic sequence*.

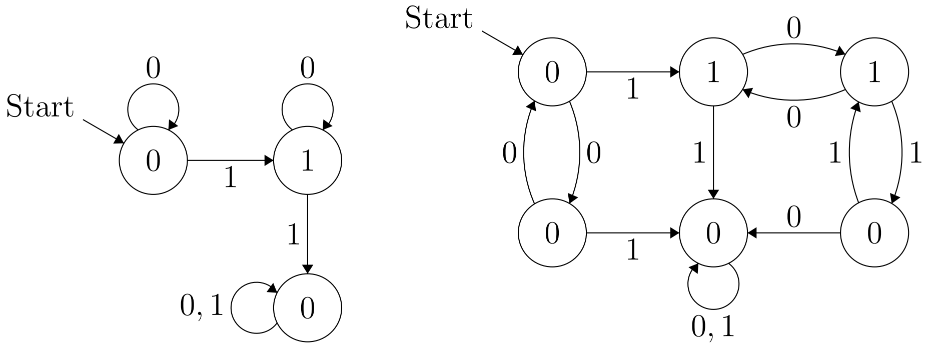
Remark. The fourth property of a p-automaton implies that an automatic sequence does not depend on which base-p expansion one takes for an integer. For instance writing $5 = 1 \cdot 1 + 0 \cdot 2 + 1 \cdot 4$ will give the same value for a_5 as writing $5 = 1 \cdot 1 + 0 \cdot 2 + 1 \cdot 4 + 0 \cdot 8$. This property is therefore also called *leading zero invariance*.

Definition 1.6. An *automatic series* is a formal power series $\sigma(t) \in \mathbb{F}_p[\![t]\!]$ such that for some automatic sequence $(a_i)_{i \geq 0}$ in $\mathbb{F}_p$ we have:

$$\sigma(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \ldots.$$

Examples. We use the two 2-automata in Figure 1.1 and write $(a_i)_{i \in \mathbb{Z}_{\geq 0}}$ and $(b_i)_{i \in \mathbb{Z}_{\geq 0}}$ respectively for the automatic sequences generated by (a) and (b).

- First of all one should check that both automata meet all the requirements of Definition 1.5.
- Suppose we are interested in calculating a_{13} and b_{13}. In binary $13 = 1101$, so we must follow the arrows labelled by $1, 0, 1$ and 1 consecutively. In (a) we end up in the vertex at the bottom labelled with 0 and in (b) in the top right vertex labelled with 1. Hence, $a_{13} = 0$ and $b_{13} = 1$.
- Using this method for all $i \in \mathbb{Z}_{\geq 0}$ we get $(a_i) = (0, 1, 1, 0, 1, 0, 0, 0, 1, \ldots)$ and $(b_i) = (0, 1, 0, 0, 1, 0, 0, 0, 0, \ldots)$. This leads to the series $t + t^2 + t^4 + t^8 + \ldots$ and $t + t^4 + t^{13} + \ldots$, respectively.
- In fact, we can see directly from the automaton in (a) that $a_i = 1$ if and only if i is a power of 2; the only paths that lead to a non-zero labelled vertex are the ones with edges labeled by $0, 0, \ldots, 0, 1(, 0, \ldots, 0)$, here the zeroes between the brackets account for the leading zero invariance. A similarly simple description of (b) is not possible, but we will see later on that it generates a so-called Klopsch's series.



(a)

(b)

Figure 1.1: Automata for the series (a) $\tau(t) = t + t^2 + t^4 + \ldots$ and (b) $\sigma_{K,3}(t)$ as given in [9, Figure 1].

The amount of structure automatic series contain is best described by a theorem of G. Christol stating that a series is automatic if and only if it is algebraic, as follows.

Definition 1.7. Let p be a prime and $\sigma(t) \in \mathbb{F}_p[\![t]\!]$. We say that σ is *algebraic* if it satisfies the polynomial equation $F(t, X) = 0$ for some $F(t, X) \in \mathbb{F}_p[t, X]$.

Examples. All series we have seen so far are actually algebraic. The following examples are all taken from $\mathscr{N}_2$.

- All polynomials $p(t)$ are algebraic since they satisfy $X - p(t) = 0$.
- The series $\sigma_{K,1}(t) = t + t^2 + t^3 + \ldots$ is algebraic since it satisfies $(t-1)X + t = 0$.
- The series $\tau^{-1}(t) = t + t^2 + t^4 + \ldots$ is algebraic since it satisfies $X^2 - X + t = 0$.

Theorem 1.8 (Christol [13])**.** *A power series in $\mathscr{N}_p$ is automatic if and only if it is algebraic.*

The proof of Christol's theorem is constructive in the sense that given an algebraic equation one can construct the automaton and vice versa. This is also how Figure 1.1b was determined.

As one can imagine, having finite compositional order also requires some structure. In [4, Remark 1.6] it was pointed out that every series of finite order is in fact automatic (and thus algebraic). Hence, in our search for elements of finite compositional order, we only have to consider automatic series.

Another interesting property elements of the Nottingham group can possess is sparseness. It tells us something about the extent of non-zero coefficients and is based on the support of an element.

Definition 1.9. Let $\sigma(t)$ be a power series in $\mathbb{F}_p[\![t]\!]$ for some prime p, then its *support* is defined to be the set

$$E(\sigma) = \{k \in \mathbb{Z} \mid a_k \neq 0\}.$$

Furthermore, we write $E(\sigma)_N$ for the set $E(\sigma) \cap [N]$.

Definition 1.10. A power series $\sigma(t)$ in $\mathbb{F}_p[\![t]\!]$ is (r-)*sparse* if

$$\#E(\sigma)_N = O(\log(N)^r)$$

for some $r \geq 0$. The infimum of such r is called the *rank of sparseness of* σ. In the case that $\sigma(t)$ is automatic, the corresponding p-automaton and automatic sequence are also called (r)-sparse.

Examples. Let $\sigma_{K,1}$ and τ be as defined in previous examples.

- Any polynomial has a finite support and is therefore sparse of rank 0.
- The support of $\sigma_{K,1}$ is $E(\sigma_{K,1}) = \mathbb{N}$ and therefore $\#E(\sigma_{K,1})_N = N$. Thus, $\sigma_{K,1}$ is not sparse.
- The support of τ is $E(\tau) = \{1, 2, 4, 8, \ldots\}$ and therefore

  $$\#E(\tau)_N = \#\{1, 2, \ldots, 2^{\lfloor \log_2(N) \rfloor}\} = O(\log(N)).$$

  Thus, τ is sparse of rank 1.

In previous examples it was quite easy to recognise whether a series was sparse or not. This might not always be the case, for instance if the support is a union of several sets all growing polylogarithmically. The following definition and proposition give one way to make this distinction and will be of relevance in the automatic sparse case. Both rely partially on Definitions 1.17 and 1.18 of the next section.

Definition 1.11. Let r be a non-negative integer and $v_0, \dots, v_r, w_0, \dots, w_r$ some p-ary words, with all w_i non-trivial. Then their *simple sparse set* is the set

$$\{v_r w_r^{x_r} \cdots v_1 w_1^{x_1} v_0 \mid x_i \in \mathbb{Z}_{\geq 0}\} \subset \mathbb{Z}_{\geq 0}.$$

If either $|v_r| \neq 0$ or $|w_r| \neq 0$, we say the simple sparse set has *rank r*.

Remark. The previous definition is not universal in the sense that different sources use different terminology. For a brief discussion see [9, p. 14].

Proposition 1.12. *Let E be a simple sparse set of rank r. Then there exists some constant $C > 0$ such that for sufficiently large N the following holds:*

$$\#(E \cap [N]) \geq C \log(N)^r.$$

Proof. By definition we can write E as $\{v_r w_r^{x_r} \cdots v_1 w_1^{x_1} v_0 \mid x_i \in \mathbb{Z}_{\geq 0}\}$ for some p-ary words $v_0, \dots, v_r, w_1, \dots, w_r$, where w_i is non-trivial for all $1 \leq i \leq r$. Write z for the sum of the lengths of $v_0, \dots, v_r$ and l_i for the length of w_i, and define $L := \max(l_i)$.

For all $N \in \mathbb{N}$ there exists an integer $k \geq 0$ such that $p^{Lk+z} \leq N < p^{L(k+1)+z}$. The number of elements in $E \cap [N]$ is then larger than or equal to the number of elements in $E \cap [p^{Lk+z}]$, so it will suffice to examine this number.
Let $(x_0, \dots, x_r) \in \mathbb{Z}_{\geq 0}^{r+1}$ so that $x_0 + \ldots + x_r = k$. Then $v_r w_r^{x_r} \cdots v_1 w_1^{x_1} v_0$ has length less than or equal to $Lk + z$ and thus, has size less than p^{Lk+z}. Note that each solution $(x_0, \dots, x_r)$ leads to a unique size of $v_r w_r^{x^r} \cdots v_1 w_1^{x_1} v_0$, so the number of such solutions gives a lower bound. There are exactly $\binom{k+r}{r}$ of such solutions and we find

$$\#(E \cap [N]) \geq \#(E \cap [p^{Lk+z}]) \geq \binom{k+r}{r} \geq \frac{1}{r!} k^r.$$

We also know that $L(k+1) + z > \log_p(N)$, so $k > \frac{1}{L}(\log_p(N) - z - L)$. Since r, L, z and p are constants we quickly see that, for sufficiently large N, we can find some constant C such that

$$\#(E \cap [N]) \geq C \log(N)^r. \qquad \square$$

From this proposition one may conclude that, if the support of a series is equal to a finite union of simple sparse sets, the series is sparse. It should be remarked, though, that the converse does not hold: there are certainly sparse series whose supports are not equal to a union of simple sparse sets. In fact, sparseness does not force a series to have any kind of predictable structure at all, but this changes when we add being automatic (or equivalently, algebraic) to the mix.

Proposition 1.13 (Szilard, Yu, Zhang and Shallit [28], [9] Cor. 3.10)**.** *A series σ is automatic and r-sparse if and only if its support is a finite union of pairwise disjoint simple sparse sets of rank at most r.*

In Chapter 2 we will see a proof of this proposition.

Remark. In the automatic case we can now give a more quantitative description of what it means for a series to be 'easy to implement'. When the simple sparse sets of a series are determined it takes $O(\log(N))$ time to calculate its first N coefficients. Whereas for a non-sparse automatic series this would take roughly $O(N \log N)$ time.

Additionally, sparseness also imposes a condition on the automaton. By a theorem of A. Cobham any automatic series is either sparse or it support grows fractionally in N. A. Szilard, S. Yu, K. Zhang and J. Shallit showed that the sparse case holds if and only if the automaton contains a tied vertex.

Definition 1.14. Let v be a vertex of some p-automaton. Then v is called *tied* if the following two properties hold:

(i) there exists a (possibly empty) walk from v to a vertex with a label other than 0.

(ii) there exist two different walks of the same length from v to itself.

Theorem 1.15 (Cobham [14], [28]). *An automatic series $\sigma(t) \in \mathbb{F}_p[\![t]\!]$ is sparse if and only if a corresponding automaton does not contain any tied vertices. Furthermore, if it is not sparse $\#E_N(\sigma) \geq N^\alpha$ for some $\alpha \in (0,1]$ and sufficiently large N.*

This theorem will also be proved in Chapter 2.

1.2 Digraphs and linear algebra

We will need some more general (algebraic) graph theory. Here we will renew the necessary definitions.

Definition 1.16. Let $G = (V, E)$ be a graph. Then a *walk* W consists of vertices $v_0, v_1, \ldots, v_n$ in V and edges $e_1, e_2, \ldots, e_n$ in E so that e_i is an edge between v_{i-1} and v_i for each i. If we also require that all vertices are distinct, W is called a *path*.

Definition 1.17. Let G be a graph and W some walk in G consisting of n edges. Then we define the *length* $l(W)$ *of* W as n.

Definition 1.18. Let W be a walk in some p-automaton, such that the edges of W are consecutively labelled by $b_1, b_2, \ldots, b_n$. Then we define the *size* $|W|$ *of* W by

$$b_1 + b_2 p + b_3 p^2 + \ldots + b_n p^{n-1}.$$

Examples. The walk with v_0 equal to the start-vertex that we use to determine a_m has size m and length $\lfloor \log_p(m) + 1 \rfloor$.
For each vertex v in a p-automaton and each integer m there is exactly one walk of size m that does not end in a zero-labelled edge and starts at v.

Definition 1.19. Let $D = (V, E)$ be a digraph (i.e. directed graph). We call D a *rooted out-tree* or an *arborescence* if there is a vertex v in V such that for each vertex $w \in V$ there is exactly one walk from v to w. The vertex v is called *the root* of D and if w has no outgoing edges, it is called a *leaf.*

Definition 1.20. Let $D = (V, E)$ be an arborescence. Then *the height* of D is the maximal length of a walk from the root v to some vertex $w \in V$.

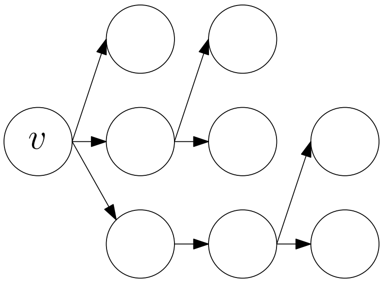


Figure 1.2: An example of an arborescence with height 3, root v and 5 leaves.

Definition 1.21. Let $D = (V, E)$ be a digraph with n vertices and suppose we have labelled the vertices by $v_1, v_2, \ldots, v_n$. For $1 \leq i, j \leq n$ define $a_{i,j}$ to be the number of edges in E going from v_i to v_j. The square matrix $A_D := (a_{i,j})$ is called the *adjacency matrix of D.*

The advantage of considering the adjacency matrix is that we can now use a bit of linear algebra to analyse digraphs. In the following we will assume all matrices to be taken over $\mathbb{R}$.

Definition 1.22. Let A be a matrix. We say A is *positive* if all entries are positive and that A is *non-negative* if all entries are non-negative. The definition is analogous for vectors.

Definition 1.23. Let A be a square matrix. The set of distinct eigenvalues is called the *spectrum* of A and denoted by $\operatorname{Spec}(A)$. The number

$$\rho(A) := \max_{\lambda \in \operatorname{Spec}(A)} |\lambda|$$

is the *spectral radius* of A. Here $|.|$ is the usual absolute value over $\mathbb{C}$.

Definition 1.24. Let A be a square matrix. We say A is a *reducible matrix* if there exists a permutation matrix P so that

$$PAP^{-1} = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix},$$

where X, Y, Z and 0 are non-trivial matrices and X and Z are square.
If A does not meet this condition, it is called an *irreducible matrix.*

Definition 1.25. A digraph $D = (V, E)$ is called *strongly connected* if for each pair of vertices u, v there exits a path from u to v. It is called *weakly connected* if there exists a path from u to v or from v to u.

Lemma 1.26. *A digraph D is strongly connected if and only of its adjacency matrix A_D is irreducible.*

Proof. Assume that A_D is reducible and that the vertices have been permuted such that $A_D = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$. Suppose X and Z have dimensions $m_x \times m_x$ and $m_z \times m_z$, receptively, and note that $m_x + m_z = n$. The zero-block in A_D tells us that there does not exist an edge from v_i to v_j for any $m_x + 1 \leq i \leq n$ and $1 \leq j \leq m_x$. Therefore, there cannot exist a path from

v_n to v_1 and D is not strongly connected.
Now suppose D is not strongly connected. There must be vertices v and u so that there is no path from v to u. Define V_z to be the set of vertices w for which there exists a path starting at v and ending at w and define V_x as the other vertices. Neither will be empty, since $v \in V_z$ and $u \in V_x$. Now permute the vertices in such a way that $v_i \in V_x$ and $v_j \in V_z$ implies $i < j$. One checks that A_D can then be written as in Definition 1.24 and is thus reducible. □

1.3 Previous work

Before we talk about the known sparse series of finite compositional order, we should briefly discuss the conjugacy classes in $\mathscr{N}_p$. It is a basic fact from group theory that the order of two elements in the same conjugacy class must be the same. For elements of order p each class can be characterised by the depth d and coefficient a_{d+1} of its elements.

Definition 1.27. The *depth* d of an element $\sigma(t)$ in $\mathscr{N}_p$ is $v_t(\sigma - t) - 1$ and therefore lies in $\mathbb{N} \cup \{\infty\}$.

Proposition 1.28 (Klopsch [21] Prop. 1.2)**.** *Let σ_1 and σ_2 in $\mathscr{N}_p$ have depth d, order p and both be of the form $t + at^{d+1} + O(t^{d+2})$ for some non-zero $a \in \mathbb{F}_p$. Then there exists some $\tau \in \mathscr{N}_p$ such that $\tau\sigma_1\tau^{-1} = \sigma_2$.*

B. Klopsch also gave a representative for each conjugacy class of order p, which exists if and only if the depth is not divisible by p.

Definition 1.29. For each prime p, integer $d \not\equiv 0 \bmod p$ and $a \in \mathbb{F}_p^\times$ we define a Klopsch's series by:

$$\frac{t}{\sqrt[d]{1 - dat^d}} = t + at^{d+1} + \dots.$$

When $p = 2$ we denote these series by $\sigma_{K,d}$.

Remark. Since Klopsch's series have finite order, they are algebraic. In fact they satisfy the equation $X^d(1 - dat^d) - t^d = 0$. For $p = 2$ it simplifies to $X^d + (tX)^d + t^d = 0$.

In our search for sparse series of finite order p we now know that they will be conjugate to some Klopsch's series and focus on finding a representative for each pair (d, a).

In the introduction it was mentioned that the only sparse series of order p known previously were of order 2 and depths $2^\mu \pm 1$. See Section 3.2 for their exact definitions.

Theorem 1.30 ([8] Prop. 10.2.1.)**.** *For all depths $d = 2^\mu \pm 1$ and $\mu \in \mathbb{N}$ there exists a sparse automatic series of order 2. These series will be denoted by $\sigma_{S,d}$.*

Example. The series $\sigma_{S,1} = t + t^2 + t^3 + t^6 + t^7 + \dots$ and $\sigma_{K,1} = t + t^2 + t^3 + t^4 + \dots$ both have order 2 and depth 2. Hence, there must exist some $\tau \in \mathscr{N}_2$ so that $\tau\sigma_{S,1}\tau^{-1} = \sigma_{K,1}$. It is in general hard to determine these τ and it is not even known if τ can or cannot be transcendental.

For elements of order higher than p a similar classification of conjugacy classes is possible, see [24]. This classification is based on the depths of $\sigma, \sigma^{\circ p}, \dots, \sigma^{\circ p^{n-1}}$, where p^n is the order of σ. However, since this thesis does not focus on any explicit examples of higher orders, we will not elaborate on it here.

Chapter 2

Automatic sparseness

In this chapter we will examine automatic series more closely. By a theorem of A. Cobham [14] any automatic series $\sigma(t)$ is either sparse or $E(\sigma)_N$ grows faster than N^α for some $\alpha > 0$. Originally, the proof of this theorem was given in the terminology of formal languages, which makes it hard to read from a graph theoretic perspective. In the first section we give a different proof using an idea of A. Szilard, S. Yu, K. Zhang and J. Shallit [28] about a correspondence between sparseness and tied vertices. From the structure of this proof we can also determine the rank of sparseness of σ or the order of growth α for $E(\sigma)_N$. Section 2 and 4 will be devoted to this, respectively. In Section 4 we will see that the supremum of values for α is always the logarithm of an algebraic integer. In Section 3 we prove all the details needed for Section 4 and in Section 5 we will calculate α for several series of finite compositional order.

2.1 Identifying sparseness by the automaton

We will prove that we can see whether an automatic series is sparse by looking solely at the automaton. Namely, the series will be sparse if and only if the automaton contains no tied vertices. Besides, we will see that in the case that an automatic series is not sparse, $E(\sigma)_N$ will grow as a fractional power of N. First we need the following two lemmas.

Lemma 2.1. *Let W_1 and W_2 be two walks in a p-automaton of length n and m respectively. Suppose that the last vertex of W_1 equals the first vertex of W_2. Then these walks can be composed to a walk with size*

$$|W_1 \circ W_2| = |W_1| + p^n |W_2|.$$

Proof. That the walks can be composed is clear. Now suppose W_1 has edges labelled by $b_1, \ldots, b_n$ and W_2 has edges labelled by $c_1, \ldots, c_m$. Then $W_1 \circ W_2$ has edges labelled by $b_1, \ldots, b_n, c_1, \ldots, c_m$ and thus its size becomes

$$b_1 + b_2 p + \ldots b_n p^{n-1} + c_1 p^n + \ldots + c_m p^{n+m-1} = |W_1| + p^n |W_2|. \qquad \square$$

Lemma 2.2. *Let v be a vertex that is not tied but does have a walk to a vertex labelled non-zero. Then there can be at most one walk from v to v that does not contain v a third time.*

Proof. We will prove this lemma by contradiction. So suppose there are two distinct walks W_1 and W_2 starting and ending in v that do not contain v a third time. Because v is not tied, these two walks cannot have the same length. Define the length of W_i to be x_i and assume without loss of generality that $x_1 < x_2$. Since W_1 and W_2 start and end in v, they can both be combined with each other. Consider the walks $W_1 \circ W_2$ and $W_2 \circ W_1$. The length of both these walks is $x_1 + x_2$ and they also still start and end in v. Furthermore, they are not the same walk, since, the first walk will have a vertex v at spot x_1 and the second walk will not, because W_2 would then contain a third v. We have constructed two different walks from v to v of the same length. This contradicts the fact that v is tied and we may conclude that no such W_1 and W_2 can exist, which proves the lemma. □

Theorem 2.3. *Let $\sigma(t)$ be an automatic series in $\mathbb{F}_p[\![t]\!]$, then one of the following must hold:*

- *the p-automaton contains a tied vertex and there is an $\alpha \in (0, 1]$ such that $\#E(\sigma)_N \geq N^\alpha$ for sufficiently large N.*
- *the p-automaton of σ contains no tied vertices and σ is sparse.*

Proof. First we will prove that if the p-automaton contains a tied vertex then we can find $\alpha > 0$ such that $\#E(\sigma)_N \geq N^\alpha$ for sufficiently large N. Secondly we will show that if the p-automaton does not contain any tied vertices then, $\#E(\sigma)_N = O(\log(N)^r)$ for some $r \geq 0$ and thus σ will be sparse. Note that this proves the theorem, since it is obvious that the p-automaton always either contains at least one tied vertex or does not contain any tied vertices at all.

Let v be a tied vertex in a p-automaton that results in the automatic series $\sigma(t)$. By connectivity we know that there must be some walk W_s from 'Start' to v. Because v is tied we also have distinct walks W_1 and W_2 going from v to v with equal length. Finally, we know that there is a walk W_e from v to a vertex labelled non-zero. We write z, x and y for the lengths of W_s, W_1 and W_e respectively. Before we start the proof we take a closer look at W_e. Suppose W_e contains only zero labelled edges. Then W_e only contains vertices with the same label as v and thus v is non-zero itself. So, W_1 and W_2 are both walks from v to a non-zero vertex. Because W_1 and W_2 are distinct, we know at least one of them must contain a non-zero labelled edge, thus we might as well choose W_e to be this walk and assume that W_e contains a non-zero labelled edge. Furthermore, if the last edge of W_e is labelled by zero, we may remove this edge because of the leading-zero-invariance and thus we may assume that W_e does not end in a 0.

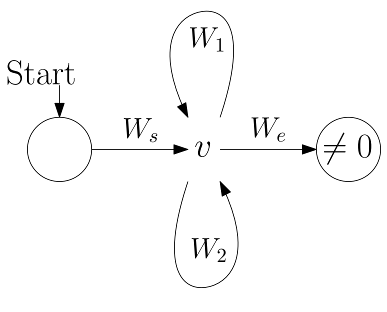


Figure 2.1: A tied vertex and its walks.

Now we can find a lower bound on the size of $E(\sigma)_N$ using the tied vertex v. Namely, for each $k \in \mathbb{Z}_{\geq 0}$ we have that the set

$$\{|W_s \circ W_{i_1} \circ W_{i_2} \circ \ldots \circ W_{i_k} \circ W_e| : i_j \in \{1, 2\} \text{ for } 1 \leq j \leq k\}$$

is a subset of $E(\sigma)$. Note here that we used that W_s starts at Start and ends in v, that W_1 and W_2 both start at and end in v, and that W_e starts at v and ends in a non-zero vertex. Also notice that all choices of k and i_j give walks of different sizes, because the walks are different and do not end with a zero-edge. Hence, by Lemma 2.1, for each k we find in this way 2^k unique elements in $E(\sigma)$ that all lie between $p^{z+y+(k-1)x}$ and p^{z+y+kx}. Even better, we find at least $2^k - 1$ elements in $E(\sigma)_N$ for $N = p^{z+y+(k-1)x}$.
Define $B := \frac{1}{x}\log_p(2)$, let $0 < \alpha < B$ and N a sufficiently large integer. Then there is some k such that N lies between $p^{z+y+(k-1)x}$ and p^{z+y+kx}. Using previous observations we find that,

$$\#E(\sigma)_N \geq 2^k - 1 = p^{xkB} - 1 \geq (p^{z+y})^\alpha p^{xk\alpha} \geq (p^{z+y+xk})^\alpha \geq N^\alpha,$$

where we use that N (and therefore k) is sufficiently large for the second inequality. We have now shown that, if the p-automaton contains a tied vertex, then $\#E(\sigma)_N \geq N^\alpha$ for some α, proving the first part of the theorem.

To prove that σ being sparse implies the automaton has no tied vertices we will first examine the structure that such a p-automaton can have.
Let v be any vertex of the automaton, then by Lemma 2.2 we have three cases:

(i) v is labelled by 0 and every walk from v goes to a 0,

(ii) there is a walk from v to a vertex labelled non-zero, but not a walk from v to v,

(iii) there is a walk from v to a vertex labelled non-zero and exactly one walk from v to v that does not contain v a third time.

In case (i) we can assume that all edges going outwards of v go to v itself. In case (iii) we know the walk from v to v is a cycle. If w is a vertex on this cycle, then the cycle also gives a walk from w to w and thus w is a case-(iii) vertex as well. Since a vertex cannot belong to two cycles, we can group case-(iii) vertices together by their cycle. In a similar manner we will group case-(ii) vertices together by arborescences. To do this we might first need to enlarge our p-automaton a bit. Suppose two distinct edges arrive at the same vertex or the same cycle and neither of these edges is part of a cycle. Then we take this vertex/cycle and all of the p-automaton that can still be reached from this vertex/cycle, and copy it. We attach one of the copies to the first edge and the other copy to the second edge. Now the two edges do not go to the same vertex or cycle any more. This can be done for all such pairs of edges throughout the p-automaton, until there are none left. Because there are only finitely many such pairs, we will still have a p-automaton in the end, that in fact produces the same automatic sequence. This construction ensures that every case-(ii) vertex has at most one ingoing edge. So, if we group together case-(ii) vertices that can be reached through each other without passing a case-(iii) vertex, these form an arborescence inside the graph. Furthermore, the construction also ensures that each cycle can be reached by at most one sequence of earlier cycles. With that in mind we can view our p-automaton as a bigger arborescence, where the vertices are the cycles made of case-(iii) vertices, the edges are the (possibly empty) arborescences made of case-(ii) vertices and the leaves are the case-(i) vertices. Also see Figure 2.2.

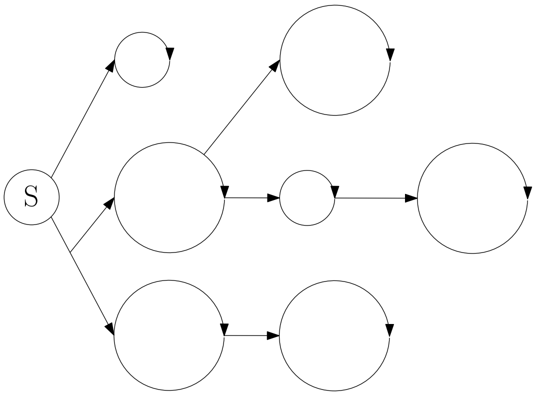


Figure 2.2: An example of a p-automaton without tied vertices viewed as cycle-arborescence. All the case-(i) vertices and their edges are not drawn.

Before we start with the computational part of the proof, we will introduce some important constants and labels based on the new structure of the p-automaton. We define s as the number of cycles there are and label them $C_1, C_2, \ldots, C_s$. For each C_i write l_i for the number of vertices it contains and define $l := \min(l_i)$. Also, let r be the height of the p-automaton viewed as cycle-arborescence, or in other words r is the maximum number of cycles one can pass through consecutively.

Now we will examine how many integers there are in $E(\sigma)_N$ with $N = p^{k+1} - 1$ for some integer $k \geq 1$. Note that the integers $m < p^{k+1}$ are exactly the integers for which the value of a_m can be determined by taking $k+1$ steps in the p-automaton. Write $k+1 = xl + y$ with $0 \leq y < l$. Since all the case-(i) vertices are labelled by 0, it will suffice to give an upper bound on the number of different walks of length $k+1$ that end in a case-(ii) or case-(iii) vertex.
Let v be any case-(ii) or -(iii) vertex. There is a unique sequence of cycles, $C_{i_1}, C_{i_2}, \ldots, C_{i_j}$, with $j \leq r$, the walk has to pass through in order to end up at v. We will proceed with induction on j. If j is either 0 or 1 (where the first can only happen if v is a case-(ii) vertex), the length $(k+1)$-walk to v must be either unique or non-existent. Now suppose we already know that there are $O(\log(N)^{j-1})$ walks to some vertex u in the cycle $C_{i_{j-1}}$ (induction hypothesis). To extend such a walk to v we can either take the shortest walk to C_{i_j} and cycle in C_{i_j} until our steps are gone, or we can cycle $1, 2, \ldots, x$ times in $C_{i_{j-1}}$ and then go to C_{i_j}. This gives $(x+1)O(\log(N)^{j-1})$ ways to end up at v. We had $N = p^{lx+y}$, so $x+1 = O(\log(N))$ and, by induction, we get that there are $O(\log(N)^j)$ walks of length $k+1$ to v. Since there are finitely many vertices in our p–automaton we also get $E(\sigma)_N = O(\log(N)^r)$ and $\sigma(t)$ is sparse. □

In the proof we just saw, we constructed a p-automaton that can be viewed almost as an arborescence but then with cycles instead of vertices and actual arborescences instead of edges. The following definition makes this precise and generalises the notion for digraphs.

Definition 2.4. Let $D = (V, E)$ be a weakly connected digraph. We call D a *cycle arborescence* if for each $v \in V$ one of the following holds:

(i) all outgoing edges of v are self-loops, and v has exactly one other ingoing edge.

(ii) any path starting at v does not end at v, and v has exactly one ingoing edge.

(iii) there exists a unique path from v to v, and on this path exactly one vertex has two ingoing edges, the others have one ingoing edge.

The only exception is the root vertex, which has one ingoing edge less.

Remark. The options for v coincide with the case-(i), -(ii) and -(iii) vertices as defined in the proof. To see this one should note that a case-(i) vertex has p self-loops and $p > 1$, so it is never simultaneously in the third category of the definition.

Remark. Theorem 2.3 can also be used to see if some element $\sigma(t)$ of $\mathbb{F}_p[\![t]\!]$ is an automatic series. Namely, if $\#E(\sigma)_N$ is not $O(\log(N)^r)$ for any r and there also does not exist any α such that $\#E(\sigma)_N \geq N^\alpha$ for sufficiently large N, then a p-automaton for σ both has to and cannot contain a tied vertex, so it does not exist.
For an example of this, define $\sigma(t) \in \mathbb{F}_p[\![t]\!]$ by $a_0 = 0, a_1 = 1, a_2 = 0, a_3 = 0$ and for $N > 3$, $a_N = 0$ if and only if $\lceil \log(N)^{\log(\log(N))} \rceil = \lceil \log(N-1)^{\log(\log(N-1))} \rceil$. Then $E(\sigma)_N$ grows as $\log(N)^{\log(\log(N))}$ which is not $O(\log(N)^r)$ and neither grows as fast as N^α for any $\alpha > 0$. So σ meets the requirements and cannot be automatic.

2.2 Determining the rank of sparseness

We will use the proof of Theorem 2.3 to look at automatic sparse power series and say something about their rank of sparseness. Recall that the rank of sparseness for a series σ is the infimum of r such that $\#E(\sigma)_N$ is in $O(\log(N)^r)$. The proof of Theorem 2.3 directly gives an upper bound for this r, namely the height of the cycle arborescence. It turns out this is also a lower bound and therefore equals the rank.
The idea will be to construct disjoint simple sparse sets of rank at most r. These grow like $\log(N)^r$, which will give us our lower bound. We will make this precise in the following lemma.

Lemma 2.5. *Suppose σ is an automatic sparse power series with a cycle arborescence of height r. Then we can write $E(\sigma)$ as a finite union of disjoint simple sparse sets of rank at most r. Moreover, at least one of those sets will have rank equal to r.*

Proof. We will first prove that we can write $E(\sigma)$ as a finite union of disjoint simple sparse sets of rank at most r and then that we must have at least one set of rank exactly r.

Let v be some vertex in the p-automaton that is labelled non-zero. By the structure of the cycle arborescence we found in the proof of Theorem 2.3, there is a unique sequence of cycles, $C_{i_1}, C_{i_2}, \ldots, C_{i_j}$, one can pass through consecutively to end up at v. Also, there are the following unique walks: a_0 from 'Start' to C_{i_1}, a_k from C_{i_k} to $C_{i_{k+1}}$ for all $1 \leq k \leq j-1$ and a_j from C_{i_j} to v. Write α_k for the walk of the cycle C_{i_k}. All the ways to reach v from 'Start' are

$$\{a_j \alpha_j^{x_j} \cdots a_1 \alpha_1^{x_1} a_0 \mid x_i \in \mathbb{Z}_{\geq 0}\}.$$

This is a simple sparse set of at most rank r, because j must be smaller or equal to the height of the cycle arborescence. We can do this for all non-zero labelled vertices. Note that the simple sparse sets we obtain must be disjoint, because they reach different end points in the p-automaton.

Now we still need to show that there is at least one vertex such that the simple sparse set we obtain has exactly rank r. Recall that we can find r cycles $C_{i_1}, \ldots, C_{i_r}$ through which we can pass consecutively by the definition of height. Also, recall that for any vertex w in C_{i_r} there must exist a walk from w to some non-zero vertex v, because, otherwise w would have been a case-(i) vertex instead of a case-(iii) vertex. So, for this particular vertex v, we find that its corresponding simple sparse set must have rank exactly r. This proves the lemma. □

Theorem 2.6. *Let $\sigma(t)$ be a sparse automatic series in $\mathbb{F}_p[\![t]\!]$ with cycle arborescence of height r. Then the rank of sparseness of σ is r.*

Proof. From the proof of Theorem 2.3 we know that $\#E(\sigma)_N = O(\log(N)^r)$, so the rank of sparseness of σ is at most r. Now we need to show that the rank cannot be smaller.

Lemma 2.5 tells us that $E(\sigma)$ contains a simple sparse set E of rank r and Lemma 1.12 then gives a constant C so that for sufficiently large N,

$$\#E(\sigma)_N \geq \#E \cap [N] \geq C \log(N)^r.$$

Since $\log(N)^r$ is not in $O(\log(N)^{r-\varepsilon})$ for any $\varepsilon > 0$ we conclude that there does not exist any $r' < r$ such that σ has rank at most r'. In other words, r is the infimum we are looking for and σ has rank of sparseness equal to r. □

Corollary 2.6.1. *If $\sigma(t)$ is an automatic sparse series, then its rank of sparseness is an integer that is attained.*

Proof. From Theorem 2.6 we know the rank is the height of a cycle arborescence, which is an integer. In the proof of Theorem 2.3 we saw that if r is the height of the cycle arborescence, then $\#E(\sigma)_N = O(\log(N)^r)$. So indeed, the rank is attained. □

Example. In Figure 1.1a we saw an automaton for a sparse power series. This automaton is already in the cycle arborescence form with two cycles (the self-loops at the top vertices). These cycles can be used consecutively, so the height is 2 and we can conclude that the rank of sparseness is 2.

2.3 Bounds on the number of walks

Here we state the definitions and lemmas we will need to give an upper bound for α in the next section. The focus will lie on bounding the number of loops with certain length for a tied vertex. Intuitively it makes sense that α depends on this; if there are more loops of a certain length we can construct more walks via the tied vertex to a non-zero vertex.

Definition 2.7. Let u and v be vertices of a (di)graph and k a non-negative integer. Define $\Omega_k(u, v)$ as the number of walks of length k, starting at u and ending at v. When $u = v$ we also write $\Omega_k(v)$. Furthermore, the number of walks of length less or equal to k is denoted by $\Omega_{\leq k}(u, v)$, that is $\Omega_{\leq k}(u, v) := \sum_{i=0}^{k} \Omega_i(u, v)$.

Remark. The number of walks of certain length in a (di)graph can also be determined by its adjacency matrix. Namely, if u and v are two vertices of some (di)graph with adjacency matrix A, then the number of walks of length k from u to v corresponds to $(A^k)_{u,v}$. Therefore, $\Omega_k(u,v) = (A^k)_{u,v}$.

The remark shows us that we can use linear algebra to analyse $\Omega_n(v)$. We will see that the spectrum and especially the spectral radius of certain matrices will be of importance.

Perron-Frobenius theory deals with the spectra of matrices that are either positive or non-negative and irreducible. The results are quite numerous and can be found for instance in [25, Chapter 8]. The following proposition assembles the parts needed in this section.

Proposition 2.8. *Let A be any non-negative irreducible matrix. Then*

(i) the spectral radius $\rho(A)$ of A is positive and contained in $\operatorname{Spec}(A)$*. Also, it is a simple eigenvalue.*

(ii) there exists a positive right-eigenvector $\vec{x}$ corresponding to the eigenvalue $\rho(A)$.

Proof. See [25, Perron–Frobenius Theorem p. 673]. □

Perron-Frobenius theory cannot be used immediately. This is because, given any p-automaton or digraph, the adjacency matrix is not necessarily irreducible. In fact, it is known that the adjacency matrix is irreducible if and only if it corresponds to a strongly connected digraph, that is a digraph such that there is a path from u to v for any two vertices u and v. For each vertex v of a certain p-automaton we will be able to look at a subdigraph that is strongly connected. The idea is based on the observation that any vertex on a walk from v to v must be strongly connected to v. Indeed, suppose W is some walk from v to v and u any vertex on W, then clearly there is a path going from v to u and a path from u to v. Even better, if w is yet another vertex strongly connected to v, we can also construct a path from u to w and w to u by going via v. What we see here is that all vertices involved in walks from v to v are strongly connected to each other. It now makes sense to look at the subdigraph that only consists of vertices that are strongly connected to v and the edges between them.

Definition 2.9. Let $D = (V, E)$ be a digraph and $v \in V$. Define

- the set of vertices $V_v := \{u \in V \mid \text{there are paths from } u \text{ to } v \text{ and from } v \text{ to } u\}$,
- the set of edges $E_v := \{(u, w) \in E \mid u, w \in V_v\}$.

Then $D_v := (V_v, E_v)$ is called the *strongly connected subdigraph of D at v*. Its adjacency matrix will be denoted by A_v, its spectral radius by ρ_v and the positive normalized right eigenvector belonging to ρ_v by $\vec{x}_v$. The spectral radius will also be called *the spectral radius of v*.

Remark. Note that previous definition relies on Proposition 2.8, since we assume $\vec{x}_v$ to exist and be unique.

Finally, before we state and prove our first theorem, it is important to note that the number of walks from v to v of length n in D_v is the same as it was in D. This follows from the observation we made before that any such walk must consist solely of vertices that are strongly connected to v. Therefore we still have that $\Omega_n(v) = (A_v^n)_{v,v}$.

Theorem 2.10. *Let $D = (V, E)$ be some p-automaton and v any tied vertex of D. Then there exists some constant $C > 0$ so that*

(i) for all $n \in \mathbb{N}$ we have $\Omega_n(v) \leq \rho_v^n$.

(ii) for infinitely many $n \in \mathbb{N}$ we have $\Omega_n(v) \geq C\rho_v^n$.

Proof. Consider the strongly connected subdigraph D_v of D at v. By Perron-Frobenius we get the equation $A_v\vec{x}_v = \rho_v\vec{x}_v$, which can easily be extended to $A_v^n\vec{x}_v = \rho_v^n\vec{x}_v$ for any $n \in \mathbb{Z}_{\geq 0}$. Because we are interested in the walks from v to v, we study the row corresponding to vertex v of this equation. Without loss of generality we may assume this to be the first row and write $u_2, \ldots, u_m$ for the other vertices in V_v. This gives us the following expression:

$$\Omega_n(v)x_{v,1} + \Omega_n(v, u_2)x_{v,2} + \ldots + \Omega_n(v, u_m)x_{v,m} = \rho_v^n x_{v,1}.$$

Since $x_{v,i} > 0$ and $\Omega_n(v, u_i) \geq 0$ for all i, it is immediate that $\Omega_n(v) \leq \rho_v^n$ for all n. The other inequality needs a bit more consideration. By the pigeon hole principle, we get that for all n we either have $\Omega_n(v)x_{v,1} \geq \frac{1}{m}\rho_v^n x_{v,1}$ or there is some $i \geq 2$ such that $\Omega_n(v, u_i)x_{v,i} \geq \frac{1}{m}\rho_v^n x_{v,1}$. For the second case we look at some path P_i from u_i to v. This path must exist, since the digraph is strongly connected. It should be clear that $\Omega_{n+l(P_i)}(v) \geq \Omega_n(v, u_i)$ which then leads to

$$\Omega_{n+l(P_i)}(v) \geq \left(\frac{x_{v,1}}{mx_{v,i}}\rho_v^{-l(P_i)}\right)\rho_v^{n+l(P_i)}.$$

The path P_i does not depend on the value of n and therefore we can fix a choice for each $2 \leq i \leq m$. Furthermore, let P_1 be the empty path. Taking the constant

$$C := \min_{1 \leq i \leq m}\left(\frac{x_{v,1}}{mx_{v,i}}\rho_v^{-l(P_i)}\right)$$

we indeed get that for infinitely many n,

$$\Omega_n(v) \geq C\rho_v^n.$$

Observe that this does not necessarily hold for any n, since we might need to take a step of length $l(P_i)$ to get there. □

Corollary 2.10.1. *Let v be a tied vertex of some p-automaton, then its spectral radius is bigger than* 1*.*

Proof. Since v is tied, we know there is some $N \in \mathbb{N}$ for which $\Omega_N(v) \geq 2$. Theorem 2.10 now implies $\rho_v \geq \sqrt[N]{2} > 1$. □

Corollary 2.10.2. *For any tied vertex we have*

$$\sup_{n\in\mathbb{N}}\left(\frac{\log_p(\Omega_n(v))}{n}\right) = \log_p(\rho_v).$$

Proof. From Theorem 2.10 we get

$$\sup_{n\in\mathbb{N}}\left(\frac{\log_p(\rho_v^n)}{n}\right) \geq \sup_{n\in\mathbb{N}}\left(\frac{\log_p(\Omega_n(v))}{n}\right) \geq \sup_{n\in\mathbb{N}}\left(\frac{\log_p(C\rho_v^n)}{n}\right).$$

The left side becomes $\log_p(\rho_v)$ and the right side $\sup_{n\in\mathbb{N}}(\frac{\log_p(C)}{n}) + \log_p(\rho_v)$. Since n gets arbitrarily large and $\log_p(C)$ is a negative constant, the first term will go to zero when taking the supremum. So both the lower and upper bound are $\log_p(\rho_v)$, which proves the corollary. □

Lemma 2.11. *Let $D = (V, E)$ be a p-automaton, $v_1, \dots, v_s$ vertices in V and k some positive integer. Define $\rho := \max_{1\leq j\leq s}(\rho_{v_j})$ and let $K_k \subset \mathbb{Z}^s_{\geq 0}$ consist of the solutions $\vec{k} = (k_1, \dots, k_s)$ to $k_1 + \ldots + k_s = k$. Then, for all $k \in \mathbb{N}$,*

$$\mathcal{W}_k := \sum_{\vec{k}\in K_k} \prod_{i=1}^{s} \Omega_{\leq k_i}(v_i) = O(k^{s-1}\rho^k).$$

Proof. We will prove this theorem by induction on the number of tied vertices s. To keep the notation uncluttered we assume $\rho_{v_1} \geq \ldots \geq \rho_{v_s}$ and therefore $\rho = \rho_{v_1}$. This can be done without loss of generality.

For $s = 1$ we get that $\mathcal{W}_k = \Omega_{\leq k}(v_1) = \sum_{i=0}^{k} \Omega_k(v_1)$. Theorem 2.10 provided ρ^i as an upper bound for each $\Omega_i(v_1)$ so that

$$\mathcal{W}_k \leq \frac{\rho^{k+1} - 1}{\rho - 1} \leq \frac{\rho}{\rho - 1}\rho^k = O(\rho^k).$$

The last step uses $\rho > 1$. This proves the induction basis.

Now suppose we know the lemma holds for any $s - 1$ tied vertices and any k (induction hypothesis). We will prove that it also holds for s tied vertices. First we take out the last factor of the product in $\mathcal{W}_k$ to get

$$\sum_{j=0}^{k} \Omega_{\leq j}(v_s) \sum_{\vec{k}\in K_{k-j}} \prod_{i=1}^{s-1} \Omega_{\leq k_i}(v_i),$$

where K_{k-j} is now a subset of $\mathbb{Z}^{s-1}_{\geq 0}$ instead of $\mathbb{Z}^s_{\geq 0}$. For each j of the summand we can bound both factors, respectively by $O(\rho^j)$ and $O(k^{s-2}\rho^{k-j})$, using the induction basis and induction hypothesis. This adds up to

$$\mathcal{W}_k = \sum_{j=0}^{k} O(k^{s-2}\rho^k) = O(k^{s-1}\rho^k).$$

By the principle of induction, this proves the claim. □

Corollary 2.11.1. *There exists a constant $C' > 0$ so that, for infinitely many k and some $1 \leq j \leq s$,*

$$\mathcal{W}_k \leq C' k^{s-1} \Omega_n(v_j).$$

Proof. Theorem 2.10 shows that for any tied vertex v there are infinitely many k such that $\rho_v^k \leq \frac{1}{C}\Omega_k(v)$. The result now follows directly from Lemma 2.11. □

2.4 An upper bound for non-sparse series

In this section we will prove two theorems. The first gives a supremum of all values α for which $\#E(\sigma)_N \geq N^\alpha$ holds. The second shows that this supremum can be expressed as the logarithm of an algebraic integer. To state and prove these results we will use the theory developed in Section 2.3.

Theorem 2.12. *Let $D = (V, E)$ be a p-automaton, $V' \subset V$ the set of tied vertices and $\sigma(t)$ its automatic series. Define*

$$B := \max_{v \in V'}(\log_p(\rho_v)).$$

Then

$$\sup\{\alpha > 0 : \#E(\sigma)_N \geq N^\alpha \text{ for sufficiently large } N\} = B.$$

Proof. In the case that V' is empty we get an empty maximum as definition for B. However, the supremum over the allowed α is empty as well, since $\sigma(t)$ will be sparse. So from now on we may assume V' is non-empty.

We will first prove that B gives an upper bound for the values that α can take. Secondly we will show that B is indeed the best we can do to give such an upper bound. To prove these two things we heavily rely on the fact that by Corollary 2.10.2 B is equal to

$$\sup_{x \in \mathbb{N}, v \in V'} \left(\frac{\log_p(\Omega_x(v))}{x} \right). \tag{2.1}$$

Let α be any value less than B. By (2.1) we know that there must exist some $v \in V'$ and $x \in \mathbb{N}$ such that $\alpha \leq \frac{\log_p(\Omega_x(v))}{x} \neq 0$. We follow the same proof as for the first part of Theorem 2.3, except now we have $\Omega_x(v)$ walks, $W_1, \ldots, W_{\Omega_x(v)}$, to choose from for $W_{i_1}, \ldots, W_{i_k}$ and therefore, there are at least

$$\frac{(\Omega_x(v))^k - 1}{\Omega_x(v) - 1}$$

distinct values in $\#E(\sigma)_N$ if $N = p^{(k-1)x+y+z}$. Since $\Omega_x(v)$ does not equal 1, we can choose k and therefore N large enough such that $\#E(\sigma)_N \geq N^\alpha$.

Now suppose that $\alpha > B$, then it suffices to prove that there exist arbitrarily large N such that $N^\alpha \geq \#E(\sigma)_N$. In other words, there must be infinitely many of such N. We will show that any $N = p^k$, where k meets the condition of Corollary 2.11.1 and is sufficiently large, has this property.

The proof will be by contradiction. So assume $\alpha = B + \varepsilon$ for some $\varepsilon > 0$ and $\#E(\sigma)_N \geq N^\alpha$ for all N sufficiently large. We know from 2.1 that $B \geq \frac{\log_p(\Omega_x(v))}{x}$ holds for all $x \in \mathbb{N}$ and $v \in V'$. In particular it holds for $x = k$. Now let $N = p^k$ and k sufficiently large, then we can rewrite this as

$$\#E(\sigma)_N \geq p^{kB} N^\varepsilon \geq \Omega_k(v) N^\varepsilon.$$

On the other hand we can also give an upper bound for $\#E(\sigma)_N$. Recall that $\#E(\sigma)_N$ counts the walks of length k that start at 'Start' and end in a non-zero vertex. Any such walk has a certain structure and we will be able to estimate the number of walks with this structure. Let $s \geq 0$ be an integer and $v_1, \ldots, v_s \in V'$ all distinct. We define the following sets:

(i) the set $\Omega(\text{Start}, v_1)(V')$ consisting of walks from 'Start' to v_1 not passing through any tied vertex.

(ii) for $1 \leq i \leq s$ the set $\Omega(v_i)$ consisting of walks from v_i to v_i.

(iii) for $1 \leq i \leq s-1$ the set $\Omega(v_i, v_{i+1})(V')$ consisting of walks from v_i to v_{i+1} not passing through any tied vertex.

(iv) the set $\Omega(v_s, \text{End})(V')$ consisting of walks from v_s to any non-zero vertex and not passing through any tied vertex.

Now let W be any walk from 'Start' to a non-zero vertex. There exists some integer $s \geq 0$ and $v_1, \ldots, v_s$ distinct in V' such that we can take a unique element of all above defined sets and that W is a composition of these walks. Note that there is only one possibility for a composition, since the begin and end vertex of each walk are known. However, the choice of $v_1, \ldots, v_s$ is not necessarily unique, but we will only need existence.

A p-automaton has a finite number of vertices and therefore of tied vertices. This implies there are at most $(|V'|+1)!$ choices for s, $v_1, \ldots, v_s$. Since we are only interested in walks of length k, one of those combinations must give the largest number of walks of length k. From now on we write s, v_1, ..., v_s for this combination. We will continue by estimating the number of walks that can be written as a composition of elements of the sets defined before. Recall from the proof of Theorem 2.3 that the number of walks between two vertices is either of order $O(\log(N)^r)$ for some integer r, or includes a tied vertex. This gives an estimate for the cardinality of the sets of (i), (iii) and (iv) when we also require the length to be at most k. This leaves $k' \leq k$ steps for the walks in the sets of (ii). Note that the number of elements of length at most k_i for any such set equals $\Omega_{\leq k_i}(v_i)$. Define $K_k := \{(k_1, \ldots, k_s) \subset \mathbb{Z}^s_{\geq 0} \mid k_1 + \ldots + k_s = k\}$, then we get

$$\#E(\sigma)_N \leq (|V'|+1)! O(\log(N)^r) \sum_{\vec{k} \in K_k} \prod_{i=1}^{s} \Omega_{\leq k_i}(v_i).$$

Now we can use Corollary 2.11.1 to see that

$$\#E(\sigma)_N \leq C k^{s-1} \Omega_k(v) \log_p(N)^r = C\Omega_k(v) \log_p(N)^{r+s-1}$$

for some tied vertex v, constant $C > 0$ and infinitely many k. Putting everything together we have

$$\Omega_k(v) N^{\varepsilon} < \#E(\sigma)_N \leq C\Omega_k(v) \log_p(N)^{r+s-1},$$

for infinitely many $N = p^k$. However, since $\log_p(N)^{r+s-1}$ will grow slower than N^{ε}, there will exist some sufficiently large N such that $\Omega_k(v) N^{\varepsilon} > C\Omega_k(v) \log_p(N)^{r+s-1} \geq \#E(\sigma)_N$, a contradiction. We conclude that there does not exist a sufficiently large N such that for all $n \geq N$ it holds that $\#E(\sigma)_n \geq n^{\alpha}$ if α is bigger than B. Hence B is indeed the supremum of all α for which this condition does hold. □

Remark. Theorem 2.12 tells us that any $\alpha < B$ works and that any $\alpha > B$ does not work. However, it says nothing about $\alpha = B$. As can be seen in Figures 2.3a and 2.3b, either situation occurs. In both examples we have $B = \frac{\log_2(2)}{2} = \frac{1}{2}$ and in the first we see $\#E(\sigma)_N \geq N^{\frac{1}{2}}$ for all $N \geq 1$, whereas in the second $\#E(\sigma)_N \leq N^{\frac{1}{2}}$ for infinitely many N.

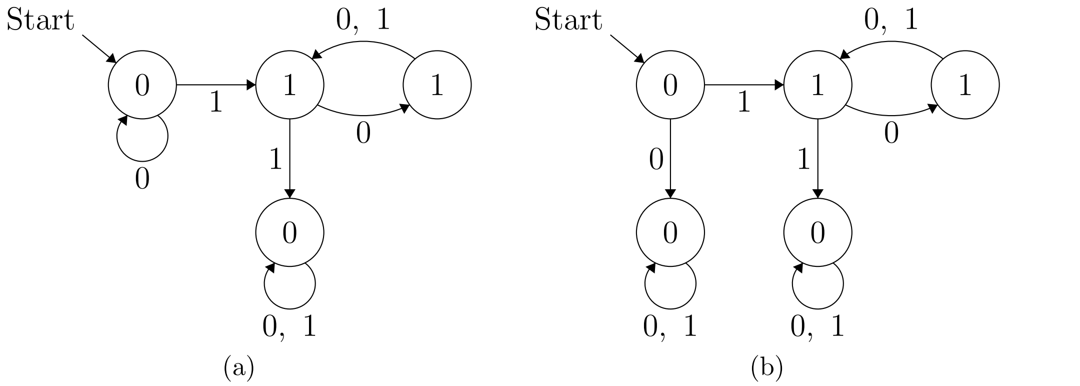


Figure 2.3: Two 2-automata for which the corresponding power series $\sigma(t)$ has either (a) $\#E(\sigma)_N \geq N^{\frac{1}{2}} = N^B$ for all N or (b) $\#E(\sigma)_N < N^{\frac{1}{2}} = N^B$ for infinitely many N. In case (a) this can be seen by noting $\#E(\sigma)_{4^k+1} = 2^{k+1} = \sqrt{4^{k+1}}$ for all $k \in \mathbb{Z}_{\geq 0}$, and in case (b) by noting $\#E(\sigma)_{4^k} = 2^{k-1}$ for all $k \in \mathbb{Z}_{\geq 0}$.

Proposition 2.13. *Let B be defined as in Theorem 2.12. Then*

(i) B can be written as $\log_p(\beta)$ with β an algebraic integer.

(ii) B is either transcendental or rational.

Proof. The first claim follows quite directly from Theorem 2.12. We know that $B = \log_p(\rho_v)$ for some tied vertex v of the automaton. Also, ρ_v is in the spectrum of A_v by Proposition 2.8 and therefore a zero of the characteristic polynomial $p_v(X)$ of A_v. A characteristic polynomial is always monic and $p_v(x)$ must have integer coefficients, because A_v only has integer entries by definition. So ρ_v is integral and taking $\beta = \rho_v$ proves the first part.

For the second part we use the Gelfond-Schneider Theorem of which a proof can be found in [22, Appendix 1, Corollary 2]. It states that for any two algebraic numbers a, b such that $a \neq 0, 1$ and b is irrational, a^b is transcendental. Since $p^B = \beta$ is algebraic, we find that B cannot be an irrational algebraic number and the second claim follows. □

Example. Klopsch's series $\sigma_{K,3}$ has three tied vertices and all have a spectral radius of $\sqrt{2}$ (See Figure 1.1b). Hence, we find that $B = \log_2(\sqrt{2}) = \frac{1}{2}$.

Definition 2.14. Let $\rho > 1$ be an algebraic integer in $\mathbb{R}$ such that all other roots of its minimal polynomial have smaller absolute value than $|\rho|$. Then ρ is called a *Perron number*.

Remark. It is actually possible to say slightly more about which values β can take. By [23, Theorem 3], it must be an integer root of a Perron number and each such value can be attained for some p that is sufficiently large. The proof of this theorem relies on Ergodic theory, which is why we left the result out of Proposition 2.13.

2.5 Non-sparse series of finite order

In this section we briefly discuss some specific values B and β can take when the automatic series is of finite compositional order. First we look at Klopsch's series and then at some examples that behave less nicely.

Proposition 2.15. *Let p be a prime, d a positive integer coprime to p and suppose $-1/d$ has p-adic expansion $(b_0, b_1, b_2, \ldots)$ with period M. Define M_j, for $0 \leq j \leq p-1$, as the number of b_i with $i < M$ that equal j. Then for any Klopsch's series with depth d we have*

$$B = \sum_{j=1}^{p-1} \frac{M_j}{M} \log_p(j+1).$$

Proof. Recall that Klopsch's series are defined by

$$\frac{t}{\sqrt[d]{1-dat^d}} = t + at^{d+1} + \ldots,$$

with $a \in \mathbb{F}_p^{\times}$. Using the p-adic expansion of $-1/d$, this can be rewritten as

$$t(1-dat^d)^{-1/d} = t \prod_{i=0} (1-dat^{dp^i})^{b_i}.$$

Since each $b_i \leq p-1$, we get that each factor $(1-dat^{dp^i})^{b_i}$ has b_i+1 terms. Now consider $E(\sigma)_N$ where σ is Klopsch's series for p, d and some fixed a and $N = p^M$. It is easy to see that $E(\sigma)_N$ has size $(b_0+1)\cdot(b_1+1)\cdots(b_{M-1}+1) = \prod_{j=0}^{p-1}(j+1)^{M_j}$. To conclude something about B or β we need to also know how the size of $E(\sigma)_N$ grows when N grows. For this we use that $(b_0, b_1, \ldots)$ is purely periodic, since $-1/d \in [-1,0] \cap \mathbb{Q}$ and d is coprime to p. Namely, let $N' = p^{nM}$ with n an integer, then we get that the size of $E(\sigma)_{N'}$ equals

$$(b_0+1)(b_1+1)\cdots(b_{nM-1}+1) = \prod_{j=0}^{p-1}(j+1)^{nM_j} = (\#E(\sigma)_{p^M})^n.$$

From this we see

$$\beta = \sqrt[M]{\#E(\sigma)_{p^M}} = \prod_{j=0}^{p-1}(j+1)^{M_j/M} \text{ , and } B = \sum_{j=0}^{p-1} \frac{M_j}{M} \log_p(j+1).$$

Note that β and B are independent of a. □

Examples.

- When $d = p^n - 1$, we get $\frac{-1}{d}$ equals $(\overbrace{1,0,\ldots,0}^{n},1,0,\ldots)$, an expansion with period n, $M_0 = n-1$, $M_1 = 1$ and $M_j = 0$ for all other j. So $\beta = 2^{\frac{1}{n}}$ and $B = \frac{1}{n}\log_p(2)$ and thus B can become arbitrarily small, even for series of finite compositional order.

- Since $p-1$ is always a divisor of p^n-1, we can also give a general expression for $d = \frac{p^n-1}{p-1}$. Then $\frac{-1}{d} = (p-1)(\overline{1,0,\ldots,0})$, which has period n and $M_0 = n-1, M_{p-1} = 1$ and $M_j = 0$ for all other j. This leads to $\beta = p^{\frac{1}{n}}$ and $B = \frac{1}{n}$.

- Using the relation $(p^{2n}-1) = (p^n+1)(p^n-1)$ we find that, if $d = p^n+1$, then $(-1/d) = (p-1, p-1, \ldots, p-1, 0, \ldots, 0, p-1, \ldots)$ with $M = 2n$, $M_{p-1} = n$ and $M_0 = n$. In this case we have $\beta = p^{\frac{1}{2}}$ and $B = \frac{1}{2}$.

- Other divisors of $p^n - 1$ can be described using a similar reasoning. For example, take $p = 5$ and divisor 3. Then 3 does not divide $p - 1 = 4$, but it does divide $p^2 - 1 = 24$. Taking $d = 24/3 = 8$ we get $M = 2$, $M_0 = 1$ and $M_3 = 1$, so $\beta = 4^{1/2}$ and $B = \frac{1}{2}\log_5(4) = \log_5(2)$. However, for $p = 5$ there does not exist any d such that $\beta = 4$ and $B = \log_5(4)$: the only possible periodic expression for this would be $(\overline{3})$, which equals $\frac{-3}{4}$ and cannot be expressed as $\frac{-1}{d}$.

From the series in Table 2.1 we know what their automata and order are and thus we can calculate B and β. The series $\sigma_{K,d}$ are Klopsch's series, the series $\sigma_{G,d}$ are greedy series as defined in Section 3.1 and all other series can be found in the Appendix.

series	order	B	β
$\sigma_{K,2^m-1}$	2	$\frac{1}{m}$	$\sqrt[m]{2}$
$\sigma_{K,2^m+1}$	2	$\frac{1}{2}$	$\sqrt{2}$
$\sigma_{K,23}$	2	$\frac{4}{11}$	$\sqrt[11]{2^4}$
$\sigma_{CS}^{\circ 2}$	2	1	2
$\sigma_{V,1}$	2	≈ 0.79845	real root of $x^5 - x^4 - x^2 - x - 2$
$\sigma_{V,2}$	2	≈ 0.79845	real root of $x^5 - x^4 - x^2 - x - 2$
$\sigma_{V,3}$	2	≈ 0.79845	real root of $x^5 - x^4 - x^2 - x - 2$
$\sigma_{G,1}$	2	≈ 0.69424	$\phi = \frac{1+\sqrt{5}}{2}$
$\sigma_{G,3}$	2	≈ 0.69424	$\phi = \frac{1+\sqrt{5}}{2}$
$\sigma_{G,5}$	2	≈ 0.55146	real root of $x^3 - x^2 - 1$
$\sigma_{\min}$	4	1	2
σ_{CS}	4	1	2
σ_J	4	1	2
$\sigma_J^{\circ 3}$	4	1	2
$\sigma_{(1,5)}$	4	≈ 0.85933	real root of $x^8 - 2x^7 + x^5 - x^3 - x^2 + 2x - 2$

Table 2.1: Some series of finite order and their B and β.

Remark. The fact that $\sigma_{V,1}, \sigma_{V,2}$ and $\sigma_{V,3}$ all have the same value for B is not trivial. They all commute, but we also know examples of sparse series that commute with non-sparse series, such as $\sigma_{CS}^{\circ 3}$, which commutes with σ_{CS}. It could be interesting to examine more commuting series and their values for B and β.

Chapter 3

The structure of the support

The goal of this chapter is to look if series of finite compositional order can be constructed by analysing what the order implies about the structure of the support. In the first three sections we examine the case $p = 2$ and give some proofs and results. In the fourth section we try to generalise the ideas to $p > 2$. It turns out, however, that $p = 2$ is a lot more straightforward than the rest. The main result is given at the end of Section 2, where we describe new automatic sparse series of order 2 and depth $2^\mu - 3$.

3.1 Structural lemmas and a greedy algorithm

In this first section we look at what we can say about the support of a series of order 2 in general. This leads to some lemmas and propositions on which we base a greedy algorithm to generate series of order 2. We prove that all greedy series exist and are algebraic, but will later see they are not sparse.

Lemma 3.1. *Let $\sigma(t) \in \mathbb{F}_2[\![t]\!]$ and $k \in \mathbb{N}$. Write k uniquely as $\sum_{i=i}^m r_i$, with r_i all powers of 2 such that $r_1 < r_2 < \ldots < r_m$. Then*

$$\sigma(t)^k = \sum_{e_1,\ldots,e_m \in E(\sigma)} t^{e_1 r_1 + \ldots + e_m r_m}.$$

Proof. Since $k = r_1 + r_2 + \ldots + r_m$, we can rewrite $\sigma(t)^k$ as

$$\sigma(t)^{r_1+r_2+\ldots+r_m} = \prod_{i=1}^m \sigma(t)^{r_i} = \prod_{i=1}^m \sigma(t^{r_i}).$$

The last equality follows because all r_i are powers of two. Expanding $\sigma(t^{r_i})$ for each i as $\sum_{d_i \in E(\sigma)} t^{e_i r_i}$ now proves the result. □

Corollary 3.1.1. *If $\sigma(t)$ has depth d, then the two smallest values in $E(\sigma(t)^k)$ are k and $R(k,d) := k + 2^{v_2(k)} d$.*

Proposition 3.2. *Suppose $\sigma(t) \in \mathscr{N}_2$ has compositional order 2 and depth d. Then there is an integer $m \in E(\sigma) \backslash \{1\}$ such that $v_2(m) < v_2(d+1)$.*

Proof. We will prove this by contradiction. Suppose for all $k \in E(\sigma)\backslash\{1, d+1\}$ we have $v_2(k) \geq v_2(d+1)$, then also $R(k,d) > R(d+1,d)$. Recall that $\sigma^{\circ 2}(t) = \sum_{k\in E(\sigma)} \sigma(t)^k$ and note that $\sigma(t)$ cancels all terms t^k of each $\sigma(t)^k$ if $k > 1$. Hence the expansion of $\sigma^{\circ 2}(t)$ is of the form $t + t^{R(d+1,d)} + O(t^{R(d+1,d)+1})$ and σ is not of order 2. □

Proposition 3.2 yields a new proof for a known result about possible depths of finite order elements.

Corollary 3.2.1. *If $\sigma(t)$ is of finite compositional order, its depth cannot be even.*

Proof. From the Proposition 3.2 it is clear that a series of compositional order 2 cannot have an even depth. We know that any element in $\mathscr{N}_2$ of finite order has order 2^r for some $r \in \mathbb{N}$. Suppose that σ has even depth d, by a similar reasoning as in previous proof we find that the depth of $\sigma^{\circ 2}$ is $R(d+1,d) - 1$. Clearly $R(d+1,d) - 1$ is still even and therefore $\sigma^{\circ 2}$ also cannot have order 2. By an inductive argument on r this shows that σ cannot have any order of the form 2^r and we conclude σ cannot have finite compositional order if its depth is even. □

In the case that $d+1$ is not divisible by 4 we can even explicitly give an odd element in the support.

Proposition 3.3. *Let $\sigma(t)$ in $\mathscr{N}_2$ have compositional order 2 and suppose its depth is d with $v_2(d+1) = 1$. Then $2d+1$ is the smallest odd number in $E(\sigma)\backslash\{1\}$.*

Proof. From Proposition 3.2 we know $E(\sigma)\backslash\{1\}$ contains an odd element and thus a smallest odd element m. First suppose $m < 2d+1$, then $R(m,d) < R(d+1,d)$. It is easy to see that for all $m' \in E(\sigma)$ with $m' \neq 1, m$ we have $R(m',d) > R(m,d)$ and therefore the term $t^{R(m,d)}$ will not be canceled in $\sigma^{\circ 2}$, a contradiction. Now suppose $m > 2d+1$, then $R(m,d) > R(d+1,d)$ and, similarly as in the previous case, the term $t^{R(d+1,d)}$ will not be canceled. We conclude that $m = 2d+1$ and indeed, $R(2d+1,d) = R(d+1,d)$. □

Definition 3.4. Let $p(t) \in \mathbb{F}_p[t]$ be a polynomial of degree $n \in \mathbb{N}$ and $\sigma(t) \in \mathbb{F}_p[\![t]\!]$. We say $p(t)$ can be *extended to* $\sigma(t)$ if $p(t) \equiv \sigma(t) \mod t^{n+1}$.

Lemma 3.5. *Let $p(t) \in t + t^2\mathbb{F}_2[t]$ of depth d and degree n, and define d_2 as the depth of $p(p(t))$. Assume $d_2 < n + d$, then $p(t)$ cannot be extended to an element in $\mathscr{N}_2$ of order two.*

Proof. Suppose that $d_2 < n + d$, then clearly $p(t)$ itself is not an element of order two. Define the polynomial $q(t) = p(t) + t^m$ with $m > n$. We will prove that $q(q(t))$ still has depth $d_2 < n + d < m + d$. By induction this proves that, for any power series $\sigma(t)$ that extends $p(t)$, the depth of $\sigma(\sigma(t))$ has to be the same as for $p(p(t))$ and thus $\sigma(t)$ cannot be an order 2 element.

Note that $q(q(t)) = \sum_{k\in E(q)}(q(t))^k$. Let k be an integer in the support of q unequal to 1. We know that $k \geq d+1$ and $m > 1$ and therefore, using the binomium of Newton,

$$q(t)^k = p(t)^k + O(t^{m+k-1}) = p(t)^k + O(t^{m+d}).$$

When $k = 1$ we obviously get $p(t) + t^m$. Plugging this into the first equation we find

$$q(q(t)) = t^m + p(p(t)) + p(t)^m + O(t^{m+d}).$$

Since $p(t) = t + t^{d+1} + O(t^{d+2})$, we find in a similar way that $p(t)^m = t^m + O(t^{m+d})$ and because $m + d > n + d \geq d_2 + 1$, we get $q(q(t)) = p(p(t)) + O(t^{m+d}) = t + t^{d_2+1} + O(t^{d_2+2})$. Hence the depth of $q(q(t))$ is still d_2 independent of the choice for m. □

Suppose we have a certain polynomial $p(t)$, which we suspect to be extendable to an order 2 element $\sigma(t)$ of $\mathscr{N}_2$. The goal is to find $\sigma(t)$ and the first step is to find the term of smallest degree of $\sigma - p$, in other words the next monomial t^m. The proof of Lemma 3.5 shows that $m \leq d_2 - d + 1$ needs to hold. Ideally we would also like $\sigma(t)$ to be sparse and thus have relatively few terms. A straightforward strategy, based on greedy choice algorithms, would now be to take m as big as possible for each following term. This motivates the following definition.

Definition 3.6. Let $\sigma(t) \in t + t^2\mathbb{F}_2[\![t]\!]$ and define d_2 as the depth of $\sigma(t)^{\circ 2}$. We call $\tau(t) = \sigma(t) + t^{d_2-d+1}$ or just $d_2 - d + 1$ the *greedy choice* for $\sigma(t)$.

The next theorem gives an idea of when we can, cannot or must use the greedy choice.

Theorem 3.7. *Let $p(t) \in t + t^2\mathbb{F}_2[t]$ of depth $d > 0$ and degree n, so that $p(p(t))$ has depth d_2. Suppose there is an $m \in \mathbb{N}$ for which $p(t) + t^m$ can be extended to an order two power series. Then one of the following holds:*

(i) m is even and strictly smaller than $d_2 - d + 1$.

(ii) m is odd, equals $d_2 - d + 1$ and the depth of $q(q(t))$ is strictly larger than d_2.

Proof. The depths of $p(t)$ and $p(t) + t^m$ are the same, so Corollary 3.2.1 tells us that d is odd and $d_2 - d + 1 \equiv d_2 \mod 2$. We already concluded that $m \leq d_2 - d + 1$ from Lemma 3.5. For the first statement it therefore suffices to show that if d_2 is even, choosing $m = d_2 - d + 1$ gives a contradiction. This contradiction will also follow from Lemma 3.5, because the new polynomial will have degree $d_2 - d + 1$, but its second composite keeps depth d_2 and $d_2 - d + 1 + d > d_2$. The proof of the second statement uses a similar contradiction for odd $m < d_2 - d + 1$.

Assume d_2 is even and let $q(t) = p(t) + t^m$ be the greedy choice, so $m = d_2 - d + 1$. By the same argument as in the proof of Lemma 3.5, we find $(p(t) + t^m)^k = p(t)^k + O(t^{k+d_2-d})$. We consider three cases, $k = 1, k = d + 1$ and $k > d + 1$. When $k = 1$ it is immediate that we get $p(t) + t^m$. When $k > d + 1$ we find $p(t)^k + O(t^{d_2+2})$, which is enough to show that the depth of $q(q(t))$ stays d_2. For $k = d + 1$ we look at the binomial expansion and find

$$(p(t) + t^m)^{d+1} = p(t)^{d+1} + \binom{d+1}{1} p(t)^d t^m + O(t^{2m+d-1}) = p(t)^{d+1} + O(t^{d_2+2}).$$

For the second equality we used that $d+1$ is even and $m \geq 2$, so $2m+d-1 = m+d_2 \geq d_2+2$. We can now express $q(q(t)) = \sum_{k \in E(q)} q(t)^k$ as:

$$q(q(t)) = t^m + p(p(t)) + p(t)^m + O(t^{d_2+2}).$$

Recall that the smallest two elements of $E(p(t)^m)$ are m and $R(m, d) = m + 2^{v_2(m)}d > m + d = d_2 + 1$. Therefore $p(t)^m = t^m + O(t^{d_2+2})$ and we find

$$q(q(t)) = p(p(t)) + O(t^{d_2+2}).$$

We see that indeed, the depth of $q(q(t))$ stays d_2 and we have arrived at our contradiction.

To prove the second statement we will use the same tactics as we did for the first. We did not use that m was even until the very last step. Therefore, if we take m odd and not necessarily the greedy choice, we still find

$$q(q(t)) = t^m + p(p(t)) + p(t)^m + O(t^{d_2+2}).$$

However, in this case the smallest elements of the support of $p(t)^m$ are m and $R(m, d) = m+d$. If m is smaller than d_2-d+1, we find $R(m, d) < d_2+1$ and thus the depth of $q(q(t))$ becomes $m + d - 1$. Obviously $m + d - 1 < m + d$, which contradicts that $q(t)$ can be extended by Lemma 3.5. If m equals $d_2 - d + 1$, we find that $R(m, d) = d_2 + 1$, so the second term of $p(t)^m$ cancels the second term of $p(p(t))$ and therefore $q(q(t))$ indeed has higher depth than $p(p(t))$ had. Finally, we already saw that m cannot be bigger than $d_2 - d + 1$, which proves the second part. □

Remarks.

- Theorem 3.7 shows that, if m is an odd element of the support of some σ of order 2 and if $p(t)$ is a polynomial of degree at most $m-1$ such that $p(t) \equiv \sigma(t) \mod t^m$, then m must be the greedy choice for p. In other words, any odd integer in the support must have been the greedy choice at some point and the even elements of the support uniquely determine the odd elements.
- Theorem 3.7 also gives a condition to be able to apply the greedy choice. Namely, whenever d_2 is even we cannot use it and will instead be obliged to take m even. If the minimum of $d_2, R(m, d) - 1, R(d + 1, m - 1) - 1$ and $R(s, m - 1) - 1$, with s the smallest odd element of the support, then occurs an odd number of times, it is equal to the depth of $q(q(t))$. Except for d_2, all these options are odd.

Corollary 3.7.1 (Greedy Algorithm). *Let $d \in 2\mathbb{N} - 1$ and $p_0(t) = t + t^{d+1} \in \mathbb{F}_2[t]$. For all $n \geq 0$ define $p_{n+1}(t) = p_n(t) + t^{m_n}$, where m_n is the greedy choice for p_n. Suppose m_n is always odd, then $\lim_{n\to\infty} p_n(t) = \sigma(t) \in \mathscr{N}_2$ has order two.*

Definition 3.8. For each positive odd integer d we define $\sigma_{G,d}$ as the series of depth d that results from the greedy algorithm as defined in Corollary 3.7.1. This series is called *the greedy series of depth d.*

From Corollary 3.7.1 we see that greedy series of odd depth d is well defined if and only if m_n is odd for all n. This seems a bit of a strong requirement, since $p_n(t)^{\circ 2}$ will also always contain terms of even degree. However, in Theorem 3.12 we will prove the existence of $\sigma_{G,d}(t)$, which implies that m_n must always be odd. The strategy will be to guess an algebraic equation $F(t, X) = 0$ and show that $X = \sigma_{G,d}(t)$ must be a solution.

I used the greedy algorithm to write a computer program in C# and calculate $\sigma_{G,d}$ for all odd $d < 500$ modulo t^N with $N = 2^{16} + 1$. These series do not seem to be sparse for any depth and a proof of this will be given in Chapter 4. For each n-th power of two I checked how many terms σ contains that are of degree less than 2^n. So far this seems to always grow like β^n with $\beta \in (1, 2)$. See Table 3.1 for some of the computational results. Before we prove the general case we will first look at the example $\sigma_{G,1}$. In Table 3.1 this series stands out, because the computed terms of the sequence $|E_{2^n}(\sigma_{G,1})| - 1$ equal Fibonacci numbers. It

d n	1	3	5	7	9	11	31	33	99
1	1	0	0	0	0	0	0	0	0
2	2	1	0	0	0	0	0	0	0
3	3	1	1	0	0	0	0	0	0
4	5	2	2	1	1	1	0	0	0
5	8	3	3	1	2	1	1	0	0
6	13	5	5	2	3	2	1	1	0
7	21	7	7	3	4	2	1	2	1
8	34	12	10	5	6	3	1	3	1
9	55	19	17	9	9	4	1	5	2
10	89	31	25	13	13	7	2	8	3
11	144	49	38	19	22	11	3	11	5
12	233	80	58	32	33	18	5	17	8
13	377	129	87	52	53	28	9	26	11
14	610	209	131	85	83	46	17	41	18
15	987	337	196	139	131	73	33	62	25
16	1597	546	293	222	205	118	49	96	38

Table 3.1: For several depths d the number of elements in $E(\sigma_{G,d})_{2^n}\backslash\{1\}$.

therefore seems likely that $\beta = \phi$, the golden ratio. To prove that this is indeed the case we will need the following lemma.

Lemma 3.9 ([8], Lemma 10.2.2)**.** *If a polynomial $F(t,X) = 0 \in \mathbb{F}_2[t,X]$ is symmetric in t and X, that is $F(t,X) = F(X,t)$, and when regarded as an algebraic equation in X over $\mathbb{F}_2((t))$, has, for some $d \geq 1$, a unique solution $\sigma \in \mathscr{N}_2$ of depth d, then σ is of order 2.*

Proposition 3.10. *Define $F(t,X) = X^3 + (tX)^2 + t^3$. Then $\sigma_{G,1}$ exists, $F(t,\sigma_{G,1}(t)) = 0$ and $\beta(\sigma_{G,1}) = \phi$.*

Proof. Since we want to be able to calculate β, it will be convenient to have the automaton of $\sigma_{G,1}$. Using the Magma program in [7], we start by calculating what the automaton of a solution of depth 1 to $F(t,X) = 0$ is. This gives the unique automaton as shown in Figure 3.1. It now suffices to show that this automaton generates the series $\sigma_{G,1}$.

First note that if we follow the edges labeled by $(0,0)$, we end up in a vertex labeled by 0, for which both outgoing edges are loops. Hence any integer divisible by 4 cannot be in the support. We also end up in this vertex if we follow $(0,1,0,\dots,0,1)$, showing that 2 is the only integer in the support that is 2 modulo 4. Therefore, besides 2, the support consists solely of odd integers. Additionally we know that the automaton must generate a series of order 2, because of Lemma 3.9 and the uniqueness of the solution. Theorem 3.7 implies that any odd integer in the support of a series of order 2 must have been the greedy choice at some point. Thus a series of order 2 having the described support structure must be a greedy series. We conclude that $\sigma_{G,1}$ exists and equals the series generated by the automaton. It is immediate that $F(t,\sigma_{G,1}) = 0$ and easy to calculate $\beta = \phi$ using Theorem 2.12. □

Remark. The same procedure can be followed for any depth d, as long as Magma can still calculate the automaton. For $d = 3$ and $d = 5$ this has been done, using the polynomial $X^{d+2} + (tX)^{d+1} + t^{d+2}$. This led to $\beta = \phi$ and β equal to the real root of $x^3 - x^2 - 1$, respectively.

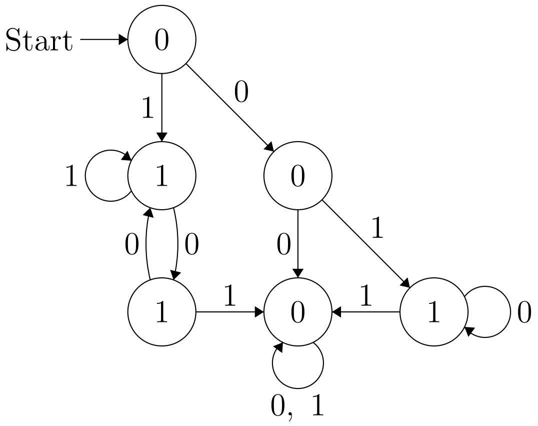


Figure 3.1: The automaton for the series $\sigma_{G,1}$, which is a solution to $X^3 + t^2X^2 + t^3$.

For the general case we need a specific formulation of Hensel's Lemma.

Theorem 3.11 (Hensel's Lemma, [17] Theorem 7.3)**.** *Let R be a ring that is complete with respect to the ideal $\mathfrak{m}$, and let $F(x) \in R[x]$ be a polynomial. If a is an approximate root of F in the sense that*

$$F(a) \equiv 0 \mod F'(a)^2\mathfrak{m}$$

then there is a root b of F near a in the sense that

$$F(b) = 0 \text{ and } b \equiv a \mod F'(a)\mathfrak{m}.$$

If $F'(a)$ is a non zero-divisor in R, then b is unique.

Corollary 3.11.1 ([5], Theorem 3.4)**.** *Let k be a perfect field with characteristic $p > 0$. Let $F(t, X)$ be a polynomial in $k[t, X]$, and denote its derivative with respect to X by F'. We assume that we are given a nonnegative integer ρ and a polynomial $\bar{f}(t)$ such that $F(t, \bar{f}(t)) \equiv 0 \mod t^{2\rho+1}$ and $F'(t, \bar{f}(t)) \not\equiv 0 \mod t^{\rho+1}$.*
There exists a unique series $f(t) \in k[\![t]\!]$ congruent to $\bar{f}(t)$ modulo $t^{\rho+1}$ for which $F(t, f(t)) = 0$.

Proof. Take R to be the ring $k[\![t]\!]$, which is complete with respect to the maximal ideal (t). Let g be the smallest integer that is not congruent to -1 modulo p, such that the coefficient $a_{g+1}(t)$ of X^{g+1} in $F(t, X)$ is non-zero. Such a g must exist, because otherwise $F'(t, X)$ is always zero. Then $F'(t, X) = a_{g+1}(t)X^g + O(X^g)$ and we have two possibilities, either $\bar{f}(0) = 0$ and $v_t(a_{g+1}) + g \leq \rho$ or $\bar{f}(0) \neq 0$.

First, suppose $\bar{f}(0) = 0$. Any element in $k[\![t]\!]$ of the form $a + O(t)$ with $a \in k\backslash\{0\}$ is a unit. Hence $F'(t, \bar{f}(t))^2(t)$ equals the ideal $(t)^{2g+1} \supset (t)^{2\rho+1}$. This shows that $\bar{f}(t)$ is an approximate root of F and because $F'(t, \bar{f}(t))$ is unequal to zero, Hensel's Lemma provides a unique series $f(t) \equiv \bar{f}(t) \mod t^{g+1}$ that is an actual root of F. Left to prove is that these series are also congruent modulo $t^{\rho+1}$. Define k to be the largest integer for which $f(t) \equiv \bar{f}(t) \mod t^{k+1}$, we will prove $k \geq \rho$. Write

$$\bar{f}(t) \equiv f_1 t + \ldots + f_{2\rho} t^{2\rho} \text{ and } f(t) \equiv f_1 t + \ldots + f_k t^k + h_{k+1} t^{k+1} + \ldots + h_{2\rho} t^{2\rho}.$$

Suppose p divides some integer i and look at the difference between $\bar{f}(t)^i$ and $f(t)^i$. Taking such a power can be seen as summing over the terms $f_{j_1} t^{j_1} \cdots f_{j_i} t^{j_i}$ for all possibilities to make i choices in the set $\{1, \ldots, 2\rho\}$. Whenever the only chosen integers are smaller than $k+1$, this term will also exist for the other series and thus be canceled. Besides, if some integer in $\{1, \ldots, 2\rho\}$ is chosen exactly once, the number of permutations that lead to the same result will be divisible by i. In other words the coefficient in front of this term will be divisible by the characteristic and disappear. Hence, the term of smallest degree in $\bar{f}(t)^i - f(t)^i$ will have at least degree $2k + i \geq k + g + i > k + g$.
Now, since both these series are a root of F modulo $t^{2\rho+1}$, we know that $F(t, \bar{f}(t)) - F(t, f(t))$ is also zero for this modulo. However, we also find that

$$F(t, \bar{f}(t)) - F(t, f(t)) = a_{g+1}(t)(g+1) f_1^g (f_{k+1} - h_{k+1}) t^{k+g+1} \mod t^{k+g+2}.$$

So $k + v_t(a_{g+1}) + g + 1 \geq 2\rho + 1$ and because $v_t(a_{g+1}) + g$ was at most ρ, we may conclude that $k \geq \rho$.

Now assume $\bar{f}(0) = a \neq 0$ and define $\bar{h}(t) := \bar{f}(t) - a$ and $H(t, X) := F(t, X + a)$. One can easily check that $H(t, \bar{h}(t)) \equiv 0 \mod t^{2\rho+1}$ and $H'(t, \bar{h}(t)) \not\equiv 0 \mod t^{\rho+1}$. The first case shows that there exists a unique series $h(t) \equiv \bar{h}(t) \mod t^{2\rho+1}$ such that $H(t, h(t)) = 0$. Clearly $f(t) = h(t) + a$ is then the unique root of $F(t, X)$ that is congruent to $\bar{f}(t)$. □

Theorem 3.12. *Let d be a positive even integer and define $F_{G,d}(t, X) := X^{d+2} + (tX)^{d+1} + t^{d+2}$. Then $\sigma_{G,d}$ exists and $F_{G,d}(t, \sigma_{G,d}(t)) = 0$.*

Proof. Since no ambiguity can arise we will drop the subscript of $F_{G,d}(t, X)$ and simply write $F(t, X)$.

The strategy will be to first prove $F(t, X)$ has an order 2 solution of depth d. Then we will show that this solution can have no even elements in the support, besides $d+1$. By a uniqueness argument this will imply the solution equals $\sigma_{G,d}$.

Lemma 3.9 tells us that if we show that $F(t, X)$ has a unique root in $\mathscr{N}_2$ of depth d, this series will have order 2. The existence and uniqueness both follow from Corollary 3.11.1. For let $\bar{\sigma}(t) = t + t^{d+1}$ and $\rho = d + 1$, then

$$F'(t, \bar{\sigma}(t)) = (t + t^{d+1})^{d+1} \equiv t^{d+1} \not\equiv 0 \mod t^{d+2}, \text{ and}$$

$$F(t, \bar{\sigma}(t)) = (t+t^{d+1})^{d+2} + t^{d+1}(t+t^{d+1})^{d+1} + t^{d+2} \equiv t^{d+2} + t^{2d+2} + t^{2d+2} + t^{d+2} \equiv 0 \mod t^{2d+3},$$

where we used Lemma 3.1 to calculate the powers modulo t^{2d+3}. The general solution will be denoted by $\sigma(t)$.

We will prove the second part by contradiction. Suppose $E(\sigma)\backslash\{d+1\}$ contains an even integer and let N be the smallest. Define the integers $i_1 < \ldots < i_n$, such that $d + 1 =$

$2^{i_1} + \ldots + 2^{i_n}$ and note that $i_1 > 0$, because $d+1$ is even. By Lemma 3.1 the terms $\sigma(t)^{d+2}$ and $(t\sigma(t))^{d+1}$ can now be rewritten as:

$$\begin{aligned}\sigma(t)^{d+2} &= \sum_{e_0,\ldots,e_n \in E(\sigma)} t^{e_0+2^{i_1}e_1+\ldots+2^{i_n}e_n}, \text{ and}\\ (t\sigma(t))^{d+1} &= \sum_{e_1,\ldots,e_n \in E(\sigma)} t^{d+1+2^{i_1}e_1+\ldots+2^{i_n}e_n}.\end{aligned}$$

Since $\sigma(t)$ is a zero of $F(t, X)$, any term of the form t^x must occur an even number of times. Thus, if $x \neq d+2$, we want there to be an even number of solutions to the problem:

$$\begin{aligned} x &= e_0 + 2^{i_1}e_1 + \ldots + 2^{i_n}e_n \quad \text{with } e_0,\ldots,e_n \in E(\sigma), \text{ or}\\ x &= d+1+2^{i_1}e_1 + \ldots + 2^{i_n}e_n \text{ with } e_1 \ldots, e_n \in E(\sigma).\end{aligned}$$

Since $d+1 \in E(\sigma)$, we can pair any solution $(e_1, \ldots, e_n)$ of the second equation with a solution $(d+1, e_1, \ldots, e_n)$ of the first equation and simplify the statement. Now we want there to be an even number of solutions to:

$$x = e_0 + 2^{i_1}e_1 + \ldots + 2^{i_n}e_n \text{ with } e_0, \ldots, e_n \in E(\sigma), \text{ and } e_0 \neq d+1.$$

Consider the case $x = N+d+1$. This value is even and therefore, any solution $(e_0, \ldots, e_n) = (E_0, \ldots, E_n)$ requires E_0 to be even. By definition of N this implies $E_0 \geq N$. This only leaves room for the solution; $(e_0, \ldots, e_n) = (N, 1, \ldots, 1)$, which indeed exists. Hence the number of solutions is odd, which shows that N cannot exist and that $E(\sigma)\backslash\{d+1\}$ consists solely of odd integers.

We now have two bits of important information about σ. It is a series of order 2 and depth d and it has, besides $d+1$, only odd elements in its support. In Theorem 3.7 we saw that any odd element of the support must have been the greedy choice at some point. Hence, any series with these properties must be the greedy series $\sigma_{G,d}$. This proves that the greedy series exists for all odd d and $F(t, \sigma_{G,d}(t)) = 0$. □

3.2 Sparse series of order 2

In this section we will generalise the greedy algorithm a bit to describe a new method for finding series of order 2. This method yields new sparse series of depths $2^\mu - 3$. To have some more examples we will first define the known sparse series of order 2 that were also mentioned in Theorem 1.30.

Definition 3.13 ([8] Proposition 10.2.1.)**.** Take $\mu \in \mathbb{N}$. For $d = 2^\mu - 1 > 1$, define $\sigma_{S,d}$ as the power series with support

$$E(\sigma_{S,d}) = \{1\} \cup \left\{ \frac{d+1}{d-1}\left(d \cdot \left(\frac{d+1}{2}\right)^{k-1} - 1\right) \,\middle|\, k \geq 1 \right\}.$$

For $d = 2^\mu + 1$, define $\sigma_{S,d}$ as the power series with support

$$E(\sigma_{S,d}) = \left\{ 1 - d + d\sum_{j\in J} 2^j (d-1)^{k(j)} \,\middle|\, \emptyset \neq J \subset \{0, 1, \ldots, \mu-1\}, k : J \to \mathbb{Z}_{\geq 0} \right\}.$$

For $d = 2$, define $\sigma_{S,d}$ as the power series with support

$$E(\sigma_{S,d}) = \{1\} \cup \{2^n - 2, 2^n - 1 \mid n \geq 2\}.$$

The ranks of sparseness equal $1, \mu$ and 1, respectively.

The ideas for the greedy algorithm followed from Theorem 3.7. This theorem also shows that there is a clear distinction between even and odd elements in the support of an order 2 series. The following consequence marks the importance of this result and gives a generalisation of the greedy algorithm.

Proposition 3.14. *Let $\mathscr{E}$ be any subset of the positive even integers. Then there can be at most one subset $\mathscr{O}$ of the positive odd integers so that, $\sigma \in \mathscr{N}_2$, $E(\sigma) = \mathscr{E} \cup \mathscr{O}$ and $\sigma^{\circ 2}(t) = t$. Moreover, $N \in \mathscr{O}$ if and only if it is the greedy choice for $\sum_{i \in E(\sigma)_{N-1}} t^i$.*

Examples.

- Taking $\mathscr{E}$ equal to $\{d + 1\}$ reduces this proposition to the greedy algorithm (Corollary 3.7.1) and thus leads to $\sigma_{G,d}$.
- Taking $\mathscr{E} = \{2^n - 2 \mid n \geq 2\}$ we will find the series $\sigma_{S,2}$. Here $\mathscr{E}$ can also be expressed as the simple sparse set $\{1^k 10 \mid k \in \mathbb{Z}_{\geq 0}\}$.
- When $d = 2^\mu - 1$, the set $\mathscr{E} = E(\sigma_{S,d}) \backslash \{1\}$ leads to $\sigma_{S,d}$. Here $\mathscr{E}$ equals the simple sparse set $\{10^{\mu-1}(10^{\mu-2})^k 0 \mid k \in \mathbb{Z}_{\geq 0}\}$.
- For $d = 2^\mu + 1$ we may take $\mathscr{E}$ as the subset of $E(\sigma_{S,d})$ with either $0 \notin J$ or $k(0) \neq 0$ to get $\sigma_{S,d}$. Here $\mathscr{E}$ consists of the union of multiple simple sparse sets.

An important question to ask is: which choices for $\mathscr{E}$ lead to sparse series? Obviously $\mathscr{E}$ itself should be sparse and, since σ has finite order and therefore is automatic, it should even be a union of simple sparse sets by Lemma 2.5. This is also what we saw in previous examples. At the moment the best method we have is just making guesses. Taking $\mathscr{E} = \{2^\mu - 2\} \cup (2^{\mu-1}\mathscr{E} + 2^{\mu-1} - 2)$ for $\mu \in \mathbb{N}$, which equals $\{2, 4, 8, 16, \ldots, \}$ when $\mu = 1$, led to the following result.

Theorem 3.15. *Let $\mu \geq 2$, $d = 2^\mu - 3$ and define*

$$\sigma_{M,d}(t) := t + \sum_{i \geq 1} \sum_{j \geq 0} t^{2^j (2^{(\mu-1)i} - 1) \frac{2^\mu - 3}{2^{\mu-1} - 1} + 1}.$$

Then $\sigma_{M,d}$ is an order two power series of depth d that is 2-sparse.

Proof. Note that $\sigma_{M,d}(t)$ is a zero of

$$F(t, X) = X^{2^\mu} + t^{2^{\mu-1}} X^{2^{\mu-1}} + tX^2 + t^2 X + t^{2^\mu}.$$

By Corollary 3.11.1 for $\rho = 2$, we find that $\sigma_{M,d}(t)$ is the unique series of depth d for which this is the case. Also, this polynomial is symmetric in X and t, so, by Lemma 3.9, $\sigma_{M,d}(t)$ has order 2.

That $\sigma_{M,d}(t)$ is 2-sparse can be seen directly from the formula. □

The simple sparse sets when $\mu \geq 3$ are

$$\{1(1^{\mu-2}0)^k \mid k \in \mathbb{Z}_{\geq 0}\} \text{ and } \{1(1^{\mu-2}0)^{k_1}1^{\mu-3}010^{k_2}1 \mid k_1, k_2 \in \mathbb{Z}_{\geq 0}\}.$$

For $\mu = 2$ we get

$$\{10^k \mid k \in \mathbb{Z}_{\geq 0}\} \text{ and } \{11^{k_1}0^{k_2}1 \mid k_1, k_2 \in \mathbb{Z}_{\geq 0}\}.$$

In both cases the support satisfies the relation $E = \{1\} \cup \mathscr{E} \cup (2E - 1)$, which in a moment we will be able to call a support relation.

3.3 Support relations

For some $d \in 2\mathbb{N} - 1$ there have been found sparse series of depth d and order 2 in $\mathscr{N}_2$, but not for all. It remains unknown whether such series even exist for all odd d. In this section we analyse $\sigma_{S,d}$ for $d = 2^\mu \pm 1 > 1$, and show their supports have a structure where the odds determine the evens. Since the odds are all greedy choices, this structure determines the series completely. For other d we give possible ways to extend this structure and then formulate a conjecture based on this.

Definition 3.16. Let $E \subset \mathbb{N}$ and suppose there exists $\mathscr{O} \subset \mathbb{N}$ and $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathbb{N}$ such that

$$E = \mathscr{O} \cup (\alpha_1 E + \beta_1) \cup \ldots \cup (\alpha_n E + \beta_n), \tag{3.1}$$

with $(\alpha_i E + \beta_i)$ for $1 \leq i \leq n$ pairwise disjoint. We call (3.1) a *support relation.* If E is the support of some $\sigma(t) \in \mathscr{N}_p$, we say *$\sigma$ satisfies a support relation.*

Proposition 3.17. *The series $\sigma_{S,d}$ with $d = 2^\mu + 1$ is the unique series of order 2 that satisfies the support relation $E = \mathscr{O} \cup (2E + 2^\mu)$ with $\mathscr{O} \subset 2\mathbb{N} - 1$.*

Proof. We will first prove that $\sigma_{S,d}$ satisfies the support relation and then show uniqueness.

Since $d - 1$ and 2 are both even, it suffices to show that any even m is in the support of $\sigma_{S,d}$ if and only if $m = 2e + d - 1$ for some e in the support. Every element m in $E(\sigma_{S,d})$ is determined by a set $J \subset \{0, \ldots, \mu - 1\}$ and a function $k := J \to \mathbb{Z}_{\geq 0}$. One can check that m is even if and only if $0 \notin J$ or $k(0) \neq 0$ and we will consider these two cases separately. First assume $0 \notin J$, then we can rewrite m as

$$m = 2\left(1 - d + d\sum_{j \in J} 2^{j-1}(d-1)^{k(j)}\right) + d - 1.$$

Define $J' := J - 1 \subset \{0, \ldots, \mu - 1\}$ and $k' : J' \to \mathbb{Z}_{\geq 0}$ by $k'(j') = k(j' + 1)$. The element m' of the support determined by J' and k' indeed meets the requirement $2m' + d - 1 = m$. Now suppose $0 \in J$ and $k(0) \neq 0$, we can rewrite m as

$$m = 2\left(1 - d + d(d-1)^{k(0)} + d\sum_{j \in J \setminus \{0\}} 2^{j-1}(d-1)^{k(j)}\right) + d - 1.$$

Substituting one factor $(d - 1)$ of $d(d - 1)^{k(0)}$, shows that it is equal to $d2^{\mu-1}(d - 1)^{k(0)-1}$. Define $J' := ((J - 1)\backslash\{-1\}) \cup \{\mu - 1\} \subset \{0, \ldots, \mu - 1\}$ and $k' : J' \to \mathbb{Z}_{\geq 0}$ by $k'(j') = k(j' + 1)$

if $j' \neq \mu - 1$ and $k'(\mu - 1) = k(0) - 1$. The element m' of the support determined by J' and k' also satisfies $2m' + d - 1 = m$.
We have now shown that any even m is in the set $2E(\sigma_{S,d}) + d - 1$. The other implication says that any element in $(2E(\sigma) + d - 1)$ is in the support and can be proven in a similar way by going from J to $J + 1$ and $k(j)$ to $k'(j + 1)$. This proves that $\sigma_{S,d}$ satisfies the support relation.

To prove uniqueness we will argue that there can be at most one set $\mathscr{O} \subset 2\mathbb{N} - 1$ such that a set E satisfying the support relation $E = \mathscr{O} \cup (2E + d - 1)$ is the support of a series of order 2. To see this we use Theorem 3.7, which said that any odd integer in the support of an order-two series is uniquely determined by the smaller elements in the support. Since an even element is in E if and only if it is in $2E + d - 1$, the same can be said about the even integers. Hence any element in the support is uniquely determined by the smaller elements and thus E must be unique. □

Proposition 3.18. *The series $\sigma_{S,d}$ with $d = 2^\mu - 1$ is the unique series of order 2 that satisfies the support relation $E = \mathscr{O} \cup (2^{\mu-1}E + 2^{\mu-1})$ with $\mathscr{O} \subset 2\mathbb{N} - 1$. Furthermore, in this case, $\mathscr{O} = \{1\}$.*

Proof. The proof of uniqueness follows exactly the same argument as in Proposition 3.17. Therefore, we will concentrate on proving $\sigma_{S,d}$ satisfies this support relation. It again suffices to show that any $m \in E(\sigma_{S,d})$ is even if and only if it is in $2^{\mu-1}E(\sigma_{S,d}) + 2^{\mu-1}$.

To minimize notation, define $m(k) := \frac{d+1}{d-1}(d \cdot (\frac{d+1}{2})^{k-1} - 1)$. From the definition of $\sigma_{S,d}$ it is clear that the support now consists of 1 and all $m(k)$ with $k \geq 1$. All $m(k)$ are even, so we will show that they are exactly the set $(2^{\mu-1}E(\sigma_{S,d}) + 2^{\mu-1})$. When $k = 1$ we find $m(k) = d + 1$, which equals $2^{\mu-1} \cdot 1 + 2^{\mu-1}$. When $k > 1$ one can easily check that $m(k) = 2^{\mu-1}m(k-1) + 2^{\mu-1}$. We may conclude $E(\sigma_{S,d}) = \{1\} \cup (2^{\mu-1}E(\sigma_{S,d}) + 2^{\mu-1})$. □

We said before that we wanted to find order-two series σ such that the odd values in $E(\sigma)$ determine the even values. This is exactly what is stated in the previous two propositions, the evens of $\sigma_{S,d}$ are determined by a support relation $E = \mathscr{O} \cup (aE + b)$ because $\mathscr{O} \subset 2\mathbb{N} - 1$ and a, b are both even. To get a better idea of what happens exactly and how to implement this we will take a look at the example $\sigma_{S,5}$. Here we will also see how the odd integers are determined by the greedy choice.

Example. For $\sigma_{S,5}$ we are given the support relation $E = \mathscr{O} \cup (2E + 4)$. Since 1 is always an element of the support, we immediately get $\{1, 6, 16, 36, 76, \ldots\} \subset E(\sigma_{S,5})$. Define the power series $\tau_0(t) = t + t^6 + t^{16} + \ldots$ and set $\mathscr{O}_0 = \{1\}$, then $\tau_0(t)$ satisfies $E(\tau_0) = \mathscr{O}_0 \cup (2E(\tau_0) + 4)$. This series does not have order 2 and the depth of $\tau_0(\tau_0(t))$ equals 15. This gives a greedy choice of $15 - 4 = 11$, which shows $\{11, 26, 56, 116, \ldots\}$ is also a subset of the support of $\sigma_{S,5}$. Define $\mathscr{O}_1 = \{1, 11\}$ and let $\tau_1(t)$ be the series that satisfies the support relation $E = \mathscr{O}_1 \cup (2E + 4)$. Then $\tau_1(t) = t + t^6 + t^{11} + t^{16} + t^{26} + \ldots$ and the depth of $\tau_1(\tau_1(t))$ is 45. The greedy choice now becomes 41 and we may define $\mathscr{O}_2 = \mathscr{O}_1 \cup \{41\}$.
Generally speaking, for each $n \geq 1$, we calculate $d_2(\tau_{n-1})$ (the depth of $\tau_{n-1}^{\circ 2}$), define $\mathscr{O}_n = \mathscr{O}_{n-1} \cup \{d_2(\tau_{n-1}) - 4\}$ and τ_n as the series that satisfies the support relation $E = \mathscr{O}_n \cup (2E + 4)$. Then $\sigma_{S,5}$ will be the limit of τ_n as n goes to infinity.

The previous algorithm will also work for all other $\sigma_{S,d}$ with $d = 2^\mu \pm 1$ and, more interestingly, for other support relations $E = \mathscr{O} \cup (aE + b)$ of which we do not yet know if there exists an order-two series that satisfies the relation. If these series exist, they will be of depth $a + b - 1$. We might, though, need to make one slight addition to the algorithm: if $\tau_n^{\circ 2}$ has infinite depth, we define $\mathscr{O}_{n+1} = \mathscr{O}_n$, or equivalently $\sigma = \tau_n$. Also the same caution is necessary as with the greedy algorithm: a series that satisfies the support relation only exists if $d_2(\tau_n)$ is always odd. We can therefore now state the following conjecture, which we will later on prove for all cases where a is a power of 2 and $b \geq \frac{-a+2}{3}$.

Conjecture 3.19. *Let $a > 0$ and $b > -a$ be even integers. Then there exists a unique element in $\mathscr{N}_2$ of order 2 that satisfies the support relation $E = \mathscr{O} \cup (aE+b)$ with $\mathscr{O} \subset 2\mathbb{N}-1$.*

In Propositions 3.17 and 3.18 we already proved the uniqueness in the specific cases $a = 2, b = 2^\mu$ and $a = b = 2^{\mu-1}$. This argument is relatively simple to generalise, which we will do in the next proof. Besides, we will give a new criterion to determine the sparseness of a series, depending only on the set $\mathscr{O}$ of the support relation.

Theorem 3.20. *Let $a > 0$ and $b > -a$ be even integers, then there exists at most one element σ in $\mathscr{N}_2$ of order 2 that satisfies the support relation $E = \mathscr{O} \cup (aE + b)$ with $\mathscr{O} \subset 2\mathbb{N} - 1$. Furthermore, σ is sparse if and only if $\mathscr{O}$ is sparse.*

Proof. Suppose such a series σ and set $\mathscr{O}$ exist. Since all elements in $\mathscr{O}$ are odd, Theorem 3.7(i) tells us that an odd integer m is in $\mathscr{O}$ if and only if it was the greedy choice for the polynomial $p(t) \equiv \sigma(t) \mod t^m$ of degree at most $m - 1$. Hence, if for all integers smaller than m we have decided whether they are in the support, then it is also uniquely determined whether m should be in it. The same conclusion holds if m is even; it is determined by a, b and the elements of $\mathscr{O}$. Therefore we find that the support, and thus σ, is uniquely determined.

For the second claim we note that the left implication is immediate: $\mathscr{O}$ is a subset of the support of σ and thus, sparse whenever σ is. For the other implication we may assume $\mathscr{O}$ to be sparse of some rank r. In mathematical notation: $|\mathscr{O} \cap [N]| = O(\log(N)^r)$. Define a' strictly between 1 and a and suppose N is equal to $(a')^k$. To bound the number of even elements, we look at how many iterations each odd element m can have gone through, while staying smaller than N. After j iterations we get an element of at least size $a^j m + \frac{a^j-1}{a-1} b$. Since $b \geq -a + 2$ this is always bigger than

$$a^j m - a^j + 1 + \frac{a^j - 1}{a - 1} > a^j(m - 1) + a^{j-2}.$$

When $m \geq (a')^i + 1$, there can be at most $k - i$ iterations. If k is large enough, we can assure the same thing for $m = (a')^i$, and in particular $m = 1$. We find that the number of odd elements that went through at least j iterations is at most $|\mathscr{O}_{(a')^j}| = O(j^r)$. There can have been at most k iterations so far and we find $|E(\sigma)_N| = O((k + 1)^r) + O(k^r) + O((k - 1)^r) + \ldots + O(2^r) + O(1) = O(k^{r+1}) = O(\log(N)^{r+1})$, from which we conclude that σ is sparse of rank at most $r + 1$. □

To check if the existence claim of the conjecture might fail, I wrote a computer program in C#. The program follows the steps as described in previous example about $\sigma_{S,5}$. For all $a + b - 1 = d < 100$ and $a \leq 32$ it is possible to calculate these series modulo t^N for $N = 2^{11} + 1$. See Table 3.2 for some of the results.

(a,b)	$(2,0)$	$(4,-2)$	$(6,-4)$	$(6,-2)$	$(8,-4)$	$(4,2)$	$(4,6)$	$(4,8)$	$(6,6)$	$(8,6)$
$n \backslash d$	1	1	1	3	3	5	9	11	11	13
1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	1	1	1	1	1	1	1
3	4	3	3	1	1	1	1	1	1	1
4	8	4	6	2	2	2	1	1	1	1
5	16	6	9	2	3	3	2	1	1	2
6	32	9	15	2	5	5	3	2	2	3
7	64	14	35	7	7	8	5	3	3	4
8	128	21	72	16	11	13	8	4	7	6
9	256	32	134	33	17	21	13	6	10	9
10	512	48	266	75	24	34	21	8	21	13
11	1024	73	494	151	35	55	34	12	47	19

Table 3.2: For several series σ that satisfy a support relation $E = \mathscr{O} \cup (aE + b)$ so that $\mathscr{O} \subset 2\mathbb{N} - 1$ and σ might be of order 2, the depth d and the number of elements in $\mathscr{O} \cap [2^n]$ are given.

Examples.

- $(a,b) = (2,0)$: In Section 3.3 we will see that taking a a power of two and $b = 0$ will lead to a Klopsch's series. In Section 2.6 we saw that in those cases $\beta = \sqrt[m]{2}$, which is what we also see here.
- $b = 6$: Based on this data alone it seems likely that in the cases that $b = 6$ the value of β should be 2. Whereas in all the other cases displayed (besides $(2,0)$) it looks like $\beta < 2$. Considering also the data not displayed, it seems likely that all supports where a is not a power of two grow like $O(2^n)$ and all series where a is a power of two grow like $O(\beta^n)$ with $\beta < 2$.
- $(a,b) = (4,2)$ and $(4,6)$: Just like with $\sigma_{G,1}$ we recognise the Fibonacci-series and thus would expect $\beta = \phi$. This is indeed the case as we will see at the end of this section.

For $\sigma_{S,d}$ it is already known that they are algebraic by the equations $F(t,X) = (tX)^{2^{\mu-1}} + X + t$ when $d = 2^\mu - 1 > 1$ and $F(t,X) = (tX)^{2^\mu} + X^{2^\mu - 1} + t^{2^\mu - 1}$ when $d = 2^\mu + 1$. We will generalise this, with a new proof, in the next two theorems for a any power of two and $b \geq \frac{-a+2}{3}$.

Proposition 3.21. *Let a be a power of two and $b \geq a$ an even integer. Then $F_{(a,b)} := (tX)^b + X^{b-a+1} + t^{b-a+1}$ has a unique solution in $\mathscr{N}_2$ of depth $a + b - 1$. This solution has order 2 and satisfies the support relation $E = \mathscr{O} \cup (aE + b)$ with $\mathscr{O} \subset 2\mathbb{N} - 1$.*

Proof. Before we begin the reader should remind oneself of the notation $..\widehat{\ }..$, introduced at beginning of this thesis. Furthermore, the polynomial $F_{(a,b)}$ is denoted by F and this proof will follow the same strategy as the proof of Theorem 3.12.

By Lemma 3.9 and Hensel's Lemma we find the existence and uniqueness of a series $\sigma(t) \in \mathscr{N}_2$ of depth $a + b - 1$, order 2 and so that $F(t, \sigma(t)) = 0$. Corollary 3.11.1 is here used with

$\bar{\sigma}(t) = t + t^{a+b}$ and $\rho = b - a$. Then

$$F'(t, \bar{\sigma}(t)) = \bar{\sigma}(t)^{b-a+1} \equiv t^{b-a} \not\equiv 0 \mod t^{b-a+1} \text{ and,}$$

$$F(t, \bar{\sigma}(t)) \equiv 0 + t^{b-a+1} + t^{b-a+1} \equiv 0 \mod t^{2b-2a+1}.$$

(This only shows that $\sigma(t) \equiv \overline{\sigma}(t) \equiv t \mod t^{b-a+1}$, but it can easily be shown, by analysing $F(t, X)$, that the depth of σ must be $a + b - 1$.)

We will prove by contradiction that this σ satisfies the support relation $E = \mathscr{O} \cup (aE + b)$. It suffices to show that an element in $E(\sigma)$ is even if and only if it is in $aE(\sigma) + b$. Let N be the smallest even integer such that one of the following holds:

(i) $N \in E(\sigma)$ but $N \notin aE(\sigma) + b$.

(ii) $N \notin E(\sigma)$ but $N \in aE(\sigma) + b$.

In other words, N is the smallest contradiction to the previous statement. Furthermore, define $M := \frac{N-b}{a}$. In case (i) $M \notin E(\sigma)$ and in case (ii) $M \in E(\sigma)$.

Now let $i_1 < i_2 < \ldots < i_n$ and l all be integers, such that $b = 2^{i_1} + \ldots + 2^{i_n}$ and $a = 2^l$. Define j to be the smallest index for which $i_j \geq l$ and $m := i_j - l$. We can now write $b - a$ as:

$$2^{i_1} + ..\widehat{2^{i_j}}.. + 2^{i_n} + (2^{i_j-1} + 2^{i_j-2} + \ldots + 2^l).$$

Note that the term between brackets equals zero if $i_j = l$. We will first only consider the case where $i_j > l$ and thus $m \geq 1$.
Using this equation we can write $(t\sigma(t))^b$ and $(\sigma(t))^{b-a+1}$ as follows:

$$(t\sigma(t))^b = \sum_{e_1,\ldots,e_n \in E(\sigma)} t^{b+2^{i_1}e_1+\ldots+2^{i_n}e_n}, \text{ and}$$

$$(\sigma(t))^{b-a+1} = \sum_{e_0,..\widehat{e_j}..,e_{n+m} \in E(\sigma)} t^{e_0+2^{i_1}e_1+..\widehat{2^{i_j}e_j}..+2^{i_n}e_n+2^{i_j-1}e_{n+1}+\ldots+2^l e_{n+m}}.$$

Because $\sigma(t)$ is a solution to $F(t, X) = 0$, any term of the form t^x must occur an even number of times. Thus, for all $x \neq b - a + 1$ there have to be an even number of solutions to:

$$x = b + 2^{i_1}e_1 + \ldots + 2^{i_n}e_n \qquad \text{with } e_1, \ldots, e_n \in E(\sigma), \text{ or}$$
$$x = e_0 + 2^{i_1}e_1 + ..\widehat{2^{i_j}e_j}.. + 2^{i_n}e_n + 2^{i_j-1}e_{n+1} + \ldots + 2^l e_{n+m} \text{ with } e_0, ..\widehat{e_j}.., e_{n+m} \in E(\sigma).$$

Consider $x = b - a + N$. Then in case (i) we find the solution $e_0 = N, e_1 = ..\widehat{e_j}.. = e_{n+m} = 1$ and in case (ii) the solution $e_0 = a + b, e_1 = ..\widehat{e_j}.. = e_{n+m-1} = 1, e_{n+m} = M$. We will now prove that all other possible solutions come in pairs. Suppose therefore that there is another solution and first consider the case that it solves the first equation. Then the solution is of the form $(e_0, \ldots, e_n) = (E_0, \ldots, E_n)$. Because $2^{i_j} > a$, we must have $E_j < M$. By the definition of N this implies $aE_j + b \in E(\sigma)$. One can now check that $e_0 = aE_j + b$, $e_{n+1} = \ldots = e_{n+m} = E_j$ and $e_i = E_i$ for all $1 \leq i \leq n$ and $i \neq j$, gives a different solution. Secondly consider the case that the solution solves the second equation but is not of the form $e_{n+1} = \ldots = e_{n+m}$ and $e_0 = ae_{n+1} + b$. Then the solution can be written as $(e_0, ..\widehat{e_j}.., e_{n+m}) = (E_0, ..\widehat{E_j}.., E_{n+m})$. Since $b - a + N$ is even, so is E_0, and $a + b \leq E_0 < N$

and $E_{n+m} < M$. Therefore, by definition of N, there exists some E_0' in the support such that $E_0 = aE_0' + b$ and $aE_{n+m} + b$ must be in $E(\sigma)$. Let $1 \leq k \leq m$ be the largest index for which $E_{n+k} \neq E_0'$, this index exists because we excluded the specific case where it does not. One may check that $e_0 = aE_{n+k}+b$, $e_i = E_i$ for all $1 \leq i \leq n+k-1$ and $i \neq j$, $e_{n+k} = E_0$, $e_i = E_k$ for all $i \geq n+k+1$ is a different solution. Note that if we had started with the constructed solution instead of $(E_0, ..\widehat{E_j}.., E_{n+m})$, this method would have constructed $(E_0, ..\widehat{E_j}.., E_{n+m})$, so we have now paired all solutions.
Before we draw any conclusions we go back to the case $m = 0$. We then want an even number of solutions to

$$\begin{aligned} x &= b + 2^{i_1}e_i + \ldots + 2^{i_n}e_n \qquad\quad \text{with } e_1, \ldots, e_n \in E(\sigma), \text{ or} \\ x &= e_0 + 2^{i_1}e_1 + ..\widehat{2^{i_j}e_j}.. + 2^{i_n}e_n \text{ with } e_1, ..\widehat{e_j}.., e_n \in E(\sigma). \end{aligned}$$

Let $x = b - a + N$ again, in case (i) the solution $e_0 = N, e_1 = ..\widehat{e_j}.. = e_{n+m} = 1$ still holds, but in case (ii) we now need to take $e_1 = ..\widehat{e_j}.. = e_n = 1$ and $e_j = M$. We can pair a solution $(E_1, \ldots, E_n)$ with $E_j < M$ to $(aE_j + b, E_1, ..\widehat{E_j}.., E_n)$ with $aE_j + b < N$. This describes all possible solutions.

Thus, when $x = b - a + N$, we have an odd number of terms of the form t^x, showing that $F(t, \sigma(t)) \neq 0$. This is a contradiction and we may conclude that N does not exist and therefore $\sigma(t)$ must satisfy the support relation $E = \mathscr{O} \cup (aE + b)$ with $\mathscr{O} \subset 2\mathbb{N} - 1$. □

Proposition 3.22. *Let a be a power of two and $a > b \geq \frac{-a+2}{3}$ an even integer. Then $F_{(a,b)}(t, X) := (tX)^{a-1} + X^{a-b-1} + t^{a-b-1} = 0$ has a unique solution in $\mathscr{N}_2$ of depth $a + b - 1$. This solution has order 2 and satisfies the support relation $E = \mathscr{O} \cup (aE+b)$ with $\mathscr{O} \subset 2\mathbb{N}-1$.*

Proof. By Lemma 3.9 and Corollary 3.11.1 we find the existence and uniqueness of a series $\sigma(t) \in \mathscr{N}_2$ of depth $a+b-1$, order 2 and so that $F(t, \sigma(t)) = 0$. To see this take $\bar\sigma(t) = t+t^{a+b}$ and $\rho = a - b - 1$, then

$$F'(t, \bar\sigma(t)) = t^{a-1}(t + t^{a+b})^{a-2} + (t + t^{a+b})^{a-b-2} \equiv t^{a-b-2} \not\equiv 0 \mod t^{a-b-1} \text{ and,}$$

$$F(t, \bar\sigma(t)) = t^{2a-2}+t^{3a+b-3}+O(t^{4a+2b-4})+t^{a-b-1}+t^{2a-2}+O(t^{3a+b-3})+t^{a-b-1} \equiv 0 \mod t^{3a+b-3}.$$

Since $\rho \geq a - b - 2$ and $2\rho + 1 \leq 3a + b - 3$, it works. (Actually, we only get $\sigma(t) \equiv t \mod t^{a-b}$, but, with a simple calculation, one may show that $\sigma(t) \equiv t + t^{a+b} \mod t^{a+b+1}$ as well.)

Define $\mathscr{O} = E(\sigma) \cap 2\mathbb{N} - 1$ and assume $\sigma(t)$ does not satisfy the support relation $E = \mathscr{O} \cup (aE + b)$. Then there exists a smallest even integer N such that one of the following holds:

(i) $N \in E(\sigma)$ but $N \notin E(\sigma) + b$.

(ii) $N \notin E(\sigma)$ but $N \in aE(\sigma) + b$.

Define M as $\frac{N-b}{a}$, then in case (i) $M \notin E(\sigma)$ and in case (ii)$M \in E(\sigma)$.

Note that because $\sigma(t)$ is a zero of $F(t, X)$, it is also a zero of $(tX)F(t, X) = (tX)^a + tX^{a-b} + t^{a-b}X$. When $X = \sigma(t)$ we want each term t^x to occur an even number of times. We will examine this for $x = a - b + N$. The expansion of $t\sigma(t)^{a-b}$ can only consist of monomials with odd degree, since $a - b$ is even. For the other two terms we know:

$$\begin{aligned}(t\sigma(t))^a &= \sum_{e \in E} t^{a+ae}, \text{ and} \\ t^{a-b}\sigma(t) &= \sum_{e \in E} t^{a-b+e}.\end{aligned}$$

Hence, we want there to be an even number of solutions to the problem:

$$\begin{aligned}x &= a + ae \qquad \text{with } e \in E, \text{ or} \\ x &= a - b + e' \text{ with } e' \in E.\end{aligned}$$

When $N \in E(\sigma)$ we have $e' = N$ as the only solution and when $M \in E(\sigma)$ we only have $e = M$. Thus the number of solutions is odd instead of even, a contradiction. We may conclude that N does not exist and therefore that σ must satisfy the support relation $E = \mathscr{O} \cup (aE + b)$. □

Definition 3.23. Denote the unique solution to $F_{(a,b)}(t, X) = 0$ of depth $a+b-1$ by $\sigma_{(a,b)}(t)$.

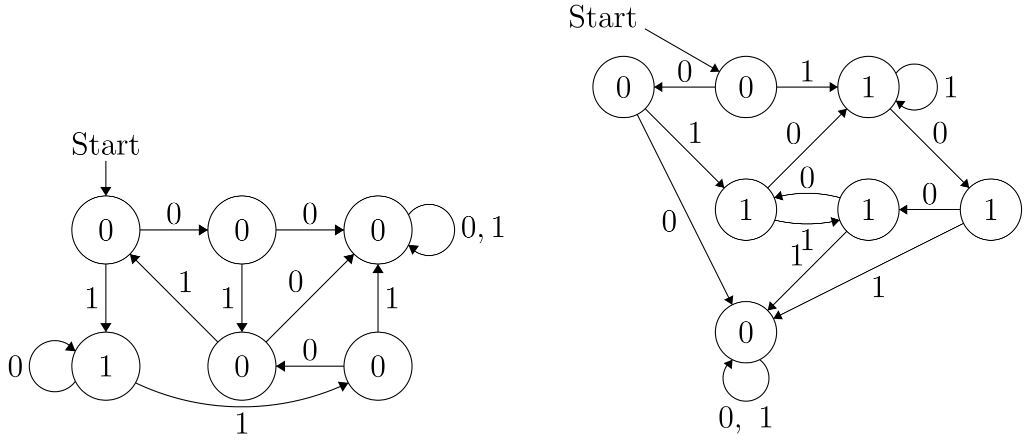
Remarks.

- The polynomials in Theorems 3.21 and 3.22 only differ by a factor $(tX)^{b-a+1}$, which is necessary to make all powers positive.
- In Proposition 3.22 the cases $\frac{-a+2}{3} > b > -a$ are excluded. This is because Hensel's Lemma is in general not able to provide the unique solution when we input $\overline{\sigma}(t) = t + t^{a+b}$. However, when $a + b - 1 \equiv 2^x - 1 \mod 2^x$ for some $x \geq 2$, we can calculate the next $2^x - 1$ monomials of $\overline{\sigma}(t)^{a-b-1}$ leading to a bigger upper bound for $2\rho+1$. One may prove that for those cases the bound for b can be improved to $b \geq \frac{-(2^x-1)(a-1)-1}{2^x+1}$. As an example take $a = 16$ and $b = -8$. We have $\frac{-a+2}{3} > b > \frac{-3a+2}{5}$ and find $a - b - 1 = 23 \equiv 3 \mod 4$. Hence $F_{(16,-8)}(t, X) = (tX)^{15} + X^{23} + t^{23} = 0$ has a unique solution in $\mathscr{N}_2$ of depth 7, order 2 and satisfying the support relation $E = \mathscr{O} \cup (16E - 8)$ with $\mathscr{O} \subset 2\mathbb{N} - 1$.
- Using Proposition 3.3 we may also improve the bound for b in the case that $v_2(b) = 1$. Defining $\overline{\sigma}(t)$ now as $t + t^{a+b} + t^{2a+2b-1}$ and noting $a - b - 1 \equiv 1 \mod 4$ one can show that $F_{(a,b)}(t, \overline{\sigma}(t)) \equiv 0 \mod t^{5a+3b-5}$. Thus any $b \geq \frac{-3a+2}{5}$ works.

Examples. As we did with the greedy series, the algebraic equations for the series satisfying support relations can be used to calculate their automata. With the help of an automaton it can then be checked if a certain series is sparse and, if it is not, what its growth numbers B and β are. For the examples that were already discussed earlier in this section, this has been done and the results and two of the (smaller) automata can be found in Table 3.3 and Figure 3.2. None of the series were sparse, unfortunately. Noteworthy is that (a, b) equal to $(4, 2)$ and $(4, 6)$ indeed have $\beta = \phi$. Also, for $(a, b) = (8, 6)$ the minimal polynomial could have been guessed, because in Table 3.2 we find the recursive relation $A_n = A_{n-1} + A_{n-3}$ when $n \geq 5$. Furthermore, in Figure 3.2a it is possible to spot the support relation $E = \mathscr{O} \cup (8E + 6)$, because when we follow the edges labeled $(0, 1, 1)$ from 'Start', we end up where we began and thus $e \in E$ if and only if $8e + 6 \in E$. The automaton in Figure 3.2b is a bit harder

(a,b)	d	Minimal polynomial of β	β	B
$(2,0)$	1	$x-2$	2	1
$(4,-2)$	1	$x^4-x^3-x^2+x-1$		≈ 0.59729
$(8,-4)$	3	$x^6-x^5-x^3+x^2-1$		≈ 0.44240
$(4,2)$	5	x^2-x-1	ϕ	≈ 0.69424
$(4,6)$	9	x^2-x-1	ϕ	≈ 0.69424
$(4,8)$	11	$x^6-x^4-x^3-x^2+1$		≈ 0.48673
$(8,6)$	13	x^3-x^2-1		≈ 0.55146
$(16,14)$	29	x^4-x^3-1		≈ 0.46496

Table 3.3: For several series that satisfy a support relation $E = \mathscr{O} \cup (aE+b)$ and are the solution to $F_{(a,b)}(t,X)=0$ as given in either Theorem 3.21 or Theorem 3.22, their depth d and their B and β.



(a)

(b)

Figure 3.2: The automata for the series which are a solution to (a) $(tX)^7 + X + t = 0$ and satisfies the support relation $E = \mathscr{O} \cup (8E+6)$, and (b) $X^5 + (tX)^3 + t^5 = 0$ and satisfies the support relation $E = \mathscr{O} \cup (4E-2)$.

to dissect this way, but also doable. Firstly, note that there are no elements divisible by 4 in the support. Therefore we can distinguish two cases for $e \in E$: either e is odd or it is 2 modulo 4. In the first case compare following the edge (1) with following $(0,1,0)$ and in the second case compare $(0,1)$ with $(0,1,1,0)$. In both cases these edge sequences lead to the same vertex and therefore, $e \in E$ if and only if $4e-2 \in E$.

Remark. In Table 3.3 the case $(a,b) = (16,14)$ has been included to illustrate a peculiarity. Whenever $a = 2^n$ and $b = a-2$, it seems the minimal polynomial for β equals $x^n - x^{n-1} - 1$. Note that these series exactly have depths $2^{n+1}-3$ and that the sparse series in Theorem 3.15 had even support of the form $\mathscr{E} = \{d+1\} \cup (a\mathscr{E}+b)$. If this is true for all n, it is remarkable that these a and b have a lot of structure in both cases.

3.4 Generalisations to $p > 2$

In this section we look which ideas that we found for $p=2$ can be generalized to other primes p.

Proposition 3.24. *Let $\sigma(t) = a_0 + a_1 t + a_2 t^2 + \ldots \in \mathbb{F}_p[\![t]\!]$ and $k \in \mathbb{N}$. Write k uniquely as $k = \sum_{i=i}^{m} r_i$ with $r_1 \leq r_2 \leq \ldots \leq r_m$ all powers of p such that each power occurs at most $p-1$ times. Then*

$$\sigma(t)^k = \sum_{e_1,\ldots,e_m \in E(\sigma)} \prod_{i=1}^{m} a_{e_i} t^{e_i r_i}.$$

Proof. Since $k = r_1 + r_2 + \ldots + r_m$, we can rewrite $\sigma(t)^k$ as

$$\sigma(t)^{r_1+r_2+\ldots+r_m} = \prod_{i=1}^{m} \sigma(t)^{r_i} = \prod_{i=1}^{m} \sigma(t^{r_i}).$$

The last equality follows because all r_i are powers of p. Expanding $\sigma(t^{r_i})$ as $\sum_{r_i \in E(\sigma)} a_{e_i} t^{e_i r_i}$ now proves the result. □

Corollary 3.24.1. *If $\sigma(t)$ has depth d, then the two smallest values in $E(\sigma(t)^k)$ are k and $R(k,d) := k + p^{v_p(k)} d$.*

Proposition 3.25. *Let $\sigma(t)$ in $\mathscr{N}_p$ have compositional order p^n and depth d with $v_p(d+1) = 0$. Then,*

(i) either $2d+1 \in E(\sigma)$, or $2d+1 \in E(\sigma^{-1})$.

(ii) if $e \in E(\sigma)$ and $1 \leq e < 2d+1$, then $e \in E(\sigma^{-1})$.

Proof. (i) Using that the inverse of $\sigma(t)$ is equal to $\sigma(t)^{\circ p^n - 1}$ it is easy to check that $\sigma^{-1}(t)$ also has depth d and order p^n. Now suppose $2d+1$ is not in the support of σ or of σ^{-1} and consider the following equation:

$$t = \sigma(\sigma^{-1}(t)) = \sum_{i \in E(\sigma)} (\sigma^{-1}(t))^i.$$

In the case that $i = 1 \in E(\sigma)$ we do not get a term of the form at^{2d+1}, but in the case $i = d+1 \in E(\sigma)$ we do get such a term, because $R(d+1,d) = 2d+1$. Any $i > d+1$ also does not give such a term, because $R(i,d) > 2d+1$ in those cases and $i \neq 2d+1$. Hence we find a contradiction and may conclude $2d+1$ is in either the support of σ or of σ^{-1}.

(ii) Let $e < 2d+1$ be the smallest element in $E(\sigma) \cup E(\sigma^{-1})$ that is not in $E(\sigma) \cap E(\sigma^{-1})$. Without loss of generality, assume e to be in $E(\sigma^{-1})$ and not in $E(\sigma)$ then, $\sigma(t) = \sigma^{-1}(t) - at^e + O(t^{e+1})$. Now consider

$$t = \sigma(\sigma^{-1}(t)) = \sum_{i \in E(\sigma)} (\sigma^{-1}(t))^i = \sigma^{-1}(t) + \sigma^{-1}(t) - t - at^e + O(t^{e+1}).$$

The last equation holds because $e < 2d+1$, so $(\sigma^{-1}(t))^i = t^i + O(t^{R(i,d)}) = t^i + O(t^{e+1})$ if $i \geq d$. Hence we get the equality $t = t + at^e + O(t^{e+1})$, which is obviously false. Therefore e cannot exist and the proof is finished. □

Remark. The previous proposition does not really tell us anything about $p = 2$, since we know that the depth d can never be divisible by p.

In the case of $p = 2$ we looked at series that satisfied a support relation of the form $E = \mathscr{O} \cup (aE + b)$. This splits up the support in odd and even elements. Similarly we might be able to split up the support for other primes into sets $\mathscr{O}_1 \cup \mathscr{O}_2 \cup \ldots \cup \mathscr{O}_p$, where $\mathscr{O}_i$ contains all elements that are congruent to i modulo p. Obviously 1 has to always be in $\mathscr{O}_1$ and, to be consistent, we might want to write the other $\mathscr{O}_i$ as $aE + b$ again. The next Lemma tells us that such a series cannot be sparse if there are multiple $i \neq 1$ such that $\mathscr{O}_i \neq \emptyset$.

Proposition 3.26. *Let $E = \mathscr{O} \cup (\alpha_1 E + \beta_1) \cup (\alpha_2 E + \beta_2) \neq \emptyset$ be a support relation with $\alpha_1, \alpha_2 > 1$. Then E is not sparse.*

Proof. Since E is nonempty there must be some element m in $\mathscr{O}$. Let $f(x) = \alpha_1 x + \beta_1$, $g(x) = \alpha_2 x + \beta_2$ and $N = \max\{f^{\circ k}(m), g^{\circ k}(m)\}$ for some $k \in \mathbb{N}$. Note that $O(\log(N)) = O(k)$ and consider the set E_N. This set must contain all elements $h_1 \circ h_2 \circ \cdots \circ h_k(m)$, where h_i is either f or g. Because $(\alpha_1 E + \beta_1)$ and $(\alpha_2 E + \beta_2)$ are disjoint, all these elements are distinct. So E_N contains at least 2^k elements and thus $|E_N| \neq O(k^r) = O(\log(N)^r)$. In other words, E is not sparse. □

From this proposition we may conclude that the only sparse series in $\mathscr{N}_p$ that satisfy a support relation must satisfy $E = \mathscr{O} \cup (aE + b)$ for some $a, b \in \mathbb{N}$ and with $1 \in \mathscr{O}$.

Another problem we run into while generalising the case $p = 2$, is that we do not have an expression for the greedy choice anymore. There might though be a 'logical choice' that can be expressed in d and d_p, the depth of the pth composite of some series. In the following we will first dissect the support relations for Klopsch's series and then check whether or not $\mathscr{O}$ could be defined as the set of 'logical choices'.

Example: Klopsch's series that satisfy a support relation

Expanding a Klopsch's series as was done in the proof of Proposition 2.15, we may conclude that its support does not depend on the coefficient $a_{d+1} \in \mathbb{F}_p^\times$. Therefore, given some prime p, we may speak about $E(\sigma_{K,d})$ without specifying which Klopsch's series of depth d is taken.

Define $\mathscr{O} = E(\sigma_{K,d}) \cap (p\mathbb{N} - (p-1))$ and $b_0 \in \{1, \ldots, p-1\}$ such that $b_0 \equiv d^{-1} \mod p$. Then for all depths $d \not\equiv 0 \mod p$, it holds that $E(\sigma_{K,d}) = \mathscr{O} \cup (\mathscr{O} + d) \cup \ldots \cup (\mathscr{O} + b_0 d)$. This follows because b_0 is the first coefficient of the p-adic expression of $\frac{-1}{d}$ and thus the congruence class of an element in the support is determined by which term we take from the first factor $(1 + t^d)^{b_0} = 1 + b_0 t^d + \ldots + t^{b_0 d}$.
In the next proposition we will answer the question: when can we write each set $(\mathscr{O} + kd)$ as $(aE + b)$?

Proposition 3.27. *Let $\sigma_{K,d}$ be a Klopsch series for some prime p and define b_0 and $\mathscr{O}$ as above. Suppose $0 < k \leq b_0$ is an integer. Then $(\mathscr{O} + kd)$ can be written as $(aE(\sigma) + b)$ if and only if $\frac{-1}{d} = \overline{b_0 0 \ldots 00}$.*

Proof. We know that the smallest element of $\mathscr{O} + kd$ is $1 + kd$ and if $i > 0$ is the smallest index such that $b_i \neq 0$, we get $1 + kd + p^i$ as the second smallest element. So $a + b = 1 + kd$ and $a(1+d) + b = 1 + kd + p^i d$. From these equations we easily solve $a = p^i$ and $b = 1 + kd - p^i$. Now let j be the smallest index such that $b_j \neq 0$ and $b_{j+i} \neq b_j$. Then we find $1 + p^j(b_j d) \in E(\sigma)$. This implies that $p^{i+j}(b_j d) + 1 + kd \in E(\sigma)$ and we conclude that $b_{j+i} > b_j$. In particular this shows that b_{j+i} is non-zero and thus $p^{i+j}(b_{i+j}d)+1+kd$ is in the support. Now we can reverse the iteration to get another element of the support:

$$\frac{1}{p^i}(p^{i+j}(b_{i+j}d)+1+kd-1-kd+p^i)=p^j(b_{i+j}d)+1.$$

This would however imply that b_j is greater than b_{i+j}, a contradiction. The only possibility is that the period of $\frac{-1}{d}$ must be divisible by i and this implies, since i was also the smallest index such that $b_i \neq 0$, that the p-adic expression must indeed be of the form $\overline{b_0 0\dots 00}$.

Conversely, assume that $\frac{-1}{d}=\overline{b_0 0\dots 00}$ of period i. Let $a=p^i$ and $b=1+kd-p^i$ and take $x\in(\mathscr{O}_1+kd)$. In this case we can express Klopsch's series (for $a_{d+1}\equiv -d^{-1}$) by

$$t\prod_{i=0}^{\infty}(1+t^d)^{b_0p^i}=t(1+t^d)^{b_0}\prod_{i=1}^{\infty}(1+t^{dp^i})^{b_0}.$$

Because $x\in(\mathscr{O}+kd)$, we can write x as $1+kd+c_1dp^i+c_2dp^{2i}+\ldots+c_ndp^{ni}$, with $c_j\in\{0,1,\ldots,b_0\}$ and therefore $\frac{x-b}{a}=1+c_1d+c_2dp^i+\ldots+c_ndp^{(n-1)i}\in(\mathscr{O}+c_1d)\subset E(\sigma_{K,d})$. From this we may conclude that $(\mathscr{O}+kd)\subset(aE(\sigma_{K,d})+b)$.
For the other inclusion we take x an element in $(aE(\sigma_{K,d})+b)$. We may write $x=1+kd-p^i+p^i(1+c_0d+c_1dp^i+\ldots+c_ndp^{ni})$, which is indeed an element of $(\mathscr{O}+kd)$. Hence $(\mathscr{O}+kd)=(aE+b)$ as we wanted. □

Corollary 3.27.1. *If $d=\frac{p^i-1}{b_0}\in\mathbb{N}$ for some $i>0$ and $b_0\in\{1,\ldots,p-1\}$, the series $\sigma_{K,d}$ satisfies a support relation.*

Examples.

- For $p=2$ these depths are 2^i-1 and they satisfy the support relation $E=\mathscr{O}\cup(2^iE)$. Here $\mathscr{O}$ consists solely of greedy choices as explained in Section 3.3.
- For $p=3$ these depths are 3^i-1 and $\frac{3^i-1}{2}$ with b_0 equal to 1 and 2 respectively.
- For $p=5$ these depths are 5^i-1, $\frac{5^i-1}{2}$, $\frac{5^i-1}{4}$ and $\frac{5^{2i}-1}{3}$ with b_0 equal to $1,2,4$ and 3 respectively.

Examples. The goal here is to find a 'logical choice' how to determine $\mathscr{O}$ for a series σ that satisfies a support relation. We use the following method. Let $\sigma_{K,d}$ be the Klopsch's series for $p=3$ with all coefficients either equal to 0 or 1 and suppose d is as in Corollary 3.27.1. Then $\sigma_{K,d}$ satisfies the support relation $\mathscr{O}\cup(\alpha_1E+\beta_1)\cup(\alpha_2E+\beta_2)$ for some $\alpha_1,\alpha_2,\beta_1,\beta_2\in\mathbb{Z}$. Write $1=x_0<x_1<\ldots$ for the elements in $\mathscr{O}$ and define $\mathscr{O}_n:=\{x_0,\ldots,x_n\}$ for all $n\geq 0$. Let $\tau_n(t)$ be the power series that satisfies the support relation $E=\mathscr{O}_n\cup(\alpha_1E+\beta_1)\cup(\alpha_2E+\beta_2)$ such that all its coefficients are either 0 or 1. Define $d_3(\tau_n)$ as the depth of $\tau_n^{\circ 3}$. How do x_{n+1},d and $d_3(\tau_n)$ relate?

- ($d=1$) We have the support relation $E=\mathscr{O}\cup(3E-1)\cup(3E)$ and find:

$$\begin{aligned}\mathscr{O}&=\{1,4,7,10,13,16,19,22,25,28,31,\ldots\},\\ d_3(\tau_n)+1&=(8,11,14,23,20,23,32,29,32,59,\ldots),\\ d_3(\tau_n)+1-x_{n+1}&=(4,4,4,10,4,4,10,4,4,28,\ldots).\end{aligned}$$

What stands out is that those are all of the form $1+3^k$ for some k and that $d_3(\tau_n)+1-1, d_3(\tau_n)+1-4, d_3(\tau_n)+1-10 \in \mathscr{O}$ for all n. However, a general structure is not clear.

- ($d=2$) We have the support relation $E = \mathscr{O} \cup (3E)$ and find:

$$\begin{aligned} \mathscr{O} &= \{1, 7, 19, 25, 55, 61, 73, 79, \ldots\}, \\ d_3(\tau_n)+1 &= (27, 69, 33, 177, 69, 93, 87, \ldots), \\ d_3(\tau_n)+1-x_{n+1} &= (20, 50, 8, 122, 8, 20, 8, \ldots). \end{aligned}$$

  Here we may note that all are $2 \mod 3$ and that any integer of the form $d_3(\tau_n)+1-8$ is in $\mathscr{O}$ but not necessarily equal to x_n.

- ($d=4$) We have the support relation $E = \mathscr{O} \cup (9E-4) \cup (9E)$ and find $x_{n+1} = d_3(\tau_n) - 15 = d_3(\tau_n)+1-4d$ for all $x_n < 3000$ (or equivalently $n < 11$).

- ($d=8$) We have the support relation $E = \mathscr{O} \cup (9E)$ and find:

$$\begin{aligned} \mathscr{O} &= \{1, 73, 649, 721, 5833, 5905, 6481, 6553, 52489, \ldots\}, \\ d_3(\tau_n)+1 &= (729, 5985, 801, 52641, 5985, 7137, 6633, 472545, \ldots). \end{aligned}$$

  Here all integers that equal $d_3(\tau_n)+1-80$ are in $\mathscr{O}$ (but not necessary equal to x_{n+1}).

It seems hard to define any ‘logical’ and consistent choices based on these examples. However, we did see that $d_3(\tau_n)+1-4d$ is always in the support and also $d_3(\tau_n)+1-10d$ might be an interesting option.

Chapter 4

The power of a Galois group

In [1] Albayrak and Bell gave a field-theoretic way to characterise sparseness. In Section 1 we state and prove one direction of the theorem. In the second section we see that by the Galois theoretic part of the characterisation we can, for all greedy series and some series satisfying a support relation (as defined in previous chapter) prove that they are not sparse. The proofs heavily rely on Newton polygons to determine the degree of the minimal polynomial.

4.1 Field theoretic conditions for sparseness

To prove the field theoretic conditions we will need the following two lemmas.

Lemma 4.1. *Let $F' = F(y)/F$ be a finite separable extension of function fields, and let $f(X) \in F[X]$ such that it is monic and $f(y) = 0$. Suppose P is a prime of F and satisfies*

$$f(X) \in \mathscr{O}_P[X] \text{ and } v_{P'}(f'(y)) = 0$$

for all primes P' of F' above P. Then P is unramified in F'/F.

Proof. This is a slight generalisation of Corollary 3.5.11 in [27]. We will assume this corollary and prove the lemma follows.

Write $\phi(X)$ for the minimal polynomial of y over F. Then we can write $f(X) = \phi(X)g(X)$ for some monic polynomial $g(X)$ in $F[X]$. We will prove that $\phi(X) \in \mathscr{O}_P[X]$ and $v_{P'}(\phi'(y)) = 0$ whenever f meets the criteria proposed in the lemma.
We can write

$$\phi(X) = X^n + a_{n-1}X^{n-1} + \ldots + a_0 \text{ and } g(X) = X^m + b_{m-1}X^{m-1} + \ldots + b_0,$$

and define $a_n = b_m = 1$ and $a_k, b_k = 0$ when $k > n$ or $k > m$ respectively. Now let $i \geq 0$ and $j \geq 0$ be the largest integers for which $v_P(a_i)$ and $v_P(b_j)$ are minimal. Whenever $\phi(X), g(X) \in \mathscr{O}_P$ we will have $i = n$ and $j = m$ and therefore $0 \geq v_P(a_i) + v_P(b_j)$. Now consider the coefficient of X^{i+j} in the expansion of f. This coefficient can be written as

$$a_{i+j}b_0 + \ldots + a_ib_j + \ldots + a_0b_{i+j}.$$

The P-adic valuation of this expression is greater or equal to the minimum of $v_P(a_kb_{i+j-k})$ for $0 \leq k \leq i + j$ with equality if and only if the minimum is unique. By the definitions of

i and j we see that the P-adic valuation must be $0 \geq v_P(a_i) + v_P(b_j) \geq 0$. Thus we may conclude that both $\phi(X)$ and $g(X)$ are in $\mathscr{O}_P[X]$.

For the second claim consider the derivative $f'(X) = \phi(X)g'(X) + \phi'(X)g(X)$. Because $\phi(y) = 0$ we get $f'(y) = \phi'(y)g(y)$ and $v_{P'}(\phi'(y)) + v_{P'}(g(y)) = 0$. Note that also $v_{P'}(y) \geq 0$, because otherwise

$$\infty = v_{P'}(0) = v_{P'}(\phi(y)) = nv_{P'}(y) < 0.$$

Using that $\phi'(X)$ and $g(X)$ are both elements of $\mathscr{O}_{P'}$ we now find that $v_{P'}(\phi'(y)), v_P(g(y)) \geq 0$, from which we conclude that $v_{P'}(\phi'(y)) = 0$. □

Lemma 4.2. *Let K be a field of characteristic $p > 0$ that contains $\mathbb{F}_{p^n}$ and take $k \in K$. The splitting field of $X^{p^n} - X + k$ over K then is Galois and has degree some power of p.*

Proof. Suppose α is a root of the polynomial and take $j \in \mathbb{F}_{p^n}$. Then, $\alpha + j$ is also a root of the polynomial, which shows that the polynomial is separable. Furthermore, $K(\alpha) = K(\alpha + j)$ and thus $K(\alpha)$ is the splitting field of $X^{p^n} - X + k$. We immediately get that $K(\alpha)$ is Galois, since it is the splitting field of a separable polynomial.
Clearly the degrees of $K(\alpha + j)$ over K for $j \in \mathbb{F}_{p^n}$ are all the same, which shows their minimal polynomials must have the same degree. Hence, the degree of $K(\alpha)$ over K must divide p^n and is thus a power of p. □

Proposition 4.3 (Albayrak–Bell [1]). *Let $\sigma \in \mathscr{N}_p$ denote a power series that is algebraic over $\mathbb{F}_p(t)$. Consider the field*

$$\mathscr{F} = \bigcup_{p \nmid \ell \geq 1} \overline{\mathbb{F}}_p(t^{1/\ell}),$$

where $\overline{\mathbb{F}}_p$ is an algebraic closure of $\mathbb{F}_p$. If σ is sparse, then the following conditions hold:

(i) σ is integral over $\overline{\mathbb{F}}_p[t, t^{-1}]$;

(ii) the extension $\overline{\mathbb{F}}_p(t)(\sigma)/\overline{\mathbb{F}}_p(t)$ is unramified outside of $0, \infty$;

(iii) the splitting field of σ over $\mathscr{F}$ has degree a power of p.

Proof. Whenever σ is sparse its support can be written as a finite disjoint union of simple sparse sets as we saw in Lemma 2.5. Thus, we may then also view σ as the sum of multiple series, that all have support equal to a simple sparse set. We will call these series simple sparse series. With an inductive argument we will show that it suffices that properties (i)-(iii) hold for simple sparse series, after which we prove that this is indeed the case. Notice that these simple sparse sets do not necessarily contain 1 and therefore the proof needs to work for all series in $\mathbb{F}_p[\![t]\!]$ and not only for those in $\mathscr{N}_p$.

Write $\sigma = \sigma_1 + \ldots + \sigma_r$, where all σ_i are simple spares series and suppose properties (i)-(iii) hold for all σ_i. Property (i) clearly also holds for σ, since the sum of a finite number of integral elements is also integral. For the second property we use Corollary 3.9.3 from [27]. We know that

$$\overline{\mathbb{F}}_p(t)(\sigma) \subset \prod_{i=1}^{r} \overline{\mathbb{F}}_p(t)(\sigma_i),$$

where $\prod$ is used to denote taking the compositum of fields, not the product. The corollary implies that the right side is unramified outside of 0 and ∞, because this is the case for each

individual field. The same then holds for any subfield.
Finally, for the third property, suppose we have proven that the splitting fields of $\sigma_1 + \ldots + \sigma_{r-1}$ and of σ_r over $\mathscr{F}$ have as degree a power of p. Write F_1 and F_2 for these splitting fields respectively and define $F := F_1 \cap F_2$. The following degrees are now all powers of p: $[F_2 : F], [F_1 : F]$ and $[F : \mathscr{F}]$. By Proposition 14.4.19 in [16], we know F_1F_2 is Galois and

$$[F_1F_2 : \mathscr{F}] = [F_1 : F][F_2 : F][F : \mathscr{F}],$$

which is thus also a power of p. The field $\mathscr{F}(\sigma)$ is a subfield of F_1F_2 and because F_1F_2 is Galois, so is the Galois closure of $\mathscr{F}(\sigma)$. With an inductive argument this shows σ meets property (iii) when all simple sparse series do.

Now we will prove that all three properties hold for some simple sparse series σ. We know that $E(\sigma) = \{v_r w_r^{x_r} \cdots v_1 w_1^{x_1} v_0 \mid x_i \in \mathbb{Z}_{\geq 0}\}$ for some fixed p-ary words $v_0, \ldots, v_r, w_1, \ldots, w_r$. When $r = 0$, σ is a monomial and thus all properties hold. Now suppose it holds for all ranks up to $r - 1 \geq 0$. Let $k_0 = \ell(v_0), k_1 = \ell(w_1), m_0 = |v_0|$ and $m_1 = |w_1|$. Define τ to be the simple sparse series with support equal to $\{v_r w_r^{x_r} \cdots w_2^{x_2} v_1 0^{k_0} \mid x_i \in \mathbb{Z}_{\geq 0}\}$. This series has rank $r-1$, so by the induction hypothesis the properties hold for τ. We will now prove the following relation between σ and τ:

$$t^{(p^{k_1}-1)m_0 - p^{k_0}m_1}\sigma(t) - \sigma(t)^{p^{k_1}} = t^{p^{k_1}m_0 - p^{k_0}m_1}\tau(t).$$

Note that we may write the support of σ as:

$$\{p^{k_1x_1}j + (p^{k_1(x_1-1)+k_0} + \ldots + p^{k_1+k_0} + p^{k_0})m_1 + m_0 \mid x_1 \in \mathbb{Z}_{\geq 0}, j \in E(\tau)\}.$$

Hence we can write the support of the first term on the left side of the equation as:

$$\{p^{k_1x_1}j + (p^{k_1(x_1-1)+k_0} + \ldots + p^{k_1+k_0} + p^{k_0})m_1 - p^{k_0}m_1 + p^{k_1}m_0 \mid x_1 \in \mathbb{Z}_{\geq 0}, j \in E(\tau)\}.$$

Similarly, the second term of the equation has support:

$$\{p^{k_1x_1}j + (p^{k_1(x_1-1)+k_0} + \ldots + p^{k_1+k_0})m_1 + p^{k_1}m_0 \mid x_1 \in \mathbb{Z}_{\geq 1}, j \in E(\tau)\}.$$

Which are exactly the same, except for the fact that in the first set $x_1 \geq 0$ and in the second $x_1 \geq 1$. Furthermore, for each n in both supports the coefficient $a_n \neq 0$ in front of t^n will be the same. So, if we subtract the second term from the first, only the cases with $x_1 = 0$ will remain. This gives the support $\{j - p^{k_0}m_1 + p^{k_1}m_0 \mid j \in E(\tau)\}$, which is equal to the support of the (only) term on the right. Here a similar remark about equal coefficients holds and we conclude that the relation between σ and τ indeed holds.

Because this relation is integral, it now immediately follows that σ is integral over $\overline{\mathbb{F}}_p(\tau)[t, t^{-1}]$ and since property (i) already holds for τ it now also holds for σ.

The relation can also be used to prove the second property. Use Lemma 4.1 with the function $f(t, X) = X^{p^{k_1}} - t^{(p^{k_1}-1)m_0 - p^{k_0}m_1}X + t^{p^{k_1}m_0 - p^{k_0}m_1}\tau(t)$. For any prime P outside of 0 and ∞ and P' above P, we see that $f(t, X) \in \mathscr{O}_P[X]$ if and only if $v_P(\tau(t)) \geq 0$ and $v_{P'}(f'(\sigma)) = v_{P'}(t^{(p^{k_1}-1)m_0 - p^{k_0}m_1}) = 0$. Hence, P is unramified if P does not lie above 0 or ∞ and $v_P(\tau(t)) \geq 0$. The last condition always holds by the following inductive argument. If the rank of τ is zero (and therefore τ is a monomial) it is clear and else the induction hypothesis tells us that there exists some polynomial $g(X) \in \mathscr{O}_P[X]$ such that $g(\tau) = 0$. As is shown in the proof of Lemma 4.1 this implies $v_P(\tau) \geq 0$.

Defining $c = -m_0 + \frac{p^{k_0}m_1}{p^{k_1}-1}$ and $d = \frac{p^{k_0}m_1}{p^{k_1}-1}$ we can rewrite the relation between σ and τ to:

$$(t^c\sigma)^{p^{k_1}} - (t^c\sigma) = t^d\tau.$$

Hence, $t^c\sigma$ is the zero of a polynomial of the form $X^{p^n} - X + k$ over $\mathscr{F}(\tau)$, as in Lemma 4.2. Since $\mathscr{F}(\tau, \sigma)$ is equal to the splitting field of this polynomial (also see the proof of the lemma), it is Galois and has degree a power of p over $\mathscr{F}(\tau)$. Furthermore, we already assume that $\mathscr{F}(\tau)$ has a degree over $\mathscr{F}$ that is power of p. This shows that the third property must indeed hold for $\mathscr{F}(\sigma)$. □

Going back to the series of order 2 that we found in Chapter 3, we will check if they meet the criteria of Proposition 4.3. Recall that we defined two types of series: greedy series and series that satisfied a support relation $E = \mathscr{O} \cup (aE + b)$. By the algebraic relations we found for greedy series and if a is a power of two, it is immediately clear that they all meet the first criterium. About the other two we will say something in the next section on the basis of their Newton polygons.

4.2 Newton polygons

Definition 4.4. Let F be a non-archimedean field that is complete with respect to some valuation v and $f(X) \in F[X]$ equal to

$$a_nX^n + \ldots + a_1X + a_0, \text{ with } a_n \neq 0.$$

We define the *Newton Polygon of* f, or NP(f), to be the lower convex hull in $\mathbb{R}^2$ of the finite set of points

$$\{i, v(a_i) \mid 0 \leq i \leq n, a_i \neq 0\}.$$

For each line segment we can determine its *slope* and *multiplicity*, i.e. the length of its projection to the horizontal axis.

Since each slope is unique, it makes sense to talk about the multiplicity of a certain slope.

Example. Take the field $\mathbb{F}_2[\![t]\!]$, which is complete with respect to (t), and consider the polynomial $F_{(8,6)}(t, X) = (tX)^6 + X + t$. The set of points becomes $\{(0, 1), (1, 0), (6, 6)\}$ and in Figure 4.1 we can now see its complex hull. It consists of two line segments, one of slope -1 and multiplicity 1 and one of slope $\frac{6}{5}$ and multiplicity 5.

Theorem 4.5 ([2] p. 8)**.** *Let $f \in F[x]$, and $\Lambda \subset \mathbb{Q}$ the set of slopes of* NP(f)*. We may write*

$$f = a_n \prod_{\lambda \in \Lambda} f_\lambda, \text{ with, } \quad f_\lambda(x) := \prod_{\substack{f(y)=0,\\ v(y)=-\lambda}} (x - y).$$

Then, $f_\lambda(x) \in F[x]$, and its degree is equal to the multiplicity of λ.

Theorem 4.6 (Dumas [15])**.** *Let (a, b) and $(a + c, b + d)$ be the start and end point of some line segment of* NP(f) *with slope $\lambda = \frac{d}{c}$. Suppose there exist $g_\lambda(x), h_\lambda(x) \in F[x]$ of degrees d_g and d_h such that*

$$f_\lambda(x) = g_\lambda(x)h_\lambda(x)$$

then, the point $(a + d_g, b + \lambda d_g)$ must lie in $\mathbb{Z}^2$.

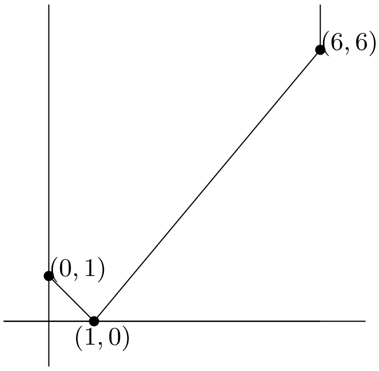


Figure 4.1: The Newton polygon of $F(t, X) = X^6 + X + t$ at the prime (t).

Example. In the case that $F(t, X) = (tX)^6 + X + t$ we can now conclude that $F(t, X)$ has one zero with t-valuation 1 and 5 zeroes with t-valuation $\frac{6}{5}$. Also, we can write the polynomial as $F(t, X) = t^6(F_{-1}(t, X)F_{\frac{6}{5}}(t, X))$, where $F_{-1}(t, X)$ has degree 1 and $F_{\frac{6}{5}}(t, X)$ has degree 5 in X. Both these polynomials are irreducible, since the polygon does not cross any additional points on the lattice $\mathbb{Z}^2$.

Proposition 4.7. *All greedy series and support relations series with a equal to a power of 2 meet the second criterium of 4.3.*

Proof. Let P be some prime in $\overline{\mathbb{F}}_p(t)$ besides 0 and ∞, σ the series in question and define F as $F_{G,d}$ or $F_{(a,b)}$ divided by their lead coefficient. Then $F(t, X) \in \mathscr{O}_P[X]$ and $F'(t, \sigma) = \sigma^n$ for some $n \in \mathbb{Z}$. We will show that $v_{P'}(\sigma) = 0$ for all $P'|P$ and then it follows from Lemma 4.1 that P is unramified.
Note that the Newton polygon of $F(t, X)$ at P consists solely of a straight line segment on the x-axis. Thus, by Lemma 4.5, any zero of $F(t, X)$ has P-adic valuation 0. The series σ is a zero of $F(t, X)$ and since P' lies above P this shows $v_{P'}(\sigma) = 0$. □

Theorem 4.8. *Suppose a is a power of two, b an even integer, $d > 0$ an odd integer and $\ell > 0$ a square free integer. Define $G(t, X) := F_{G,d}(t, X)$ and $F(t, X) := F_{(a,b)}(t, X)$, then:*

(i) G is irreducible over $\overline{\mathbb{F}}_2(t^{1/\ell})$.

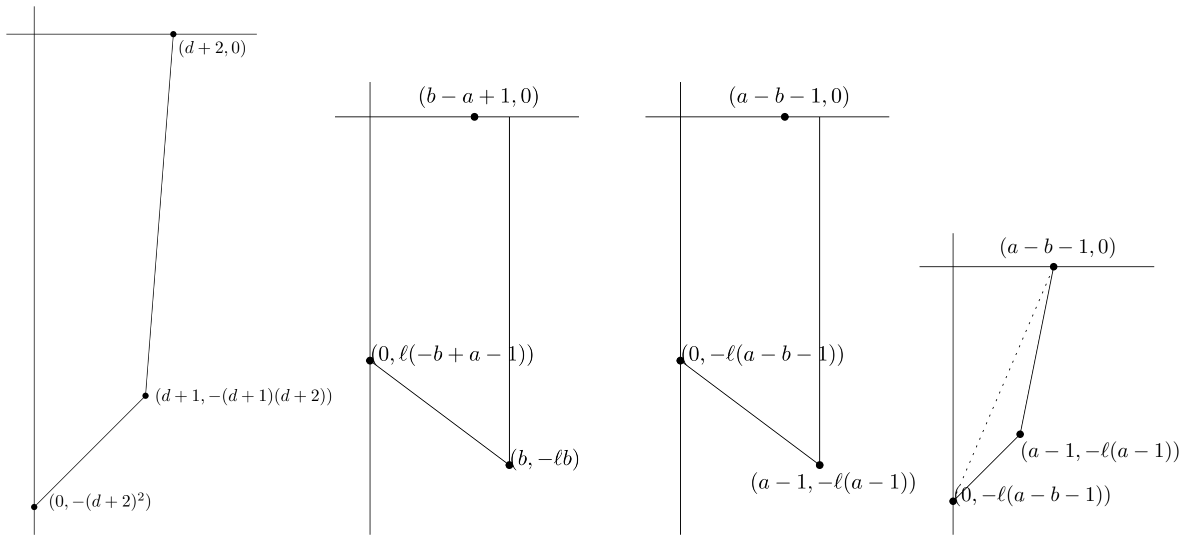
(ii) When $b \geq a$ all irreducible factors of F over $\overline{\mathbb{F}}_2(t^{-1/\ell})$ have a degree that is a multiple of $\frac{b}{\gcd(\ell(a-1),b)}$.

(iii) When $a > b > 0$ all irreducible factors of F over $\overline{\mathbb{F}}_2(t^{-1/\ell})$ have a degree that is a multiple of $\frac{a-1}{\gcd(a-1,\ell b)}$.

(iv) When $0 > b > -a$ all irreducible factors of F over $\overline{\mathbb{F}}_2(t^{1/\ell})$ have a degree that is either a multiple of $\frac{b}{\gcd(\ell(a-1),b)}$ or a multiple of $\frac{a-1}{\gcd(a-1,\ell b)}$.

Proof. Note that all irreducible factors of G or F must lie in $\overline{\mathbb{F}}_2[t^{1/\ell}, X]$ (case (i) and (iv)) or $\overline{\mathbb{F}}_2[t^{-1/\ell}, X]$ (case (ii) and (iii)) because G and F, when divided by their lead coefficients, themselves do. (Also see the argument given in the proof of Lemma 4.1.)

(i) We start by looking at the Newton Polygon of F at the prime $t^{-1/(d+2)}$, see Figure 4.2a. The polygon has two line segments, one from $(0, -(d+2)^2)$ to $(d+1, -(d+1)(d+2))$



(a) (b) (c) (d)

Figure 4.2: The Newton Polygons at $t^{-1/\ell}$ of (a) $F(t,X) = X^{d+2}+(tX)^{d+1}+t^{d+2}$ for $\ell = d+2$, (b)$F(t,X) = (tX)^b + X^{b-a+1} + t^{b-a+1}$ when $b \geq a$, (c) $F(t,X) = (tX)^{a-1} + X^{a-b-1} + t^{a-b-1}$ when $a > b > 0$, and (d) $F(t,X) = (tX)^{a-1} + X^{a-b-1} + t^{a-b-1}$ when $0 > b > -a$.

and the other from $(d + 1, -(d + 1)(d + 2))$ to $(d + 2, 0)$. Neither crosses any other integer points and their multiplicities are $d + 1$ and 1. Therefore, the only way that F could be reducible, is if it has a zero $q(t)$ in $\mathbb{F}_2[t^{1/(d+2)}]$ of order $v_{t^{-1/(d+2)}}(q) = -(d+1)$. So the term of highest degree of $q(t)$ should be $a_{\frac{d+1}{d+2}}t^{(d+1)/(d+2)}$. However, when looking at the Newton Polygon of F at the prime t, we find that all its zeroes have $v_t(q(t)) = 1$ and thus, the term of lowest degree should be a_1t. This shows that q could not be a polynomial in any fractional power of t and thus F stays irreducible over $\overline{\mathbb{F}}_2[t^{1/(d+2)}]$. Now write g for the largest divisor of ℓ that is coprime with $d + 2$. Then,

$$[\overline{\mathbb{F}}_2(t^{1/\ell}, t^{1/(d+2)}) : \overline{\mathbb{F}}_2(t^{1/(d+2)})] = g.$$

Since g and $d + 2$ are coprime and extending $\overline{\mathbb{F}}_2(t^{1/(d+2)})$ with σ has a degree of $d + 2$, we find that F stays irreducible over $\overline{\mathbb{F}}_2(t^{1/\ell}, t^{1/(d+2)})$ and thus over all fields $\overline{\mathbb{F}}_2(t^{1/\ell})$.

(ii) In Figure 4.2b the Newton Polygon of F at $t^{-1/\ell}$ is depicted. It has one line segment of degree b and slope $\frac{-\ell(a-1)}{b}$. Any point (x, y) on the line is in $\mathbb{Z}^2$ if x is divisible by $\frac{b}{\gcd(-\ell(a-1),b)}$. The statement now follows directly from Theorem 4.6.

(iii) The Newton Polygon of F at $t^{-1/\ell}$ now has a single line segment of degree $a - 1$ and slope $\frac{-\ell b}{a-1}$, see Figure 4.2c. Therefore, (x, y) is an integer point on the segment if x is divisible by $\frac{a-1}{\gcd(a-1,-\ell b)}$.

(iv) The Newton Polygon of F at $t^{-1/\ell}$ can be seen in Figure 4.2d. Note that since $b > -a$ the point $(a-1, a-1)$ will always lie on the convex side of the dotted line. The polygon has two line segments, one of degree $a - 1$ and slope $\frac{-\ell b}{a-1}$ and one of degree $-b$ and slope $\frac{\ell(a-1)}{-b}$. The statement now follows by the same ideas an in (ii) and (iii). □

Corollary 4.8.1. *(i) All greedy series are not sparse.*

(ii) Suppose $\frac{b}{\gcd(a-1,b)}$ *(when* $b \geq a$*),* $\frac{a-1}{\gcd(a-1,b)}$ *(when* $a > b > 0$*) or both (when* $0 > b > -a$*) are divisible by some odd prime squared. Then the series of order* 2 *that satisfies the support relation* $E = \mathscr{O} \cup (aE + b)$ *is not sparse.*

Proof. By Proposition 4.3 it suffices to show that the degree of the splitting field of σ over $\mathscr{F}$ is not a power of 2.

(i) The minimal polynomial of $\sigma_{G,d}$ over $\mathscr{F}$ is just $F_{G,d}$. It follows that $[\mathscr{F}(\sigma_{G,d}) : \mathscr{F}] = d + 2$ which is odd and larger than 1.

(ii) Let q be the prime that divides the relevant constant twice. Then q must divide $[\mathscr{F}(\sigma_{(a,b)}) : \mathscr{F}]$ by a similar argument. □

Proposition 4.9. *Suppose* $a - b - 1$ *is equal to* 3 *and* $a > b > 0$*. Then* $\sigma_{(a,b)}$ *is not sparse.*

Proof. In these cases the Newton polygon at $t^{1/\ell}$ consists of two line segments. A segment of slope $-\ell$ and degree 3 and a segment of slope $\lambda(\ell) := \frac{\ell(a-1)}{a-4}$ and degree $a - 4$. Since $v_t(\sigma_{(a,b)}) = 1$ we know that it is a zero $F_\ell(t, X)$ and not of $F_{-\lambda(\ell)}(t, X)$. One can prove that there is no ℓ such that F_ℓ has a zero in $\overline{\mathbb{F}}_2[t^{-1/\ell}]$. Besides, all irreducible factors of $F_{-\lambda(\ell)}$ in $\overline{\mathbb{F}}_2[\![t^{1/\ell}]\!]$ must have a degree that is divisible by $\frac{a-4}{\gcd(\ell(a-1),a-4)}$. Hence, the minimal polynomial of σ over $\mathscr{F}$ is of degree $3 + x$, where x is some multiple of $\frac{a-4}{\gcd(\ell(a-1),a-4)}$ for some odd and square free integer ℓ. Whenever ℓ is odd we find that $3 + x$ is also odd. Therefore, the minimal polynomial of $\sigma_{(a,b)}$ of $\mathscr{F}$ will be of odd degree greater or equal to 3 and $\sigma_{(a,b)}$ cannot be sparse by Proposition 4.3(iii). □

Examples. Especially the cases where $\frac{b}{\gcd(a-1,b)}$ is a power of two and $b \geq a$ or $0 > b$, seem good candidates for sparse series. In fact, we have seen examples of both answers.

- When $a = 2$ and $b = 2^\mu$ or $a = b = 2^{\mu-1}$ we get the series $\sigma_{S,2^\mu \pm 1}$, which is sparse. Note that $b \geq a$ here and $\frac{b}{\gcd(a-1,b)}$ equals 2^μ or 1 respectively.
- When $(a, b) \in \{(4, -2), (8, -4), (4, 8)\}$ we saw at the end of Section 3.3 that σ is not sparse. In all these cases $\frac{b}{\gcd(a-1,b)} = b$ which is (up to a minus sign) a power of 2.
- When $(a, b) = (4, 6)$ we also saw at the end of Section 3.3 that σ is not sparse. In this case $b \geq a$ and $\frac{b}{\gcd(a-1,b)} = 2$.

Remark. We have not discussed any of the cases where $b = 0$ here. However, in Section 3.4 we saw that those all generate Klopsch's series, which we know to not be sparse.

Chapter 5

Further research

(1) In Chapter 2 a new growth constant β was defined for non-sparse series σ. We know that β is bounded by p and always equals the nth root of a Perron number. A natural question to ask oneself is: given a certain prime p, does this describe all possible values β can take? What about when we know the compositional order of σ? A known problem in automaton theory is whether you can see what the order of an automatic series is by only looking at its automaton. If the set of values β that one can take in the finite order case is unequal to the set of values in the infinite order case, this gives a partial answer to this problem.

(2) In Chapter 3 a new method to search for order 2 power series was described, by guessing the subset of even elements of the support. So far there has not been found a non-empty even set for which such a series does not exist. It would therefore be interesting to try and prove or disprove the following conjecture:

   Conjecture 5.1. *Let $\mathscr{E} \subset 2\mathbb{N}$ non-empty, then there exists exactly one $\sigma \in \mathscr{N}_2$ such that $E(\sigma) \cap 2\mathbb{N} = \mathscr{E}$ and σ is of compositional order* 2.

   The proof of this conjecture might also lead to more insight in how the odd part of the support depends on the even part, which could help us with finding more sparse series of order 2. Even if this does not (yet) work, the new method can be used to guess more sparse even sets and see if these result in any sparse series. This is of special interest for all depths outside of $2^{\mu} \pm 1$ and $2^{\mu} - 3$.

(3) In Section 3.4 we saw that it is not completely straightforward how to generalise the method to $p = 3$. However, the way $\mathscr{O}$ was constructed for the Klopsch's series did not seem entirely without structure. Similarly there might exist structures for series of order p^n with $n \geq 2$, starting with 4.

(4) In the beginning of Chapter 4 a field theoretic characterisation of sparse series by Albayrak and Bell was mentioned. One way to find series σ of finite order is by using Witt vectors and choosing a good uniformizer, see [8]. This choice also influences field theoretic properties of σ. It would be nice to see if these two ideas could be combined to see which uniformizers might or will definitely not lead to sparse series. We saw that so far especially the third property of Proposition 4.3 was of importance, so a good place to start is by calculating the splitting field over $\mathscr{F}$ for a known sparse series.

Bibliography

[1] Seda Albayrak and Jason P. Bell. “A refinement of Christol’s theorem for algebraic power series”. In: *Mathematische Zeitschrift* 300.3 (2022), pp. 2265–2288.

[2] Matthew Baker and Oliver Lorscheid. “Descartes' rule of signs, Newton polygons, and polynomials over hyperfields”. In: *Journal of Algebra* 569 (2021), pp. 416–441.

[3] José Bertin and Ariane Mézard. “Déformations formelles des revêtements sauvagement ramifiés de courbes algébriques”. In: *Inventiones mathematicae* 141.1 (2000), pp. 195–238.

[4] Frauke M. Bleher, Ted Chinburg, Bjorn Poonen, and Peter Symonds. “Automorphisms of Harbater–Katz–Gabber curves”. In: *Mathematische Annalen* 368.1-2 (2016), pp. 811–836.

[5] Alin Bostan, Xavier Caruso, Gilles Christol, and Philippe Dumas. “Fast coefficient computation for algebraic power series in positive characteristic”. In: *The Open Book Series* 2.1 (2019), pp. 119–135.

[6] Andrew Bridy. “Automatic sequences and curves over finite fields”. In: *Algebra and Number Theory* 11.3 (2017), pp. 685–712.

[7] Andrew Bridy, Gunther Cornelissen, and Onno van Zomeren. *Solving polynomial equations in power series using automata: algorithms, bounds and experiments.* Master Thesis with algorithm based on previous article. 2020.

[8] Jakub Byszewski, Gunther Cornelissen, and Djurre Tijsma. “Automata and finite order elements in the Nottingham group”. In: *Journal of Algebra* 602 (2022), pp. 484–554.

[9] Jakub Byszewski and Jakub Konieczny. “Automatic sequences and generalised polynomials”. In: *Canadian Journal of Mathematics* 72.2 (2020), pp. 392–426.

[10] Rachel Camina. “Subgroups of the Nottingham group”. In: *Journal of Algebra* 196.1 (1997), pp. 101–113.

[11] Rachel Camina. “The Nottingham group”. In: *New Horizons in pro-p Groups.* Ed. by Marcus du Sautoy, Dan Segal, and Aner Shalev. Vol. 184. Boston, MA: Birkhäuser Boston, 2000, pp. 205–221.

[12] Ted Chinburg and Peter Symonds. *An element of order 4 in the Nottingham group at the prime 2.* 2010.

[13] Gilles Christol. “Ensembles presque periodiques k-reconnaissables”. In: *Theoretical Computer Science* 9.1 (1979), pp. 141–145.

[14] Alan Cobham. “Uniform tag sequences”. In: *Math. Systems Theory* 6 (1972), pp. 164–192.

[15] Gustave Dumas. “Sur quelques cas d’irréductibilité des polynomes à coefficients rationnels”. In: *Journal de Mathématiques Pures et Appliquées* 2 (1906), pp. 191–258.

[16] David S. Dummit and Richard M. Foote. *Abstract Algebra.* third ed. New York: Wiley, 2004.

[17] David Eisenbud. *Commutative Algebra with a View Toward Algebraic Geometry.* first ed. Vol. 150. Graduate Texts in Mathematics. Berlin and New York: Springer Verlag, 1995.

[18] Mikhail Ershov. "On the commensurator of the Nottingham group". In: *Transactions of The American Mathematical Society* 362.12 (2010), pp. 6663–6678.

[19] Sandrine Jean. *Classification à conjugaison près des séries de p-torsion.* thèse de doctorat. 2008 (last accessed July 2022). URL: `http://aurore.unilim.fr/ori-oai-search/notice/view/unilim-ori-24991`.

[20] David L. Johnson. "The group of formal power series under substitution". In: *Journal of the Australian Mathematical Society. Series A. Pure Mathematics and Statistics* 45.3 (1988), 296—302.

[21] Benjamin Klopsch. "Automorphisms of the Nottingham Group". In: *Journal of Algebra* 223.1 (2000), pp. 37–56.

[22] Serge Lang. *Algebra.* third ed. Vol. 211. Graduate Texts in Mathematics. New York: Springer Verlag, 2002.

[23] Douglas Lind. "The entropies of topological Markov shifts and a related class of algebraic integers". In: *Ergodic Theory and Dynamical Systems* 4.2 (1984), 283—300.

[24] Jonathan Lubin. "Torsion in the Nottingham group". In: *Bulletin of the London Mathematical Society* 43.3 (2011), pp. 547–560.

[25] Carl Meyer. *Matrix Analysis and Applied Linear Algebra.* Philadelphia: SIAM, 2000.

[26] Marcus du Sautoy and Ivan Fesenko. "Where the wild things are: ramification groups and the Nottingham group". In: *New Horizons in pro-p Groups.* Ed. by Marcus du Sautoy, Dan Segal, and Aner Shalev. Vol. 184. Boston, MA: Birkhäuser Boston, 2000, pp. 287–328.

[27] Henning Stichtenoth. *Algebraic Function Fields and Codes.* second ed. Vol. 254. Graduate Texts in Mathematics. Berlin: Springer-Verlag, 2009.

[28] Andrew Szilard, Sheng Yu, Kaizhong Zhang, and Jeffrey Shallit. "Characterizing regular languages with polynomial densities". In: *Mathematical Foundations of Computer Science 1992.* Ed. by Ivan M. Havel and Václav Koubek. Berlin, Heidelberg: Springer Berlin Heidelberg, 1992, pp. 494–503.

[29] Evan Wallace. *Finite State Machine Designer.* 2010 (last accessed July 2022). URL: `http://madebyevan.com/fsm/`.

Appendix

Here we will give the automata, algebraic equations and/or direct formulas of some non-sparse series of finite compositional order. All the automata were calculated in [8] and drawn using the Finite State Machine Designer [29]. When not stated otherwise, the series were also defined by J. Byszewski, G. Cornelissen and D. Tijsma in [8].

The series $\sigma_{\min}$ is defined as the unique zero in $\mathscr{N}_2$ of

$$F(t,X) = (t+1)^3X^3 + (t^3+t)X^2 + (t^3+t+1)X + t^3 + t.$$

It has order 4 and depth 1 and its automaton can be found in Figure 5.2a.

The series $\sigma_{(1,5)}$ satisfies the equation $t^2X^3 + (t+1)^3X + t^3 + t = 0$ and is of order 4 and depth 1. Its automaton can be found in Figure 5.4b.

The series $\sigma_{V,1}$ is a zero of $t^4X^4 + t^3X^3 + X^2 + (t+1)X + t^2 + t$ and $\sigma_{V,2}$ is a zero of $(t^4+1)X^4 + tX^2 + t^2X + t^4$. Then $\sigma_{V,3}$ is defined as $\sigma_{V,1} \circ \sigma_{V,2}$ and satisfies

$$t^4X^4 + (t+1)^3X^3 + t(t^2+t+1)X^2 + (t+1)^3X + t(t+1)^2 = 0.$$

All have order 2 and they have depths $1, 5$ and 1 respectively. They also commute with each other, and hence exhibit an explicit embedding of the Klein group. Their automata can be found in Figure 5.5, 5.6 and 5.1.

The series σ_{CS} was discovered by T. Chinburg and P. Symonds [12] and has order 4. Its compositional inverse, $\sigma_{CS}^{\circ 3}$ was computed by Z. Scherr and M. Zieve [4, Remark 1.5], and is in fact sparse.

$$\begin{aligned}
\sigma_{CS} &:= t + t^2 + \sum_{k\geq 0}\sum_{\ell=0}^{2^k-1} t^{6\cdot 2^k+2\ell},\\
\sigma_{CS}^{\circ 2} &= t + \sum_{k\geq 0}\sum_{\ell=0}^{2^k-1} t^{4\cdot 2^k+2\ell}.\\
\sigma_{CS}^{\circ 3} &= \sum_{k\geq 0}(t^{3\cdot 2^k-2} + t^{4\cdot 2^k-2}).
\end{aligned}$$

Their depths are $1, 3$ and 1, respectively and the automata of σ_{CS} and $\sigma_{CS}^{\circ 2}$ can be found in Figure 5.3.

Found by S. Jean [19] and a zero of $F(t, X) = (t+1)X^2 + (t^2+1)X + t$ we have the series σ_J of order 4 and its compositional inverse:

$$\sigma_J := \sum_{k\geq 0} \frac{t^{2^k}}{(t+1)^{3\cdot 2^k - 1}},$$
$$\sigma_J^{\circ 3} = \sum_{k\geq 0} \frac{t^{2^{k+1}-1}}{(t+1)^{3\cdot 2^k - 2}}.$$

They both have depth 1 and the automata can be found in Figures 5.2b and 5.4a, respectively.

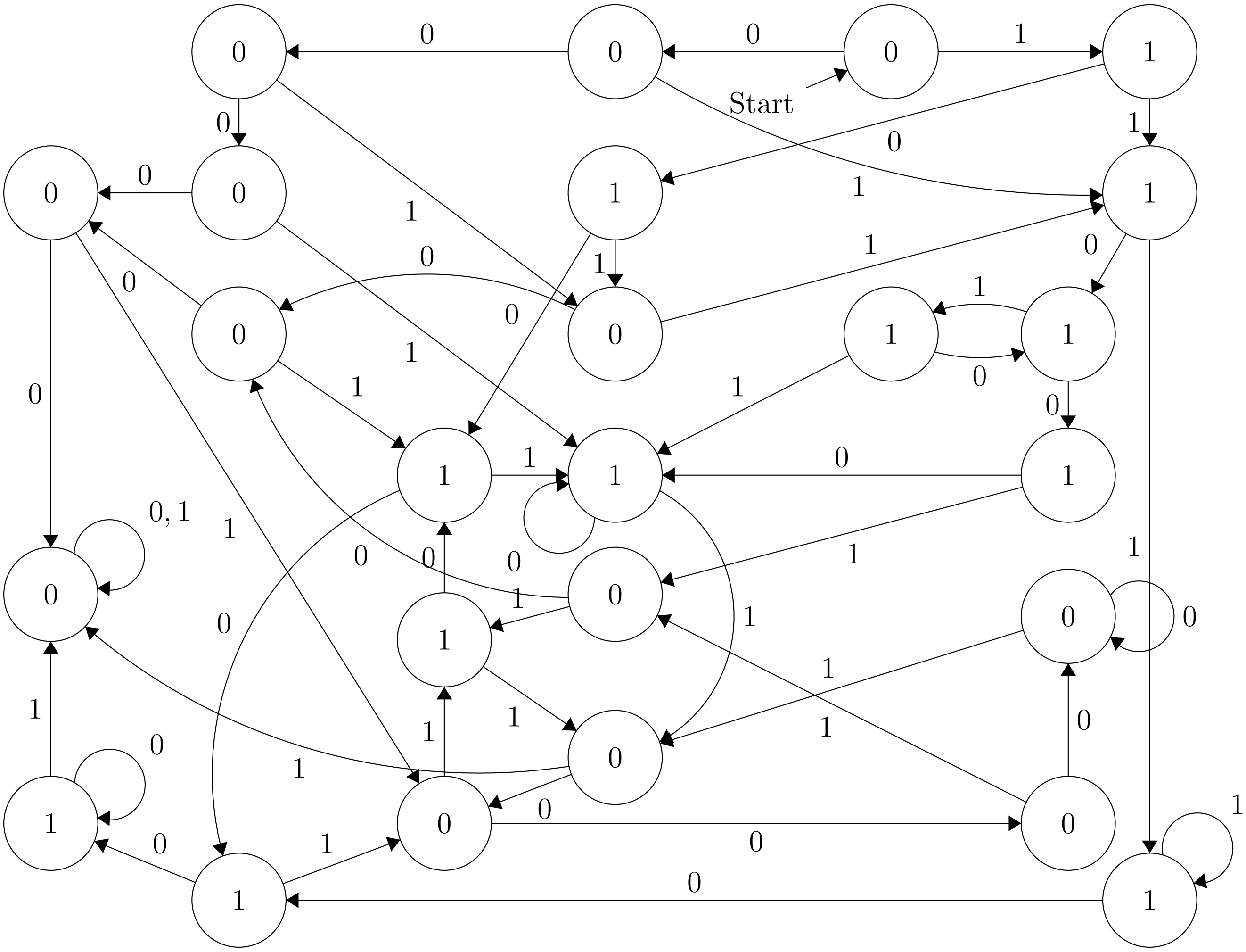


Figure 5.1: The 2-automaton of $\sigma_{V,3}$

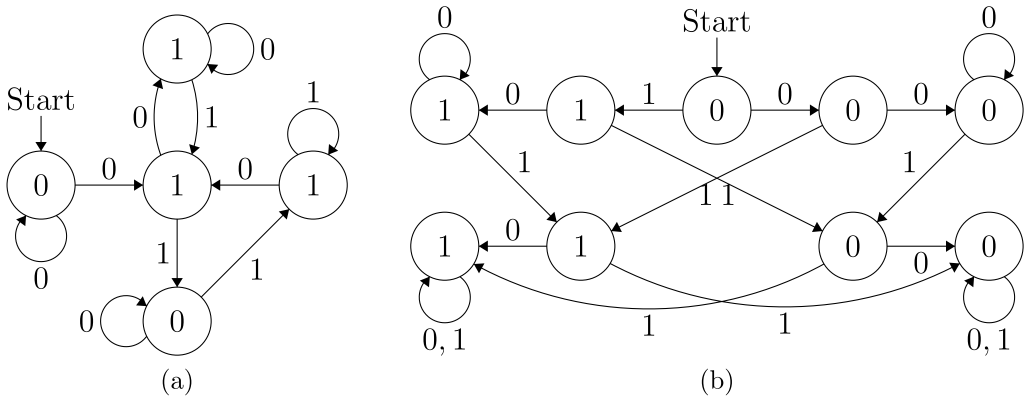


Figure 5.2: The 2-automata of (a) $\sigma_{\min}$ and (b) σ_J.

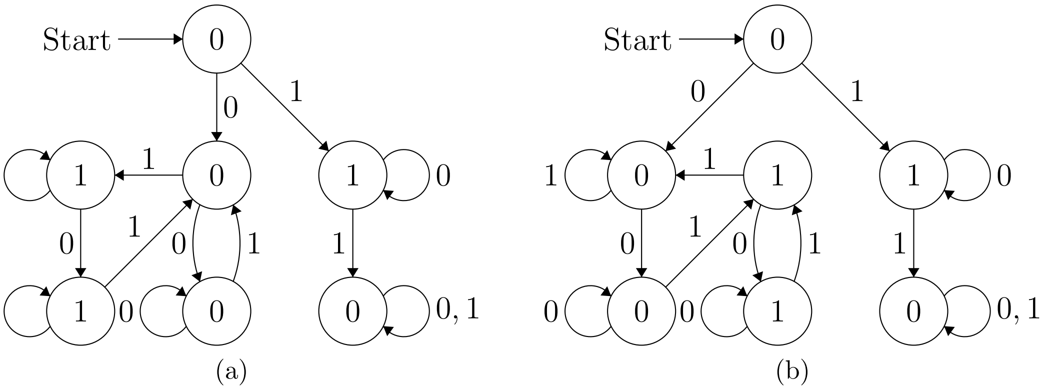


Figure 5.3: The 2-automata of (a) σ_{CS} and (b) $\sigma_{CS}^{\circ 2}$.

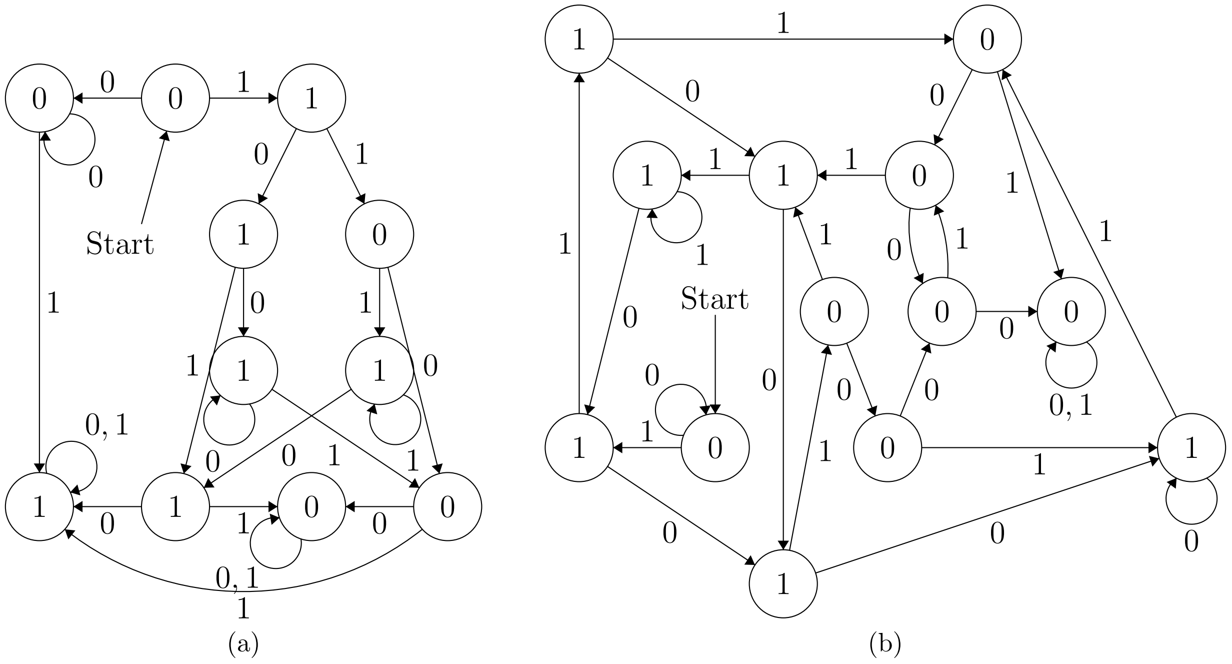


Figure 5.4: The 2-automata of (a) $\sigma_J^{\circ 3}$ and (b) $\sigma_{(1,5)}$.

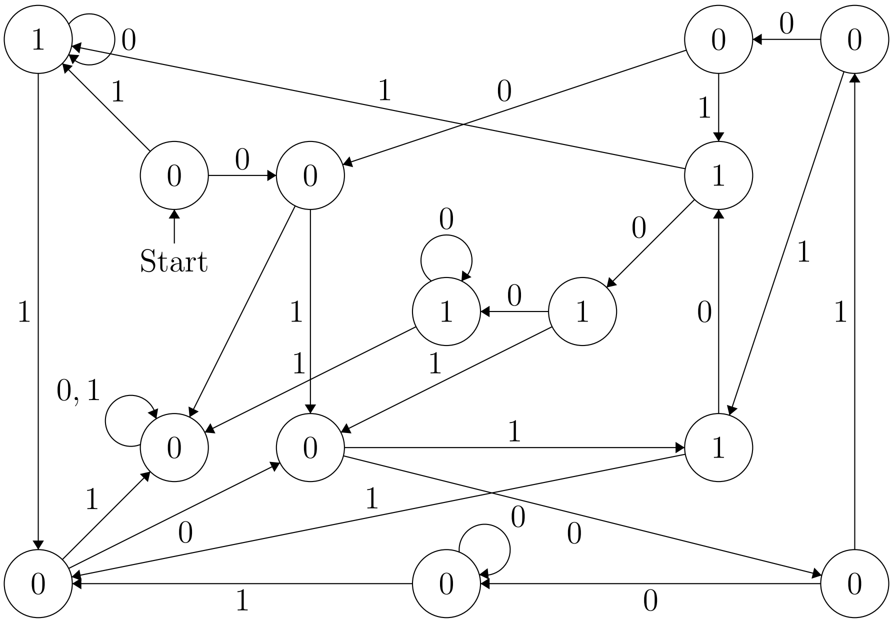


Figure 5.5: The 2-automaton of $\sigma_{V,1}$

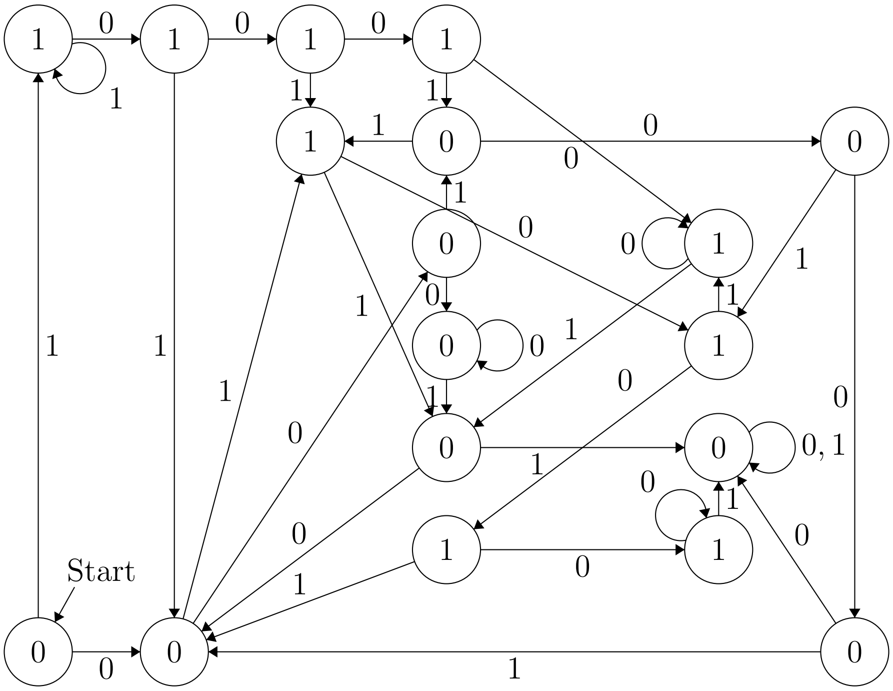


Figure 5.6: The 2-automaton of $\sigma_{V,2}$